TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY
Julkaisu 680 • Publication 680

Erkko Lehtonen

Operations on Finite Sets, Functional Composition, and Ordered Sets

Tampere 2007

Tampereen teknillinen yliopisto. Julkaisu 680
Tampere University of Technology. Publication 680

Erkko Lehtonen

Operations on Finite Sets, Functional Composition, and Ordered Sets

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Sähkötalo Building, Auditorium S2, at Tampere University of Technology, on the 8th of November 2007, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2007

ISBN 978-952-15-1849-2 (printed)
ISBN 978-952-15-1867-6 (PDF)
ISSN 1459-2045

Abstract

The unifying theme of this research work is functional composition. We study operations on a nonempty set A, i.e., mappings $f : A^n \to A$ for some $n \geq 1$, called the arity of f, an important particular case being that of Boolean functions when $A = \{0, 1\}$. The composition of an n-ary operation f with m-ary operations $g_1, \ldots, g_n$, denoted $f(g_1, \ldots, g_n)$, is the m-ary operation defined as $f(g_1, \ldots, g_n)(\mathbf{a}) = f(g_1(\mathbf{a}), \ldots, g_n(\mathbf{a}))$ for all $\mathbf{a} \in A^m$. A class of operations is a subset of the set $\mathcal{O}_A$ of all operations on A. The notion of composition can be extended to classes of operations: the composition of classes $\mathcal{C}_1$ and $\mathcal{C}_2$ is the class $\mathcal{C}_1 \circ \mathcal{C}_2$ that consists of all well-defined compositions of functions where the outer functions come from $\mathcal{C}_1$ and the inner functions from $\mathcal{C}_2$.

A clone on A is a class of operations on A that contains all projection maps and is closed under functional composition. The first part of this thesis is a study of compositions of the clones of Boolean functions. The clone of all Boolean functions can be decomposed in various ways into minimal clones, and we observe that such decompositions correspond to different normal form systems: the disjunctive normal form (DNF), conjunctive normal form (CNF), Zhegalkin polynomial, dual Zhegalkin polynomial, and so-called median normal form. These normal form systems are compared in terms of efficiency, and we establish that the median normal form system provides in a certain sense more efficient representations than the other four normal form systems mentioned above.

The second part of this thesis is a study of certain order relations on the set $\mathcal{O}_A$ of all operations on A. For a fixed class $\mathcal{C} \subseteq \mathcal{O}_A$, we say that f is a $\mathcal{C}$-subfunction of g, denoted $f \leq_{\mathcal{C}} g$, if f can be obtained by composing g from inside with operations from $\mathcal{C}$, i.e., $f = g(h_1, \ldots, h_n)$ for some $h_1, \ldots, h_n \in \mathcal{C}$ (or, equivalently, in terms of function class composition, $f \in \{g\} \circ \mathcal{C}$). We say that f and g are $\mathcal{C}$-equivalent, denoted $f \equiv_{\mathcal{C}} g$, if f and g are $\mathcal{C}$-subfunctions of each other. The $\mathcal{C}$-subfunction relation $\leq_{\mathcal{C}}$ is a quasiorder (a reflexive and transitive relation) on $\mathcal{O}_A$ if and only if the parametrizing class $\mathcal{C}$ is a clone, and if $\mathcal{C}$ is a clone, then $\equiv_{\mathcal{C}}$ is indeed an equivalence relation and $\leq_{\mathcal{C}}$ induces a partial order $\preccurlyeq_{\mathcal{C}}$ on the quotient $\mathcal{O}_A/{\equiv_{\mathcal{C}}}$.

The simplest example of $\mathcal{C}$-subfunctions is obtained when $\mathcal{C}$ is the smallest clone $\mathcal{I}_A$ of projections on A. Forming $\mathcal{I}_A$-subfunctions corresponds to simple manipulation of variables, namely addition and deletion of dummy variables, permutation of variables, and identification of variables. None of these operations increases the number of essential variables, and only variable identification may decrease this number. We study more carefully the effect of variable identification on the number of essential variables of operations on finite base sets.

