Comparing the efficiency of normal form systems to represent Boolean functions

Miguel Couceiro, Pierre Mercuriali, Romain Péchoux

Comparing the efficiency of normal form systems to represent Boolean functions

Miguel Couceiro Pierre Mercuriali Romain Péchoux

Université de Lorraine, CNRS, Inria, LORIA
F 54000 Nancy, France
{miguel.couceiro, pierre.mercuriali, romain.pechoux}@loria.fr

Abstract

In this paper we compare various normal form representations of Boolean functions. We extend the study of [4], pertaining to the comparison of the asymptotic efficiency of representations that are produced by normal form systems (NFSs) that are factorizations of the clone Ω of all Boolean functions. We identify some properties, such as associativity, linearity, quasi-linearity and symmetry, that allow the efficiency of the corresponding NFSs to be compared in terms of the non-trivial connectives used. We illustrate these results by comparing well-known NFSs such as the DNF, CNF, Zhegalkin (Reed-Muller) polynomial (PNF) and Median (MNF) representations, thereby confirming the results of [4]. In particular, we show that the MNF is of equivalent complexity to, e.g., the Sheffer Normal Form (SNF), UNF and WNF (associated with 1 and 0-separating functions respectively) and thus that the latter are polynomially as efficient as any other NFS, and are strictly more efficient than the DNF, CNF, and Zhegalkin polynomial representations.

Keywords: Boolean function; Normal form; Median; Structural representation; Efficient representation.

1 Introduction

Efficient normal form representations of Boolean functions and optimal procedures for constructing them remain active topics in engineering and circuit design as well as in data mining and knowledge representation (see, e.g., [10, 15, 20, 21]).

Classical normal form representations of Boolean functions, such as disjunctive normal form (DNF), conjunctive normal form (CNF) and Zhegalkin polynomial (PNF) representations, can be thought of as factorizations of the clone (a class containing all projections and closed under composition) Ω of all Boolean functions into compositions of minimal clones. These facts were observed in [4] where the notion of class composition was investigated. The composition of two clones may or may not be a clone, and this fact motivated the study of such compositions of clones and that culminated in a complete classification of all pairs of clones accordingly. This classification showed that each clone can be factorized into "prime" clones and it led to all possible factorizations of the clone of all Boolean functions into minimal clones.

The latter classification had interesting consequences. For instance, it lead to a formalization of the intuitive notion of Normal Form System (NFS) defined as an irredundant factorization of Ω into prime clones and capable of expressing classical normal form systems (DNF, CNF and PNF) as well as the median normal form (MNF) that has the ternary median operator as its only non-trivial connective. Moreover, this framework provided a powerful formalism in which comparisons between the different normal form systems can be carried out rigorously.

The comparative study between the classical and the median normal form systems showed that the latter provides representations of lower complexity (measured as the minimum number of non-trivial connectives in the syntactic representation of a function) than the classical ones, while the classical NFSs remain pairwise incomparable. This fact asked for algorithmic procedures for producing optimal median representations of Boolean functions, which constitutes a topic of current research (see, e.g., [2, 7]). Recently, the problem of deciding whether a given median term is optimal was considered in [9] and shown to be in $\Sigma_2^{\mathbf{P}}$.

In this paper we go beyond the formalism proposed in [4] in four main aspects:

1. we slightly adapt the notion of NFS to express it syntactically, i.e. we consider NFSs to be sets of terms along with a semantic interpretation of them;
2. we relax the strict notion of NFS in [4] since we also consider factorizations into clones that are not necessarily prime (irreducible);
3. we consider arbitrary generators of clones (that play the role of the connectives in the NFSs), e.g., we take median connectives of arbitrary arity;
4. we develop a theory of NFSs that relies on structural properties of the functions connectives are interpreted as and systems, e.g., associativity and linearity, respectively.

As we will see, several noteworthy results stem from these considerations. For instance, the fairly intuitive idea that connectives interpreted as non-associative functions encode more "information" about functions is attested in two ways: first, they induce NFSs with a single non-trivial connective; second, the corresponding NFSs produce representations that are more efficient than those using connectives interpreted as associative functions. In fact, it will be shown that, under irredundancy, NFSs either use only connectives interpreted as associative functions or a single connective interpreted as a non-associative function. In the former case, more than one connective is necessary to guarantee the representation of Boolean functions.

The paper is organized as follows: in Section 2 we recall basic notions of clone theory, followed by notions on terms and their interpretations. We adapt the notion of NFS given in [4] to focus on terms rather than functions. We recall the notion of efficiency of representations of a Boolean function. In Section 3 we focus on NFSs that are generated by a single connective. We give conditions on those connectives that guarantee we can convert terms from one system to another efficiently (i.e., with no exponential explosion of their size). In particular, we show that the MNF system is polynomially as efficient as any other NFS generated by a single connective (Theorem 3). We also consider NFSs generated by more than one connective, such as the well-known disjunctive, conjunctive, polynomial, and dual polynomial normal forms. In fact, we establish a relation between non-associativity and quasi-Shefferness: an essentially n-ary connective, that is a connective the interpretation of which depends on all of its n variables, generates an NFS (connective that we call *quasi-Sheffer*) if and only if its interpretation is a non-associative function (Theorem 2). Moreover, we give a way to convert those terms that are compositions of connectives interpreted as associative functions into terms that are compositions of a single connective interpreted as a non-associative function. As a by-product we conclude that the MNF is polynomially as efficient as any other NFS (Theorem 4), thus extending Theorem 3. We also show that the MNF, SNF, UNF and WNF are equivalently efficient.

2 Preliminaries and notation

In this section we recall basic notions of clone theory and normal forms systems in the context of Boolean functions. For a more detailed presentation of clone theory and clone composition we refer to [12, 14] and [4], respectively.

2.1 Clone theory

Let $\mathbb{B} = \{0, 1\}$. The set $\mathbb{B}^n$ is the Boolean (distributive and complemented) lattice of 2^n elements under the component-wise ordering of tuples $\preceq$. The *complement* of a tuple $\mathbf{a} = (a_1, \ldots, a_n)$ is defined as $\overline{\mathbf{a}} = (1-a_1,\dots,1-a_n)$. We denote $\mathbf{0} = (0,\dots,0)$ and $\mathbf{1} = (1,\dots,1)$. For a function $f : \mathbb{B}^n \to \mathbb{B}$, the *dual* of f is defined as $f^d(\mathbf{a}) := \overline{f(\overline{\mathbf{a}})}$.

A *Boolean function* is a map $f : \mathbb{B}^n \to \mathbb{B}$, for some positive integer n called the *arity* of f. A *class* of functions is a subset $\mathcal{C} \subseteq \bigcup_{n\geq 1} \mathbb{B}^{\mathbb{B}^n}$. For a fixed arity n, there are n different *projection maps* $(a_1,\dots,a_n) \mapsto a_i, 1 \leq i \leq n$. We now give the definition of *essential variables* (see, e.g., [19, 22, 8]). Let f be a Boolean function of arity n. The ith argument of f is said to be *essential* in f, or that f *depends on* x_i, if there is a pair of tuples

$$((a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n),(a_1,\dots,a_{i-1},b,a_{i+1},\dots,a_n)) \quad \in (\mathbb{B}^n)^2,$$

differing only on the ith component, such that

$$f(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n) \quad \neq \quad f(a_1,\dots,a_{i-1},b,a_{i+1},\dots,a_n).$$

Such a pair is called a *witness of essentiality of* x_i *in* f. Two functions f and g are *equivalent* if each one can be obtained from the other by permutation of variables and addition or deletion of inessential variables (see, e.g., [6]). The number of essential variables in f is called the *essential arity* of f. The essential arity is an invariant for the equivalency of functions: if two functions are equivalent, then they have the same essential arity.

If f is an n-ary function and $g_1,\dots,g_n$ are all m-ary functions, then the *composition* $f(g_1,\dots,g_n)$ is the m-ary function given by

$$f(g_1,\dots,g_n)(a_1,\dots,a_m) = f(g_1(a_1,\dots,a_m),\dots,g_n(a_1,\dots,a_m)),$$

for all $(a_1,\dots,a_m) \in \mathbb{B}^m$. This notion extends naturally to classes of functions $\mathcal{I}$ and $\mathcal{J}$. The *composition of* $\mathcal{I}$ *with* $\mathcal{J}$, denoted $\mathcal{I} \circ \mathcal{J}$, is defined by

$$\mathcal{I} \circ \mathcal{J} := \{f(g_1,\dots,g_n) | n,m \geq 1,\ f\ n\text{-ary in } \mathcal{I},\ g_1,\dots,g_n\ m\text{-ary in } \mathcal{J}\}.$$

A *Boolean clone* is a class $\mathcal{C}$ of Boolean functions that contains all projections and satisfies $\mathcal{C} \circ \mathcal{C} \subseteq \mathcal{C}$ (i.e., it is closed under composition). Clones of Boolean functions constitute an algebraic lattice, which was completely described by E. Post (see [18]), where the meet is the intersection, the join of two clones is the smallest clone that contains their union, where the largest clone is $\Omega = \bigcup_{n\geq 1} \mathbb{B}^{\mathbb{B}^n}$ (all Boolean functions), and where the smallest clone is the clone of all projections I_c. These clones and the lattice are often called the Post Classes and the Post Lattice, respectively. The Post Lattice is reproduced in Figure 1. Let F be a set of Boolean functions. The clone *generated* by the set F, denoted $\mathcal{C}(F)$, is defined as follows:

$$\mathcal{C}(F) = \bigcap_{\mathcal{C} \text{ a clone, } F \subseteq \mathcal{C}} \mathcal{C}.$$

In other words, $\mathcal{C}(F)$ is the smallest clone that contains F.