We then study certain order-theoretical properties of various $\mathcal{C}$-subfunction partial orders defined by larger clones $\mathcal{C}$ on finite base sets A. We are mostly concerned about the descending chain condition and the existence of infinite antichains, and as it turns out, these properties on the defining clone $\mathcal{C}$. We focus on the following cases: the clones of monotone functions with respect to a partial order on A, the clones of linear functions on finite fields, the clones containing only essentially at most unary functions on A, and the clone of 1-separating Boolean functions.

Homomorphisms of labeled posets (or k-posets) are applied in our analysis of subfunction relations defined by clones of monotone functions. The third part of this thesis is a study of the homomorphicity order of finite k-posets on its own right. We establish that this order is a distributive lattice, and furthermore, it is universal in the sense that every countable poset can be embedded into it. This result implies universality of the subfunction partial orders defined by clones of monotone functions on finite sets of more than two elements. In this way, we also obtain a new proof for the well-known universality of the homomorphicity order of graphs.

Acknowledgements

...in the flowering of a mathematical talent social environment has an important part to play.

Jean Dieudonné (1906–1992)

This research was carried out at the Institute of Mathematics at Tampere University of Technology. I am grateful to Professor Seppo Pohjolainen, head of the institute, for the opportunity to complete this work. I also acknowledge financial support from Tampere Graduate School in Information Science and Engineering (TISE).

I am greatly indebted to my supervisor, Professor Stephan Foldes, for the constant professional and personal support. He introduced me to the very interesting topic of universal algebra and clones and he turned my attention to Boolean functions and operations on finite sets. He guided me patiently throughout this work, which often amounted to working very, very late in the night.

I wish to express my warmest and deepest gratitude to my friend and colleague Miguel Couceiro for many helpful discussions and fruitful collaboration, as well as for sharing moments of success, frustration and inspiration that I have experienced in my work.

I want to thank my colleagues at the Institute of Mathematics of Tampere University of Technology for the pleasant working environment. I also feel welcome to drop by at the University of Tampere every now and then to see fellow mathematicians.

Thank you to Jorge Almeida, Endre Boros, Jim Geelen, Ian Goulden, Lauri Hella, Sven Kosub, Kerkko Luosto, Hajime Machida, Gerasimos Meletiou, Jaroslav Nešetřil, Reinhard Pöschel, Maurice Pouzet, Sándor Radeleczki, Bruce Richmond, Ivo Rosenberg, Arto Salomaa, Victor Selivanov, Ágnes Szendrei, Jenő Szigeti, Ross Willard, and Igor Zverovich, as well as to a number of anonymous referees, for their advice, constructive remarks and suggestions regarding my research. I also had many interesting discussions with Bill Cunningham and the late Peter Hammer who hosted my North American sojourn in the summer of 2006.

Very special thanks go to Eszter K. Horváth, Léonard Kwuida, and Peter Mayr for their efforts of pre-examining this thesis. I am greatly honoured to have Professor Klaus Denecke of the University of Potsdam as my opponent.

I am glad to have a great little brother Teppo, as well as a handful of good old friends who occasionally remind me that there is life outside the realm of mathematics. Antti, Antti, Jan, Karoliina, Miikka, Mikko, Mikko, Tommi, you are great guys and gals! Greetings to all my friends around the globe.

Finally, I am thankful to my parents Leena and Tarmo for their endless and unconditional support and encouragement throughout my academic career and, for that matter, throughout all my life.

Tampere and Waterloo, August 2007
Erkko Lehtonen

Contents

List of publications

This thesis consists of the following five publications, or manuscripts accepted for publication.

1. M. Couceiro, S. Foldes, E. Lehtonen, Composition of Post classes and normal forms of Boolean functions, *Discrete Math.* **306** (2006) 3223–3243.