2.2 Terms and their interpretation

We adopt the terminology of [3]. A *signature* Σ is a set of *function symbols*, also called *connectives.* Each $\alpha \in \Sigma$ is associated with a non-negative integer n called the *arity* of α and denoted by $\mathrm{ar}(\alpha)$. Let X be a countable set of variables. For a signature Σ such that $\Sigma \cap X = \emptyset$, the set $T(\Sigma, X)$ of all Σ-*terms over* X is recursively defined as follows:

- every variable in X is a term;
- the constants $\top$ and $\bot$ are terms;
- $\forall n \geq 0$, $\forall \alpha \in \Sigma$ such that $\mathrm{ar}(\alpha) = n$, $\forall t_1,\dots,t_n \in T(\Sigma, X)$, $\alpha(t_1,\dots,t_n)$ is a term.

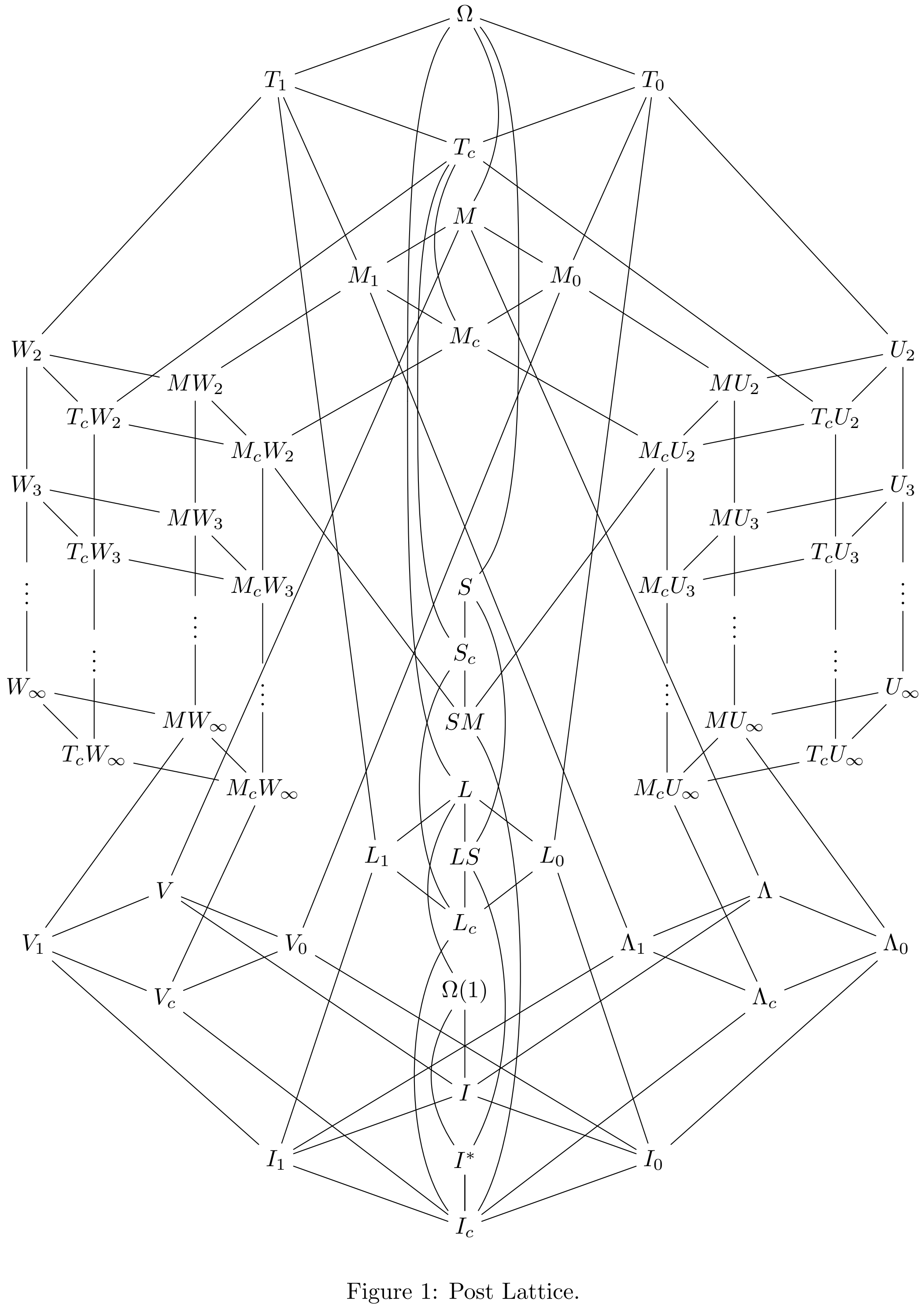

Figure 1: Post Lattice.

Remark 1. *In this paper, we are mainly interested in the number of connectives that occur in a term. As we fixed our set of variables X, we often make it implicit by writing $T(\Sigma)$ instead of $T(\Sigma, X)$.*

We introduce another notation to consider terms that follow a particular structure. To indicate the sequence $\alpha_1, \alpha_2, \dots, \alpha_n$, we adopt a string notation $\alpha_1\alpha_2\cdots\alpha_n$. Given two sequences ℓ and ℓ' viewed as strings, we indicate their *concatenation* by $\ell\ell'$. The empty string is denoted by ϵ.

Given a sequence of connectives $\alpha_1\cdots\alpha_n$, we denote by $T(\alpha_1\cdots\alpha_n, X)$ the set of terms $t_{\alpha_1\cdots\alpha_n}$ generated by the following grammar:

$$t_{\alpha\ell} ::= \alpha(\underbrace{t_{\alpha\ell}, \dots, t_{\alpha\ell}}_{\mathrm{ar}(\alpha)\text{ times}}) \mid t_\ell$$

$$t_\epsilon ::= x \mid \top \mid \bot \mid \neg x.$$

with $x \in X$ and with ℓ being a sequence of connectives.

For convenience, we also omit the set of variables X, using the notation $T(\alpha_1\cdots\alpha_n)$.

Example 1. *The set of terms $T(\wedge\vee)$ corresponding to the well-known disjunctive normal form is defined by the following grammar:*

$$\begin{aligned} t_{\wedge\vee} &::= \wedge(t_{\wedge\vee}, t_{\wedge\vee}) \mid t_\vee \\ t_\vee &::= \vee(t_\vee, t_\vee) \mid t_\epsilon \\ t_\epsilon &::= x \mid \top \mid \bot \mid \neg x. \end{aligned}$$

For instance $x\wedge(y\vee z) \in T(\wedge\vee)$, but $x\vee(y\wedge z) \notin T(\wedge\vee)$.

In this paper we will use letters $s, t, s', t', \dots$ to indicate terms, and we adopt the notation $\{t/x\}$ for the standard *substitution* of terms. For instance, $\mathrm{m}(x, y, z)\{\mathrm{m}(x, y, \top)/x\} = \mathrm{m}(\mathrm{m}(x, y, \top), y, z)$. Terms can be put in correspondence with functions by *interpreting* them as Boolean functions. We denote by $[\alpha]$ the interpretation of a connective α. The *interpretation of a term* is defined inductively on the structure of terms. The *interpretation of a set of terms $T(\alpha_1\cdots\alpha_n)$* is defined as $[T(\alpha_1\cdots\alpha_n)] = \{[t] : t \in T(\alpha_1\cdots\alpha_n)\}$. We say that the term t *represents* the Boolean function f if $[t]$ and f are equivalent. The constants $\top$ and $\bot$ are interpreted as 1 and 0, respectively.

Two terms t_1, t_2 are said to be *equivalent*, which we denote by $t_1 \equiv t_2$, if they are interpreted as the same function: $[t_1] = [t_2]$.

In this paper, we will consider the usual connectives

- $\wedge$ (that is interpreted as the binary conjunction),
- $\vee$ (that is interpreted as the binary disjunction),
- $\neg$ (that is interpreted as the negation),
- $\oplus$ (that is interpreted as the binary sum modulo 2),
- m (that is interpreted as the ternary median, or ternary majority, that can be defined as $\mathrm{m}(x, y, z) \equiv (x\wedge y)\vee(y\wedge z)\vee(z\wedge x))$,
- m_{2n+1} (that is interpreted $2n + 1$-ary median, also called majority; for instance, $\mathrm{m}_3 = \mathrm{m}$),
- $\uparrow$ (that is interpreted as the Sheffer stroke, defined as $\uparrow xy \equiv \neg(x\wedge y)$) and its dual $\downarrow$ (defined as $\downarrow xy \equiv \neg(x\vee y)$),

We will also use the connectives

- u (that can be defined as $\mathrm{u}(x, y, z) \equiv (x\vee y)\wedge z$), and

- w (that can be defined as $\mathrm{w}(x,y,z) \equiv (x\wedge y)\vee z$).

Example 2. *Both* $\mathrm{m}(x,y,z)$ *and* $(x\wedge y)\vee(y\wedge z)\vee(x\wedge z)$ *represent the same self-dual monotone function, which belongs to the clone* $\mathcal{C}(\mathrm{m})$ *of self-dual monotone functions, generated by* m*: we write* $[\mathrm{m}(x,y,z)] = [(x\wedge y)\vee(y\wedge z)\vee(x\wedge z)]$*. However,* $\mathrm{m}(x,y,z) \in T(\mathrm{m})$ *but* $(x\wedge y)\vee(y\wedge z)\vee(x\wedge z) \notin T(\mathrm{m})$*.*

Let $\mathcal{C}(\{\alpha_1,\ldots,\alpha_k\})$ denote the smallest clone containing $\{[\alpha_1],\ldots,[\alpha_k]\}$. The clone $\mathcal{C}(\{\alpha_1,\ldots,\alpha_k\})$ is said to be *generated by the set* $\{\alpha_1,\ldots,\alpha_k\}$. If a clone $\mathcal{C}$ is generated by a single connective α, which we note $\mathcal{C} = \mathcal{C}(\alpha)$ without brackets, then we say that α is a *generator* of $\mathcal{C}$.

2.3 Normal form systems

It is well-known that every Boolean function can be represented in disjunctive normal form. This fact can be restated as $\Omega = \mathcal{C}(\vee)\circ\mathcal{C}(\wedge)\circ\mathcal{C}(\neg)$, with $\mathcal{C}(\neg)$ denoting the clone of all literals (projections and negated projections). This illustrates the fact that we can express Ω as a factorization into clones, which was the basis for the notion of normal form systems proposed in [4]. We adapt this notion slightly to focus on terms instead of functions.

Definition 1 (Normal form systems)**.** *Let* $\alpha_1,\ldots,\alpha_n$ *be connectives and* $[\ ]$ *an interpretation. If* $[T(\alpha_1\cdots\alpha_n)] = \Omega$*, then the couple* $(T(\alpha_1\cdots\alpha_n),[\ ])$ *is called a* normal form system *or* NFS *for short. We may refer to the sequence of connectives* $\alpha_1\cdots\alpha_n$ *as the* generators *of the NFS.*

The NFS $(T(\alpha_1\cdots\alpha_n),[\ ])$ *is said to be* redundant*, if there exists an* $i \in \{1,\ldots,n\}$ *such that* $(T(\alpha_1\cdots\alpha_{i-1}\alpha_{i+1}\cdots\alpha_n),[\ ])$ *is an NFS. Otherwise, it is said to be* irredundant*.*

In this paper we only consider irredundant NFSs. For notational convenience, we will use the notation $T(\alpha_1\cdots\alpha_n)$ instead of the notation $T(\alpha_1\cdots\alpha_n,[\ ])$ throughout the paper, when the interpretation is clear from the context.

Remark 2. *The interpretations of the terms in* $T(\alpha_1\cdots\alpha_n)$ *are the functions in the clone* $\mathcal{C}(\alpha_1)\circ\cdots\circ\mathcal{C}(\alpha_n)\circ\Omega(1)$*:* $[T(\alpha_1\cdots\alpha_n)] = \mathcal{C}(\alpha_1)\circ\cdots\circ\mathcal{C}(\alpha_n)\circ\Omega(1)$*.*

Example 3. *The term* $\mathrm{m}(x,y,\wedge(z,t)) \in T(\mathrm{m}\wedge)$*, can already be expressed by the equivalent term of* $T(\mathrm{m})$*,* $\mathrm{m}(x,y,\mathrm{m}(z,t,\bot))$*. Hence* $T(\mathrm{m}\wedge)$ *is redundant, and we will consider instead the NFS* $T(\mathrm{m})$ *that is the Median NFS.*

A connective α is said to be *Sheffer* (resp. *quasi-Sheffer*) if $\Omega = \mathcal{C}(\alpha)$ (resp. $\Omega = \mathcal{C}(\alpha)\circ\Omega(1)$). Similarly, a clone $\mathcal{C}(\alpha)$ is said to be *complete* (resp. *quasi-complete*) if the connective α is Sheffer (resp. quasi-Sheffer). Clearly, every Sheffer connective is also quasi-Sheffer. The Sheffer stroke $\uparrow$ is Sheffer and thus quasi-Sheffer, whereas the median m is quasi-Sheffer ([4]) but not Sheffer. Indeed, since m is interpreted as a nondecreasing function, the terms it generates cannot be interpreted as strictly decreasing functions.

Here we introduce some NFSs that will be mentioned throughout the paper.

Example 4. *The well-known disjunctive, conjunctive, polynomial, and dual polynomial NFSs, denoted respectively by* $\mathbf{D},\mathbf{C},\mathbf{P}$*, and* $\mathbf{P}^d$*, are defined respectively by:*

- $\mathbf{D} = T(\vee\wedge)$*,*
- $\mathbf{C} = T(\wedge\vee)$*,*
- $\mathbf{P} = T(\oplus\wedge)$*, and*
- $\mathbf{P}^d = T(\oplus\vee)$*.*

We also consider the median, $2n+1$*-ary median, Sheffer,* 1*-separating and* 0*-separating NFSs:*

- $\mathbf{M} = T(\mathrm{m})$*,*

- $\mathbf{M}_{2n+1} = T(\mathrm{m}_{2n+1})$,
- $\mathbf{S} = T(\uparrow)$,
- $\mathbf{U} = T(\mathrm{u})$, *and*
- $\mathbf{W} = T(\mathrm{w})$.

2.4 Efficiency of representations in NFSs

Let t be a term and α a connective. We denote by $|t|_\alpha$ the number of occurrences of the symbol α in the term t. The *size* of a term t is denoted by $|t|$, and it is defined as the number of all connectives occurring in t:

$$|t| = \sum_{\alpha, \mathrm{ar}(\alpha)>1} |t|_\alpha.$$

For instance, $|x\wedge(y\vee\top)| = |x\wedge(y\vee\top)|_\wedge + |x\wedge(y\vee\top)|_\vee = 1 + 1 = 2$. We do not count litterals nor constants. This does not constitute a restriction since the number of literals and constants in a term is linear in the number of connectives in that term.

Definition 2 (**A**-complexity)**.** *Let* $\mathbf{A} = T(\alpha_1 \cdots \alpha_k)$ *be an NFS. For a function* $f \in \Omega$ *we define the* **A**-complexity *of* f*, denoted* $C_\mathbf{A}(f)$*, by*

$$C_\mathbf{A}(f) = \quad \min\{|t| : t \in T(\alpha_1 \cdots \alpha_k),\ t \textit{ represents } f\}.$$

Example 5. *Let* f *be the ternary majority function that returns* 1 *if at least two of its inputs are* 1*, and* 0 *otherwise. Then* $C_\mathbf{M}(f) = 1$ *because* $\mathrm{m}(x, y, z)$ *is the smallest term in* **M** *that represents* f*. However,* $C_\mathbf{D}(f) = 5$ *because* $(x\wedge y)\vee(y\wedge z)\vee(z\wedge x)$ *is the smallest term in* **D** *that represents* f*.*

Definition 3 (Efficiency)**.** *For two NFSs* **A** *and* **B***, we say that* **A** *is* polynomially as efficient as **B***, denoted* $\mathbf{A} \preceq \mathbf{B}$*, if there is a polynomial* P *with non-negative integer coefficients such that* $C_\mathbf{A}(f) \leq P(C_\mathbf{B}(f))$ *for all* $f \in \Omega$*.*

Remark that $\preceq$ *is a preorder on any set of NFSs, which is not total [4]. If* $\mathbf{A} \not\preceq \mathbf{B}$ *and* $\mathbf{B} \not\preceq \mathbf{A}$*, then* **A** *and* **B** *are said to be* provide representations of incomparable complexity *or, simply, that* **A** *and* **B** *are* incomparable*, and we write* $\mathbf{A}||\mathbf{B}$*. In the case when* $\mathbf{A} \preceq \mathbf{B}$ *but* $\mathbf{B} \not\preceq \mathbf{A}$*,* **A** *is said to be* polynomially more efficient than **B***, or* **A** *is said to* provide a representation of lower complexity than **B***, and we write* $\mathbf{A} \prec \mathbf{B}$*. In the case when* $\mathbf{A} \preceq \mathbf{B}$ *and* $\mathbf{B} \preceq \mathbf{A}$*,* **A** *and* **B** *are said to be* equivalent *or said to* provide representations of equivalent complexity*, and we write* $\mathbf{A} \sim \mathbf{B}$*. Thus defined,* $\sim$ *is an equivalence relation.*

Observe that a given clone $\mathcal{C}$ may have different (sets of) generators. For instance, the clone SM of self-dual monotone functions is generated by the ternary median m, as well as by any $2n+1$-ary median. This fact raises the question: which generator provides most efficient NFSs? We conjecture that any generator will induce NFSs of equivalent complexity.