2. M. Couceiro, E. Lehtonen, On the effect of variable identification on the essential arity of functions on finite sets, *Int. J. Found. Comput. Sci.,* to appear.

3. E. Lehtonen, Descending chains and antichains of the unary, linear, and monotone subfunction relations, *Order* **23** (2006) 129–142.

4. E. Lehtonen, An infinite descending chain of Boolean subfunctions consisting of threshold functions, *Contributions to General Algebra* **17**, Proceedings of the Vienna Conference 2005 (AAA 70), Verlag Johannes Heyn, Klagenfurt, 2006, pp. 145–148.

5. E. Lehtonen, Labeled posets are universal, *European J. Combin.* (2007) doi:10.1016/j.ejc.2007.02.005.

Chapter 1

Introduction

1.1 Operations and clones

...that flower of modern mathematical thought—the notion of a function.

Thomas J. McCormack (1865–1932)

Let A be an arbitrary nonempty base set. In the current work, we are mostly concerned about the cases where the base set is finite. Since it is unimportant for our purposes what the elements of the base set are, we may assume that $A = \{0, 1, \dots, k-1\}$ for some $k \geq 1$, and we keep denoting the cardinality of A by k. In any case, the definitions and notions we present in this chapter apply to arbitrary nonempty base sets A, either finite or infinite.

An *operation* on A is a mapping $f : A^n \to A$ for some positive integer n, called the *arity* of f. Operations on the set $\{0, 1\}$ are an important particular case, and they are called *Boolean functions.* We denote by $\mathcal{O}_A$ the set of all operations on A, i.e., $\mathcal{O}_A = \bigcup_{n \geq 1} A^{A^n}$.

For $1 \leq i \leq n$, the i-th n-ary *projection* is the mapping $(a_1, \dots, a_n) \mapsto a_i$, and it is denoted by x_i^n, or simply by x_i when the arity is clear from the context. Denote by $\mathcal{I}_A$ the set of all projection operations on A.

If f is an n-ary operation and $g_1, \dots, g_n$ are m-ary operations on A, then the *composition* of f with $g_1, \dots, g_n$, denoted $f(g_1, \dots, g_n)$, is the m-ary operation defined by

$$f(g_1, \dots, g_n)(\mathbf{a}) = f(g_1(\mathbf{a}), \dots, g_n(\mathbf{a}))$$

for all $\mathbf{a} \in A^m$. Composition of operations satisfies the *superassociativity* condition:

$$f(g_1, \dots, g_n)(h_1, \dots, h_m) = f(g_1(h_1, \dots, h_m), \dots, g_n(h_1, \dots, h_m)).$$

Any subset $\mathcal{C} \subseteq \mathcal{O}_A$ is called a *class* of operations on A. The *n-ary part* of a class $\mathcal{C}$ is the set $\mathcal{C}^{(n)} = \{f \in \mathcal{C} : f \text{ is } n\text{-ary}\}$. The notion of functional composition is naturally extended to classes of operations on A by defining the *composition* of classes $\mathcal{C}$ and $\mathcal{K}$ as the class

$$\mathcal{C} \circ \mathcal{K} = \{f(g_1, \ldots, g_n) : f \in \mathcal{C}^{(n)},\ g_1, \ldots, g_n \in \mathcal{K}^{(m)} \text{ for some } m,\ n\}.$$

Class composition satisfies the following associative property of [3].

Lemma 1.1. *Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be classes of operations on A.*
(i) $(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} \subseteq \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})$.
(ii) If $\mathcal{B} \circ \mathcal{I}_A \subseteq \mathcal{B}$, then $(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} = \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})$.