Conjecture 1. *Consider the NFS* $\mathbf{A} = T(\alpha)$*. Let* β *be another generator of* **A** *and let* $\mathbf{B} = T(\beta)$*. Then* $\mathbf{A} \sim \mathbf{B}$*. In other words, the choice of generator has no effect on the efficiency of the representations produced.*

As we will see (Corollary 3), this conjecture holds for the set of generators of **M** made of the $2n+1$-ary medians ($n \geq 1$). However, this conjecture remains open in general. In view of Conjecture 1, we will mainly focus on NFSs generated by connectives of smallest arity. For instance, in the case of the clone SM and the NFS **M**, the connective chosen is the median, m. In the case of the clone M_cU_∞ and the NFS **U**, it is u.

We recall the following result on the comparison of well-known NFSs with the Median NFS **M**, obtained in [4].

Theorem 1 (Theorem 5, [4])**.** *For every* $\mathbf{A}, \mathbf{B} \in \{\mathbf{D}, \mathbf{C}, \mathbf{P}, \mathbf{P}^d\}$*, we have* $\mathbf{A}||\mathbf{B}$ *whenever* $\mathbf{A} \neq \mathbf{B}$*. Furthermore, for every* **B** *in* $\{\mathbf{D}, \mathbf{C}, \mathbf{P}, \mathbf{P}^d\}$*, we have that* $\mathbf{M} \prec \mathbf{B}$*.*

3 NFSs generated by a single connective

3.1 Relation between non-associativity and quasi-Shefferness

Let us recall the notion of an associative function (see, e.g., [1, 13, 17]). An n-ary function $f : X^n \to X$, $n \geq 2$, is said to be *associative* if it verifies:

$$\begin{aligned} & f(f(x_1, \ldots, x_n), x_{n+1}, \ldots, x_{2n-1}) = \ldots \\ =& f(x_1, \ldots, x_i, f(x_{i+1}, \ldots, x_{i+n}), x_{i+n+1}, \ldots, x_{2n-1}) = \ldots \\ =& f(x_1, \ldots, x_{n-1}, f(x_n, \ldots, x_{2n-1})), \quad \text{for} \quad i = 1, \ldots, n-2. \end{aligned}$$

It is easy to see that the median m is not interpreted as an associative function, because, for instance, $\mathrm{m}(\top, \bot, \mathrm{m}(\bot, \bot, \top)) \not\equiv \mathrm{m}(\top, \mathrm{m}(\bot, \bot, \bot), \top)$, whereas $\wedge, \vee$, and $\oplus$ are interpreted as associative functions. As we will see in Subsection 3.3, other NFSs generated by connectives interpreted as non-associative functions such as the Sheffer stroke $\uparrow$ provide representations of complexity equivalent to that of $\mathbf{M}$.

Theorem 2. *An connective whose interpretation only has essential variables is quasi-Sheffer if and only if its interpretation is non-associative.*

Proof. The composition of an associative function that only has essential variables with itself is still associative, so a connective whose interpretation only has essential variables cannot generate non-associative functions. Thus, connectives whose interpretation only has essential variables cannot be quasi-Sheffer.

We now prove that a connective that is not quasi-Sheffer must be interpreted as an associative function. In order to characterize connectives that are not quasi-Sheffer, we consider all the clones $\mathcal{C}(\alpha)$ with α of smallest arity such that $\mathcal{C}(\alpha) \circ \Omega(1) \neq \Omega$. To this end we refer to the complete description of the composition $\mathcal{C}_1 \circ \mathcal{C}_2$ of two clones of the Post Lattice, given in [4]. We reproduce some of their results in Table 1, only looking at clones that are generated by a single connective of minimal arity. In the case when the clone is generated by a connective and a constant, we only consider this connective in Table 1. For instance, the clone of all conjunctions and 1-preserving functions Λ_1 is generated by the functions $\{\wedge, 1\}$, but $T(\wedge) = T(\wedge\top)$: in the context of an NFS we do not need to consider $\top$ because the sets of terms are the same. Thus for the clones $\Lambda_1, \Lambda_0, \Lambda_c$ and Λ we will only consider the terms $\wedge$.

$\mathcal{C}$	generator (no constants)	$\mathcal{C} \circ \Omega(1)$
L, L_1, L_0, L_c, LS	$\oplus$	L
$\Lambda, \Lambda_1, \Lambda_0, \Lambda_c$	$\wedge$	$\Lambda \cup \Omega(1)$ (not a clone)
V, V_1, V_0, V_c	$\vee$	$V \cup \Omega(1)$ (not a clone)
$\Omega(1), I^*$	$\neg$	$\Omega(1)$
I, I_1, I_0, I_c	*id* (identity)	$\Omega(1)$

Table 1: Portion of the clone composition table such that $\mathcal{C} \circ \Omega(1) \neq \Omega$.

The connectives considered in Table 1 are all interpreted as associative functions. Thus, for each clone $\mathcal{C}(\alpha)$ that is not quasi-complete (i.e. such that α is not quasi-Sheffer), α is interpreted as an associative function.

□

Lemma 1. *Let $n > 1$ and let $\mathbf{A} = T(\alpha_1 \cdots \alpha_n)$ be an irredundant NFS. Then, every α_i is interpreted as an associative function.*

Proof. Suppose that α_i is interpreted as a non-associative function for some i. Then by Theorem 2, α_i is quasi-Sheffer, i.e., $\mathcal{C}(\alpha_i) \circ \Omega(1) = \Omega$. This means $\mathbf{A}$ is redundant. $\square$

Example 6. *The irredundant NFSs* $\mathbf{D}, \mathbf{C}, \mathbf{P}, \mathbf{P}^d$ *are generated respectively by* $\vee$ *and* $\wedge$, $\wedge$ *and* $\vee$, $\oplus$ *and* $\wedge$, *and* $\oplus$ *and* $\vee$. *In each case, the generators are interpreted as associative functions.*

3.2 Comparing NFSs generated by a single connective

In this section we compare the complexity of the representations produced by two NFSs following some conditions on the connectives that generate these normal forms. In particular, we investigate equivalences that can be used to convert terms from one system to the other. Consider, for instance, the equivalence $\mathrm{u}(x, y, z) \equiv \mathrm{m}(\mathrm{m}(x, \top, y), \bot, z)$. This equivalence allows us to convert terms of $\mathbf{U}$ into median terms of $\mathbf{M}$ without changing their complexity. We generalize such complexity-preserving conversions by giving conditions based on the number of variables occurring in the equivalences.

Definition 4 (Linear and quasi-linear NFS relation)**.** *Consider two NFSs* $\mathbf{A} = T(\alpha)$ *and* $\mathbf{B} = T(\beta)$. *Suppose that* β *has arity* n. *We say that:*

- *there exists a* linear NFS relation, *denoted* $\textsc{lin}(\mathbf{B}, \mathbf{A})$, *if:*

$$\exists t \in T(\alpha),\ \beta(x_1, \ldots, x_n) \equiv t \quad \textit{and} \quad \forall j \in \{1, \ldots, n\},\ |t|_{x_j} = 1;$$

- *there exists a* universal quasi-linear NFS relation, *denoted* $\forall\textsc{qlin}(\mathbf{B}, \mathbf{A})$, *if:*

$$\forall j \in \{1, \ldots, n\}, \exists t_j \in T(\alpha),\ \beta(x_1, \ldots, x_n) \equiv t_j \quad \textit{and} \quad |t_j|_{x_j} = 1;$$

- *there exists an* existential quasi-linear NFS relation, *denoted* $\exists\textsc{qlin}(\mathbf{B}, \mathbf{A})$, *if:*

$$\exists t \in T(\alpha),\ \beta(x_1, \ldots, x_n) \equiv t \quad \textit{and} \quad \exists j \in \{1, \ldots, n\},\ |t|_{x_j} = 1.$$

Remark 3. *The notion of linear NFS relation is somewhat related to the notion of* read-once (Boolean) functions*: a function* f *is called* read-once *if it can be represented by a term in which each variable appears at most once (see, e.g., [5, 11]). We call such a term a* read-once term. *If a linear relation holds, we can express the connective* β *with a term* ϕ *that is exactly a read-once term for* β *or, to be precise, for the function that* β *represents.*

Fact 1. *Note that for any two NFSs* $\mathbf{A}, \mathbf{B}$, *we have the following implications:*

$$\textsc{lin}(\mathbf{B}, \mathbf{A}) \Rightarrow \forall\textsc{qlin}(\mathbf{B}, \mathbf{A}) \Rightarrow \exists\textsc{qlin}(\mathbf{B}, \mathbf{A}).$$

As we will see in Example 7 the converse implications do not hold.

Example 7. $\textsc{lin}(\mathbf{U}, \mathbf{M})$ *holds. This follows from the linear equivalence* $\mathrm{u}(x, y, z) \equiv \mathrm{m}(\mathrm{m}(x, \top, y), \bot, z)$. *Indeed,* $|\,\mathrm{m}(\mathrm{m}(x, \top, y), \bot, z)|_x = 1$. *However,* $\textsc{lin}(\mathbf{M}, \mathbf{U})$ *does not hold (as an exhaustive search may show). The weaker property* $\exists\textsc{qlin}(\mathbf{M}, \mathbf{U})$ *can be inferred from* $\mathrm{m}(x, y, z) \equiv \mathrm{u}(\mathrm{u}(x, \bot, y), \mathrm{u}(x, y, z), \top)$ *since* $|\,\mathrm{u}(\mathrm{u}(x, \bot, y), \mathrm{u}(x, y, z), \top)|_z = 1$. *As we will see (Theorem 3),* $\forall\textsc{qlin}(T(\alpha), \mathbf{M})$ *always holds.*

Proposition 1. *Consider two NFSs* $\mathbf{A} = T(\alpha)$ *and* $\mathbf{B} = T(\beta)$. *If* $\textsc{lin}(\mathbf{B}, \mathbf{A})$ *holds, then* $\mathbf{A} \preceq \mathbf{B}$.

Proof. The proof is straightforward: converting a term of $\mathbf{B}$ into a term of $\mathbf{A}$ using a linear equivalence will increase its size at most polynomially (in fact, linearly). $\square$

Example 8. *We can convert ternary median terms into* 5*-ary median terms with the equivalence* $\mathrm{m}(x, y, z) \equiv \mathrm{m}_5(\bot, \top, x, y, z)$. *For instance, the term* $t_1 = \mathrm{m}(\mathrm{m}(x, y, z), u, v)$ *can be converted into the term* $t_2 = \mathrm{m}_5(\bot, \top, \mathrm{m}_5(\bot, \top, x, y, z), u, v)$. *Furthermore we have* $|t_1| = |t_2|$.

We now give an example that highlights the crucial problem in choosing the right conversion equivalences when $\textsc{lin}(\mathbf{B}, \mathbf{A})$ does not hold.

Example 9. *Consider the median term* $t_1 = \mathrm{m}(\mathrm{m}(x_1, x_2, x_3), x_4, x_5)$ *that we would like to convert into a formula of* $\mathbf{S}$ *using the equivalences*

$$\mathrm{m}(x, y, z) \equiv (y{\uparrow}z){\uparrow}(x{\uparrow}((y{\uparrow}\top){\uparrow}(z{\uparrow}\top))) \quad \textit{or} \tag{1}$$

$$\mathrm{m}(x, y, z) \equiv (x{\uparrow}z){\uparrow}(y{\uparrow}((x{\uparrow}\top){\uparrow}(z{\uparrow}\top))). \tag{2}$$

Equivalence (1) yields the term

$$t_2 = (x_4{\uparrow}x_5){\uparrow}((x_2{\uparrow}x_3){\uparrow}(x_1{\uparrow}((x_2{\uparrow}\top){\uparrow}(x_3{\uparrow}\top))){\uparrow}((x_4{\uparrow}\top){\uparrow}(x_5{\uparrow}\top)))$$

that has size 12*, but Equivalence (2) would yield the term*

$$\begin{aligned} t_3 = ((x_1{\uparrow}x_3){\uparrow}(x_2{\uparrow}((x_1{\uparrow}\top){\uparrow}(x_3{\uparrow}\top))){\uparrow}x_5){\uparrow}(x_4{\uparrow}(((x_1{\uparrow}x_3){\uparrow}(x_2{\uparrow}((x_1{\uparrow}\top){\uparrow}(x_3{\uparrow}\top)))\\ {\uparrow}\top){\uparrow}(x_5{\uparrow}\top)))\end{aligned}$$

that has size 18 *because the subterm* $\mathrm{m}(x, y, z)$*, once converted, is duplicated in* t_3*. Furthermore, it is possible to produce terms of* $T(\mathrm{m})$ *which, once converted into terms of* $T(\uparrow)$*, have polynomial sizes using Equivalence (1) but exponential sizes using Equivalence (2).*

In fact, Proposition 1 may be strengthened and stated in terms of universal quasi-linearity.

Proposition 2. *Consider two NFSs* $\mathbf{A} = T(\alpha)$ *and* $\mathbf{B} = T(\beta)$*. Suppose that* $\forall\textsc{qlin}(\mathbf{B}, \mathbf{A})$ *holds. Then,* $\mathbf{A} \preceq \mathbf{B}$.

Proof. Let $\mathbf{A}$ and $\mathbf{B}$ be NFSs satisfying the conditions of Proposition 2. Since $\forall\textsc{qlin}(\mathbf{B}, \mathbf{A})$ holds,

$$\forall i, \exists t_i \in T(\alpha),\, \beta(x_1, \ldots, x_n) \equiv t_i \quad \text{and} \quad |t_i|_{x_i} = 1.$$

To prove that $\mathbf{A} \preceq \mathbf{B}$, we give a recursive and efficient way of converting a term of $\mathbf{B}$ into an equivalent term of $\mathbf{A}$. We then prove that the size of the converted term is polynomial in the size of the term of $\mathbf{B}$.

Let s be a term of $\mathbf{B}$. Recall that for a sequence of n reals $(a_n)_n$, $\mathrm{argmax}_i\{a_1, \ldots, a_n\}$ is the smallest integer i such that for all n, $a_i \geq a_n$. We denote by $\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)$ ("Converted from $\mathbf{B}$ to $\mathbf{A}$") the term of $\mathbf{A}$ equivalent to s recursively defined as follows.

- If s is a literal or a constant, then $\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s) = s$;
- if $s = \beta(s_1, \ldots, s_n)$, then
  $\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s) = t_\ell\{\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_i)/x_i\}_i$, with
  $\ell = \mathrm{argmax}_i(|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_i)|)$.

The idea behind this recursive conversion process is to avoid repeating the biggest subterm that has already been converted; see Example 9. This is made possible because $\forall\textsc{qlin}(\mathbf{B}, \mathbf{A})$ holds. As we will see, this is sufficient to ensure an efficient conversion. The fact that $\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s) \equiv s$ is assured by the stability of interpretations by substitution.

Let $k = \max_i\{|t_i|_\alpha\}$ and $q = \max_{i,j}\{|t_i|_{x_j}\}$.

Let s be a term of $\mathbf{B}$ that represents a Boolean function f. We will prove that $|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| \leq k|s|^q$ by induction on the structure of terms of $\mathbf{B}$.

- Suppose that s is a literal or a constant, then $|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| = 0 = |s| = k|s|^q$.

- Suppose now that $s = \beta(s_1, s_2, \ldots, s_n)$ with $s_i \in T(\beta)$ for all i. Then,
$$|s|^q = (1 + |s_1| + |s_2| + \ldots |s_n|)^q. \tag{3}$$
Suppose without loss of generality that
$$|s_1| \geq |s_2| \geq \cdots \geq |s_{n-1}| \geq |s_n| \tag{4}$$
Let ℓ be the index defined as above: $\ell = \mathrm{argmax}_i(|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_i)|)$. By construction
$$\begin{aligned} |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| &= |t_\ell\{\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_i)/x_i\}_i| & (5)\\ &= |t_\ell|_\alpha + \sum_{1\leq j\leq n} |t_\ell|_{x_j} |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_j)| \\ &\leq k + q \sum_{1\leq j\leq n,\, j\neq\ell} |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_j)| + |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_\ell)|, & (6)\end{aligned}$$
because $|t_\ell|_{x_\ell} = 1$, $|t_i|_{x_j} \leq q$ and $|t_\ell|_\alpha \leq k$. Then,
$$|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| \leq k + q \sum_{2\leq j\leq n} |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_j)| + |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_1)| \tag{7}$$
because the difference between the right hand side of (7) and the right hand side of (6), $(q - 1)(|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_\ell)| - |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s_1)|)$, is positive by definition of ℓ. Thus,
$$|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| \leq k(1 + |s_1|^q + q\sum_{i=2}^{n} |s_i|^q) \qquad \text{by induction hypothesis.} \tag{8}$$
Remark now that by (4), $|s_{i+1}|^q \leq |s_{i+1}|^{q-1}|s_i|$. Recall also the multinomial formula:
$$(\sum_{i=1}^{n} X_i)^q = \sum_{k_1+k_2+\cdots+k_n=q} \frac{q!}{k_1!k_2!\ldots k_n!} \prod_{j=1}^{n} X_j^{k_j}$$
for all non-negative integers X_i. Thus, developing the right-hand side of (3) yields, in particular, the term
$$1 + |s_1|^q + q|s_1||s_2|^{q-1} + q|s_2||s_3|^{q-1} + \cdots + q|s_{n-1}||s_n|^{q-1}$$
which once multiplied by k is an upperbound for (8). Hence, $|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| \leq k|s|^q$, which concludes the induction.

Let $f \in \Omega$ be a Boolean function. Let s be a smallest term in $\mathbf{B}$ that represents f. Then we have $C_{\mathbf{B}}(f) = |s|$. We also have $C_{\mathbf{A}}(f) \leq |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)|$. Since $|\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| \leq k|s|^q$, we have:
$$C_{\mathbf{A}}(f) \leq |\textsc{conv}_{\mathbf{B}\to\mathbf{A}}(s)| \leq k|s|^q \leq k(C_{\mathbf{B}}(f))^q.$$
Thus, $\mathbf{A} \preceq \mathbf{B}$. □

In the case when α or β are interpreted as symmetric functions, Propositions 1 and 2 can be refined. Recall that a Boolean function f of arity n is said to be *symmetric* if for every permutation $\pi \in \mathfrak{S}_n$, we have $f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$.

Proposition 3. *Consider two NFSs $\mathbf{A} = T(\alpha)$ and $\mathbf{B} = T(\beta)$. Suppose that $\exists\textsc{qlin}(\mathbf{B}, \mathbf{A})$ holds and that α or β are interpreted as symmetric functions. Then, $\mathbf{A} \preceq \mathbf{B}$.*

Proof. Here, the symmetry of either α or β allows us to exhibit quasi-linear relations in every variable from the relation induced by the fact that $\exists\textsc{qlin}(\mathbf{B}, \mathbf{A})$ holds.

We can then apply Proposition 2. □

Example 10. *The median* m *is interpreted as a symmetric function. Since the equivalence* $\mathrm{m}(x, y, z) \equiv (y{\uparrow}z){\uparrow}(x{\uparrow}((y{\uparrow}\top){\uparrow}(z{\uparrow}\top)))$ *holds, we obtain the equivalences* $\mathrm{m}(x, y, z) \equiv (x{\uparrow}z){\uparrow}(y{\uparrow}((x{\uparrow}\top){\uparrow}(z{\uparrow}\top)))$ *and* $\mathrm{m}(x, y, z) \equiv (y{\uparrow}x){\uparrow}(z{\uparrow}((y{\uparrow}\top){\uparrow}(x{\uparrow}\top)))$. *All three equivalences together mean that* $\forall\textsc{qlin}(\mathbf{M}, \mathbf{S})$ *holds: in each one there is a variable that is not duplicated (x, y and z respectively).*

3.3 Applications: efficiency of the median normal form

In this section, we illustrate the usefulness of the notions of linear or quasi-linear NFS relations and the theorems they induce by comparing NFSs generated by a single connective. In particular, we show that $T(\mathrm{m})$ is polynomially as efficient as any other $T(\alpha)$. This is due to the median decomposition scheme (9) which allows us to apply Proposition 2 by showing that $\forall\textsc{qlin}(T(\alpha),\mathbf{M})$ holds. Let us now reproduce this median decomposition scheme ([16]) adapted to terms. Let α be a connective whose interpretation is a monotone Boolean function. Then

$$\alpha(x_1,\ldots,x_{\mathrm{ar}(\alpha)}) \equiv \mathrm{m}(\alpha(x_1,\ldots,x_{\mathrm{ar}(\alpha)})\{\bot/x_k\}, x_k, \alpha(x_1,\ldots,x_{\mathrm{ar}(\alpha)})\{\top/x_k\}) \tag{9}$$

for all $1 \leq k \leq \mathrm{ar}(\alpha)$.

Example 11. *Let α be the connective defined by $\alpha(x,y,z) \equiv (x\wedge y)\wedge z$. Remark that the function it describes is monotone. With the median decomposition scheme, we can obtain a conversion equivalence as follows. First, let us decompose according to the variable x, to obtain $\alpha(x,y,z) \equiv \mathrm{m}(\alpha(\bot,y,z),x,\alpha(\top,y,z))$. After decomposing the remaining subterms according to y and z, we obtain the conversion equivalence $\alpha(x,y,z) \equiv \mathrm{m}(\mathrm{m}(\mathrm{m}(\bot,z,\bot),y,\mathrm{m}(\bot,z,\bot)),x,\mathrm{m}(\mathrm{m}(\bot,z,\bot),y,\mathrm{m}(\bot,z,\top)))$[1] in which x only occurs once.*

Let us first state a few corollaries, the first of which compares the efficiency of representing terms using another quasi-Sheffer connective u.

Corollary 1. $\mathbf{U} \sim \mathbf{M}$*. Dually,* $\mathbf{W} \sim \mathbf{M}$*.*

Proof. Consider the following equivalences that establish the equivalence between $\mathbf{M}$ and $\mathbf{U}$:

$$\mathrm{u}(x,y,z) \equiv \mathrm{m}(\mathrm{m}(x,\top,y),\bot,z), \text{ and } \mathrm{m}(x,y,z) \equiv \mathrm{u}(\mathrm{u}(x,\bot,y),\mathrm{u}(x,y,z),\top). \tag{10}$$

Remark now that:

$$\forall w \in \{x,y,z\}, \quad |\,\mathrm{m}(\mathrm{m}(x,\top,y),\bot,z)|_w = 1 \quad \text{and} \quad |\,\mathrm{u}(\mathrm{u}(x,\bot,y),\mathrm{u}(x,y,z),\top)|_z = 1.$$

We can now apply Propositions 1 and 3, using the equivalences (10). A dual reasoning can be used to prove $\mathbf{W} \sim \mathbf{M}$. □

Corollary 2. $\mathbf{M} \sim \mathbf{S}$.

Proof. We have the following equivalences:

$$x{\uparrow}y \equiv \mathrm{m}(\neg x,\top,\neg y), \tag{11}$$

$$\mathrm{m}(x,y,z) \equiv (y{\uparrow}z){\uparrow}(x{\uparrow}((y{\uparrow}\top){\uparrow}(z{\uparrow}\top))). \tag{12}$$

Remark that in both equivalences (at least) one variable occurs only once in the right-hand side of the equivalence, i.e., $|\,\mathrm{m}(\neg x,\top,\neg y)|_x = |\,\mathrm{m}(\neg x,\top,\neg y)|_y = 1$, and $|(y{\uparrow}z){\uparrow}(x{\uparrow}((y{\uparrow}\top){\uparrow}(z{\uparrow}\top)))|_x = 1$. Thus, both $\exists\textsc{qlin}(\mathbf{S},\mathbf{M})$ and $\exists\textsc{qlin}(\mathbf{M},\mathbf{S})$ hold. Remark also that both m and ↑ are interpreted as symmetric functions. From Proposition 3 it then follows that $\mathbf{M} \preceq \mathbf{S}$ by (11) and that $\mathbf{S} \preceq \mathbf{M}$ by (12). In other words, $\mathbf{M} \sim \mathbf{S}$. □

The next theorem further motivates the study of $\mathbf{M}$: it is always polynomially as efficient as any other NFS $T(\beta)$.

Theorem 3. *Consider an NFS* $\mathbf{A} = T(\alpha)$*. Then* $\mathbf{M} \preceq \mathbf{A}$*.*

[1] Remark that the right hand side of this equivalence can be simplified further into $\alpha(x,y,z) \equiv \mathrm{m}(\bot,x,\mathrm{m}(\bot,y,z))$. See, e.g., [7, 9].

Proof. In [7] a simple algorithm was provided to construct a median representation of an arbitrary Boolean function, based on the median decomposition scheme. Algorithms in [7] essentially apply the decomposition scheme (9) iteratively to each variable, thus producing a median term, just as we have done in Example 11. We can control which variable is not duplicated by choosing on which variable the decomposition scheme is applied first. In Example 11, we chose the variable x first and obtained the equivalence $\alpha(x, y, z) \equiv \mathrm{m}(\mathrm{m}(\mathrm{m}(\bot, z, \bot), y, \mathrm{m}(\bot, z, \bot)), x, \mathrm{m}(\mathrm{m}(\bot, z, \bot), y, \mathrm{m}(\bot, z, \top)))$, but if we had chosen y, we would have obtained:

$$\alpha(x, y, z) \equiv \mathrm{m}(\mathrm{m}(\mathrm{m}(\bot, z, \bot), x, \mathrm{m}(\bot, z, \bot)), y, \mathrm{m}(\mathrm{m}(\bot, z, \bot), x, \mathrm{m}(\bot, z, \top))).$$

- If α is interpreted as a monotone function, then we can convert $\alpha(x_1, \ldots, x_n)$ directly into a term of $T(\mathrm{m})$. Remark that the first variable on which the decomposition scheme above has been applied thus only appears once in the final converted term. By applying the decomposition scheme on every variable x_i, we can produce n terms $t_i \in T(\mathrm{m})$ such that $t_i \equiv \alpha(x_1, \ldots, x_n)$ and $|t_i|_{x_i} = 1$. Thus, $\forall\textsc{qlin}(\mathbf{A}, \mathbf{M})$ holds, and from Proposition 2 it follows that $\mathbf{M} \preceq \mathbf{A}$.

- If α is not interpreted as a monotone function, then we follow similar steps as [7]. Given the connective α and its interpretation $[\alpha]$, define the function g_α as follows: for all $\mathbf{a} := (a_1, \ldots, a_{2n}) \in \mathbb{B}^{2n}$, let $\mathbf{b} := (a_1, \ldots, a_n)$, and let $\mathbf{c} := (a_{n+1}, \ldots, a_{2n})$, and let g_α be defined by:

$$g_\alpha(\mathbf{a}) = \begin{cases} 0 & \text{if } w(\mathbf{a}) < n, \\ 1 & \text{if } w(\mathbf{a}) > n, \\ [\beta](\mathbf{b}) & \text{if } \mathbf{b} = \overline{\mathbf{c}}, \\ 0 & \text{otherwise.} \end{cases}$$

  Here, $w(\mathbf{a})$ denotes the *Hamming weight* of $\mathbf{a}$, i.e., the number of 1 in $\mathbf{a}$. g_α is monotone, and for all $\mathbf{b} \in \mathbb{B}^n$, $[\alpha](\mathbf{b}) = g_\alpha(\mathbf{b}, \overline{\mathbf{b}})$; hence

$$[\alpha](x_1, \ldots, x_n) = g_\alpha(x_1, \ldots, x_n, \overline{x_1}, \ldots, \overline{x_n})$$

  for all $x_1, \ldots, x_n$. Since g_α is monotone, the median decomposition scheme applied to x_1 (for instance) yields the term $\mathrm{m}(x_1, t'_\alpha, t''_\alpha)$ whose interpretation is g_α, and where t'_α and t''_α are two median terms with no occurrence of x_1. Thus,

$$\beta(x_1, \ldots, x_n) \equiv \mathrm{m}(x_1, \tilde{t}'_\alpha, \tilde{t}''_\alpha)$$

  with $\tilde{t}'_\alpha$ and $\tilde{t}''_\alpha$ two median terms, containing variables $x_2, \ldots, x_{2n}$, and such that $x_{n+i} = \neg x_i$ for $1 \leq i \leq n$.

  Since this holds for every variable, from Proposition 2 we can convert efficiently the term of $\mathbf{A}$ into a term of $\mathbf{M}$. Once the conversion is done, we can replace the duplicated variables following $x_{n+i} = \neg x_i$ for $1 \leq i \leq n$. This process does not change the size of the median term, which stays polynomial in the size of the term of $\mathbf{A}$ on input. Thus $\mathbf{M} \preceq \mathbf{A}$.

□

Corollary 3. *For every $n \geq 1$, we have that $\mathbf{M} \sim \mathbf{M}_{2n+1}$.*

Proof. First observe that for all $n \geq 1$, m_{2n+1} is interpreted as a symmetric function. From Theorem 3 it follows that $\mathbf{M} \preceq \mathbf{M}_{2n+1}$, and from Proposition 3 and the equivalence

$$\mathrm{m}(x, y, z) \equiv \mathrm{m}_{2n+1}(z, \underbrace{x, \ldots, x}_{n \text{ times}}, \underbrace{y, \ldots y}_{n \text{ times}})$$

it follows that $\mathbf{M}_{2n+1} \preceq \mathbf{M}$. □

3.4 NFSs generated by two or more connectives

We can now establish a relation between the number of connectives that generate an NFS and the efficiency of the representations they produce. Interestingly, using more connectives does not entail more efficient representations.

We now investigate irredundant NFSs of the form $T(\beta\gamma)$, and compare them with NFSs generated by a single connective $T(\alpha)$. We show that representing a Boolean function using two connectives does not yield a significant gain of size over using a single one.

Lemma 2. *Consider two connectives α and β with interpretations $[\alpha]$ and $[\beta]$, respectively. If $[\beta]$ is associative and if there exists $t_\alpha \in T(\alpha)$ such that $\beta(x_1, \ldots, x_{\mathrm{ar}(\beta)}) \equiv t_\alpha$[2], then there exists a polynomial P with non-negative integer coefficients such that for all $s \in T(\beta)$, there exists $s' \in T(\alpha)$ such that $s \equiv s'$ and $|s'| \leq P(|s|)$.*

Proof. The proof is similar to the proof that $\mathbf{M} \preceq \mathbf{P}$ (or that $\mathbf{M} \preceq \mathbf{P}^d$) in [4]. We give a way to build s' explicitly by converting a term of $T(\beta)$ efficiently by *dichotomy*, in the sense that the associativity of the function $[\beta]$ represented by β allows us to regroup the terms in a composition of β efficiently. First, let $b = \mathrm{ar}(\beta) = \mathrm{ar}(t_\alpha)$, with $t_\alpha \in T(\alpha)$ such that $\beta(x_1, \ldots, x_b) \equiv t_\alpha$, and $k_i = |t_\alpha|_{x_i}$ for $i = 1, \ldots, b$. Now, let us consider the family $(s'_n)_{n\geq 1}$ of terms of $T(\alpha)$, s'_n, inductively defined as follows:

- $s'_1(x_1, \ldots, x_b) \equiv t_\alpha(x_1, \ldots, x_b)$;
- $s'_n(x_1, \ldots, x_{b^n}) \equiv t_\alpha(s'_{n-1}(x_1, \ldots, x_{b^{n-1}}), \ldots, s'_{n-1}(x_{b^n - b^{n-1}+1}, \ldots, x_{b^n}))$

We can prove by induction on n that

$$|s'_n(x_1, \ldots, x_{b^n})|_\alpha \equiv k\frac{(\sum_{i=1}^{b} k_i)^n - 1}{(\sum_{i=1}^{b} k_i) - 1}$$

with a positive integer k. Now, let us consider a term $s = \beta(\ldots\beta(x_1, \ldots, x_m)\ldots)$ of $T(\beta)$. Let n be the smallest positive integer such that $m \leq b^n$. There exist constants $c_1, \ldots, c_{b^n-m}$ such that $s \equiv \dot{s}$, with

$$\dot{s} = \beta(\beta(\ldots\beta(x_1, x_2, \ldots, x_{\mathrm{ar}(\beta)}), x_{\mathrm{ar}(\beta)+1}, \ldots, x_{2\,\mathrm{ar}(\beta)-1}), \ldots, c_1, \ldots, c_{b^n-m}).$$

Using the dichotomy method explained above, we can efficiently convert $\dot{s}$ into a term $\dot{s}' \in T(\alpha)$ such that

$$|\dot{s}'| \leq k\frac{(\sum_{i=1}^{b} k_i)^n - 1}{(\sum_{i=1}^{b} k_i) - 1}.$$

A lower bound for the size of s is $|s|>(b^{n-1} - 1)/(b - 1)$ because $|s| = (m - 1)/(b - 1)$, and since n is the smallest integer such that $m \leq b^n$, $b^{n-1} < m$, and so $(b^{n-1} - 1)/(b - 1) < |s|$. There exists a polynomial Q with non-negative integer coefficients such that $Q(b^n) \geq (\sum_{i=1}^{b} k_i)^n$, and so there exists a polynomial P with non-negative integer coefficients such that $|\dot{s}'| \leq P(|s|)$. □

Proposition 4. *Consider two irredundant NFSs $\mathbf{A} = T(\alpha)$ and $\mathbf{B} = T(\beta\gamma)$. Then $\mathbf{A} \preceq \mathbf{B}$.*

Proof. By Lemma 1 and Theorem 2 we can assume that both β and γ are interpreted as associative functions, while α is interpreted as a non-associative function because it is quasi-Sheffer.

Let $f \in \Omega$ be a Boolean function and let $t^f_{\mathbf{B}}$ be a smallest term of $\mathbf{B}$ that represents f. Recall that $|t|_\alpha$ indicates the number of occurrences of the symbol α in the term t. Let $n' = |t^f_{\mathbf{B}}|_\beta(\mathrm{ar}(\beta) - 1) + 1$, which is the number of leaves in the composition of $|t^f_{\mathbf{B}}|_\beta\ \beta's$. There exist $s_1, \ldots, s_{n'}$ terms of $T(\gamma)$ and $t'^f_{\mathbf{B}} \in T(\{\beta\}, \{x_1, \ldots, x_{n'}\})$ with $|t'^f_{\mathbf{B}}|_{x_i} = 1$ for all i, such that $t^f_{\mathbf{B}}\{s_i/x_i\} = t'^f_{\mathbf{B}}$. Since α is quasi-Sheffer,

[2] I.e. we can convert any term of β into a term of α.

there exists a term $t_\alpha \in T(\alpha)$ such that $t_\alpha \equiv \beta(x_1, \ldots, x_{\mathrm{ar}(\beta)})$. By Lemma 2, there exists a polynomial with integer coefficients P and a term $t'_{\mathbf{B},\alpha}$ of $T(\alpha)$ such that $t'^f_{\mathbf{B},\alpha} \equiv t'^f_{\mathbf{B}}$ and $|t'^f_{\mathbf{B},\alpha}| \leq P(|t'^f_{\mathbf{B}}|)$. Similarly, for every i, $s_i \in T(\gamma)$ and thus there exists a polynomial with non-negative integer coefficients Q_i and a term $t_{i,\alpha} \equiv s_i$ such that $|t_{i,\alpha}| \leq Q_i(|s_i|)$.

Hence we can give an upper bound for the number of leaves in the term $t'_{\mathbf{B},\alpha}$, namely,

$$\sum_i |t'^f_{\mathbf{B},\alpha}|_{x_i} \leq \mathrm{ar}(\alpha)p(n').$$

In particular, the term $t'^f_{\mathbf{B},\alpha}\{s_i/x_i\}$ is equivalent to $t^f_{\mathbf{B}}$, and

$$|t'^f_{\mathbf{B},\alpha}| \leq \mathrm{ar}(\alpha)P(n') \sum_{1\leq i\leq n'} |s_i|_\gamma. \tag{13}$$

Consider now the term $t^f_{\mathbf{A}}$ which is the smallest term of $\mathbf{A}$ equivalent to $t^f_{\mathbf{B}}$.

We have

$$\begin{aligned}
|t^f_{\mathbf{A}}| &\leq |t'^f_{\mathbf{B},\alpha}\{t_{i,\alpha}/x_i\}_i| \\
&\leq \mathrm{ar}(\alpha)P(n') \sum_{1\leq i\leq n'} |t_{i,\alpha}| \quad \text{by (13)} \\
&\leq \mathrm{ar}(\alpha)P(n') \sum_{1\leq i\leq n'} Q_i(|s_i|_\gamma) \\
&\leq \mathrm{ar}(\alpha)P(n') \sum_{1\leq i\leq n'} Q(|s_i|_\gamma) \quad \text{with } Q = \sum_i Q_i \\
&\leq \mathrm{ar}(\alpha)P(n')Q(\sum_{1\leq i\leq n'} |s_i|_\gamma) \\
&\leq \mathrm{ar}(\alpha)P(n' + \sum_{1\leq i\leq n'} |s_i|_\gamma)Q(n' + \sum_{1\leq i\leq n'} |s_i|_\gamma) \quad \text{by monotonicity of } Q.
\end{aligned}$$

Now, consider the size of $t^f_{\mathbf{B}}$ given by

$$|t^f_{\mathbf{B}}| = |t^f_{\mathbf{B}}|_\beta + |t^f_{\mathbf{B}}|_\gamma = n' + \sum_{1\leq i\leq n'} |s_i|_\gamma.$$

The inequality $|t^f_{\mathbf{A}}| \leq \mathrm{ar}(\alpha)P(|t^f_{\mathbf{B}}|)Q(|t^f_{\mathbf{B}}|)$ holds, which yields the desired comparison $\mathbf{A} \preceq \mathbf{B}$. □

Example 12. *Proposition 4 is another way to obtain some results involving* $\mathbf{M} = T(\mathrm{m})$ *of Theorem 1 from [4]: since* m *is interpreted as a non-associative function but* $\vee, \wedge$, *and* $\oplus$ *are interpreted as associative functions,* $T(\mathrm{m}) \preceq T(\vee\wedge)$, $T(\mathrm{m}) \preceq T(\wedge\vee)$, $T(\mathrm{m}) \preceq T(\oplus\wedge)$, *and* $T(\mathrm{m}) \preceq T(\oplus\vee)$.

Theorem 4. *Consider an NFS* $\mathbf{A} = T(\alpha)$ *and an irredundant NFS* $\mathbf{B} = T(\beta_1 \cdots \beta_n)$ *for* $n > 1$. *Then,* $\mathbf{A} \preceq \mathbf{B}$.

Proof. Straightforward from Proposition 4 and Lemma 1. □

We can now show that $\mathbf{M}$ is polynomially as efficient as any other NFS.

Corollary 4. *Let* $\mathbf{B}$ *be an NFS. Then* $\mathbf{M} \preceq \mathbf{B}$.

Proof. This result follows from Theorem 3 when $\mathbf{B} = T(\beta)$, and it follows from Theorem 4 (applied to $\mathbf{M}$) when $\mathbf{B} = T(\beta_1 \cdots \beta_n)$ with $n > 1$. □

4 Conclusion

In this paper we considered NFSs according to properties verified by the functions their generator(s) represent, such as associativity, symmetry or, simply, the number of generators involved, and we compared the efficiency of representing Boolean functions with terms produced by these NFSs. We provided sufficient conditions for NFSs to be polynomially as efficient as others. For instance, if we exhibit

1. a linear dependency between α and β,
2. a quasi linear dependency under the assumption that α or β is interpreted as a symmetric function,
3. or quasi-linear dependencies for all variables,

then we can conclude that $T(\alpha) \preceq T(\beta)$. Another noteworthy result tied to associativity is that systems generated by a single connective are always polynomially as efficient as any system generated by two or more connectives: allowing more connectives in the terms we use to represent functions does not yield more efficient representations. Moreover, we have shown that **M** is polynomially as efficient as any other NFS, which further motivates the study of the median connective and median algebra.

These results, together with those of [4], are summarized in Figure 2.

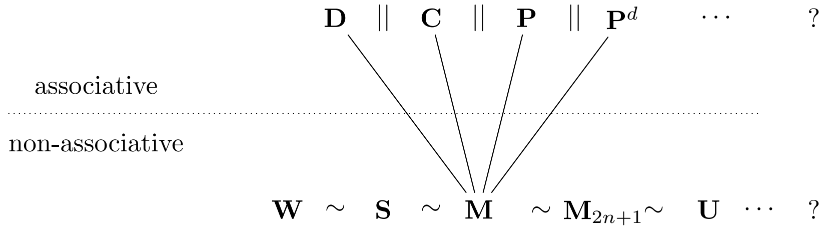
$\mathbf{D}$ || $\mathbf{C}$ || $\mathbf{P}$ || $\mathbf{P}^d$ $\cdots$?
associative
non-associative
$\mathbf{W}$ ~ $\mathbf{S}$ ~ $\mathbf{M}$ ~ $\mathbf{M}_{2n+1}$~ $\mathbf{U}$ $\cdots$?

Figure 2: Semilattice of some NFSs, ordered by $\preceq$, with a separation between NFSs based on connectives interpreted as associative functions and those, asymptotically more efficient, based on connectives interpreted as non-associative functions.

However, Figure 2 remains incomplete, and this gives rise to three challenging conjectures:

1. the strict relation between the top and the bottom levels, with the top corresponding to NFSs based on connectives interpreted as associative functions, and the bottom corresponding to NFSs based on a single connective interpreted as a non-associative function;
2. the incomparability relation between any two NFSs at the top level;
3. the equivalence between any two NFSs at the bottom level.

A positive answer to Conjecture 1 would be a valuable tool to settle the latter conjecture.

Acknowledgments

The authors wish to thank Emmanuel Hainry and Erkko Lehtonen for useful and fruitful discussions, and their insightful comments.

References

[1] János Aczél, Gary J. Erickson, and Yuxiang Zhai. The associativity equation re-revisited. In *Proc. of the AIP Conference*, volume 707, pages 195–203. AIP, 2004.

[2] Luca Amarú, Pierre-Emmanuel Gaillardon, and Giovanni De Micheli. Majority-inverter graph: A novel data-structure and algorithms for efficient logic optimization. In *Proc. of the 51st Annual Design Automation Conference*, pages 1–6. ACM, 2014.

[3] Franz Baader and Tobias Nipkow. *Term rewriting and all that.* Cambridge university press, 1999.

[4] Miguel Couceiro, Stephan Foldes, and Erkko Lehtonen. Composition of post classes and normal forms of Boolean functions. *Discrete Mathematics*, 306(24):3223–3243, 2006.

[5] Miguel Couceiro and Erkko Lehtonen. Galois theory for sets of operations closed under permutation, cylindrification and composition. *Algebra Universalis*, 67(3):25, 2012.

[6] Miguel Couceiro, Erkko Lehtonen, et al. A survey on the arity gap. In *Multiple-Valued Logic (ISMVL), 2011 41st IEEE International Symposium on*, pages 277–281. IEEE, 2011.

[7] Miguel Couceiro, Erkko Lehtonen, Jean-Luc Marichal, and Tamás Waldhauser. An algorithm for producing median formulas for Boolean functions. In *Proc. of the Reed Muller 2011 Workshop*, pages 49–54, 2011.

[8] Miguel Couceiro, Erkko Lehtonen, and Tamás Waldhauser. Decompositions of functions based on arity gap. *Discrete Mathematics*, 312(2):238–247, 2012.

[9] Miguel Couceiro, Pierre Mercuriali, Romain Pchoux, and Abdallah Saffidine. Median based calculus for lattice polynomials and monotone Boolean functions. To appear in *Proc. of the 47th IEEE International Symposium on Multiple-Valued Logic (ISMVL)*, may 2017.

[10] Yves Crama and Peter L. Hammer. *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, volume 2. Cambridge University Press, 2010.

[11] Yves Crama and Peter L. Hammer. *Boolean Functions: Theory, Algorithms, and Applications.* Cambridge University Press, 2011.

[12] Klaus Denecke and Shelly L. Wismath. *Universal Algebra and Coalgebra.* World Scientific, 2009.

[13] Wilhelm Dörnte. Untersuchungen ber einen verallgemeinerten gruppenbegriff. *Mathematische Zeitschrift*, 29:1–19, 1929.

[14] Dietlinde Lau. *Function Algebras on Finite Sets: Basic Course on Many-Valued Logic and Clone Theory.* Springer Science & Business Media, 2006.

[15] Heikki Mannila and Hannu Toivonen. Multiple uses of frequent sets and condensed representations: Extended abstract. In *Proc. of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD'96)*, pages 189–194, 1996.

[16] Jean-Luc Marichal. Weighted lattice polynomials. *Discrete Mathematics*, 309(4):814–820, 2009.

[17] Emil L. Post. Polyadic groups. *Transactions of the American Mathematical Society*, 48(2):208–350, 1940.

[18] Emil L. Post. *The Two-Valued Iterative Systems of Mathematical Logic*, volume 5, pages 1–122. Princeton, 1941.

[19] Arto Salomaa. On essential variables of functions, especially in the algebra of logic. *Annales Academiæ Scientiarum Fennicæ*, Series A I 339:11, 1963.

[20] Jilles Vreeken and Nikolaj Tatti. *Interesting Patterns*, pages 105–134. Springer International Publishing, Cham, 2014.

[21] Ingo Wegener. *Complexity Theory: Exploring the Limits of Efficient Algorithms*. Springer Science & Business Media, 2005.

[22] Ross Willard. Essential arities of term operations in finite algebras. *Discrete Mathematics*, 149(1-3):239–259, 1996.