A *clone* on A is a class $\mathcal{C}$ that contains all projections and is closed under composition (in other words $\mathcal{C} \circ \mathcal{C} \subseteq \mathcal{C}$, i.e., $f(g_1, \ldots, g_n) \in \mathcal{C}$ whenever $f, g_1, \ldots, g_n \in \mathcal{C}$ and the composition is defined). The smallest clone on A is the class $\mathcal{I}_A$ of all projections, and the largest clone on A is the class $\mathcal{O}_A$ of all operations on A. The clones on A constitute an inclusion-ordered lattice, where the lattice operations are the following: the meet of two clones is their set-theoretical intersection, and the join of two clones is the smallest clone containing their union. For a general account on clones, see [32].

A major ongoing research programme in universal algebra and multi-valued logic is the attempt to describe the structure of the lattice of clones on finite base sets. The clones of Boolean functions were completely described in the 1940s by E. Post [23], and these clones are called *Post classes* and the lattice of clones on the two-element set is called the *Post lattice.* Post's theorem has been reproved by several authors; see, e.g., [28, 33, 38] for recent shorter proofs.

While the Post lattice is countably infinite, it was shown by Yanov and Muchnik [36] and independently by Hulanicki and Świerczkowski [15] that the lattices of clones on finite base sets with more than two elements are uncountable. Little is known about the structure of clone lattices. Some parts of these lattices have been described, e.g., it is known that there are a finite number of atoms (minimal clones) and a finite number of coatoms (maximal clones). The maximal clones were characterized by Rosenberg [29]. In contrast, the classification of minimal clones is far from being completed (cf. [5, 25]).

1.2 $\mathcal{C}$-subfunction relations

One should always generalize.

Carl Jacobi (1804–1851)

For a fixed class $\mathcal{C} \subseteq \mathcal{O}_A$, we say that an operation f is a $\mathcal{C}$-*subfunction* (or a $\mathcal{C}$-*minor*) of an operation g, denoted $f \leq_{\mathcal{C}} g$, if $f \in \{g\} \circ \mathcal{C}$, i.e., $f = g(h_1, \dots, h_n)$ for some $h_1, \dots, h_n \in \mathcal{C}$. If f and g are $\mathcal{C}$-subfunctions of each other, we say that they are $\mathcal{C}$-*equivalent* and denote $f \equiv_{\mathcal{C}} g$. We have now defined families of binary relations $\leq_{\mathcal{C}}$ and $\equiv_{\mathcal{C}}$ on the set $\mathcal{O}_A$ of all operations on A, parametrized by the class $\mathcal{C}$.

If $\mathcal{C}$ and $\mathcal{K}$ are classes with $\mathcal{C} \subseteq \mathcal{K}$, then by definition $\leq_{\mathcal{C}} \subseteq \leq_{\mathcal{K}}$ and $\equiv_{\mathcal{C}} \subseteq \equiv_{\mathcal{K}}$. It is easy to verify that $\{x_i\} \circ \mathcal{C} = \mathcal{C}$ for any projection x_i and any class $\mathcal{C}$. Therefore, the relations $\leq_{\mathcal{C}}$ and $\leq_{\mathcal{K}}$ are distinct for $\mathcal{C} \neq \mathcal{K}$. However, the relations $\equiv_{\mathcal{C}}$ and $\equiv_{\mathcal{K}}$ may coincide even if $\mathcal{C} \neq \mathcal{K}$.

Lemma 1.2. *The relation $\leq_{\mathcal{C}}$ is reflexive if and only if $\mathcal{C}$ contains all projections.*

Proof. Let x_i be a projection of any arity. Assuming that $\leq_{\mathcal{C}}$ is reflexive, we have that $x_i \leq_{\mathcal{C}} x_i$, i.e., $x_i \in \{x_i\} \circ \mathcal{C} = \mathcal{C}$. Assume then that $\mathcal{C}$ contains all projections, and let f be n-ary. Then $f = f(x_1, \dots, x_n) \in \{f\} \circ \mathcal{C}$, i.e., $f \leq_{\mathcal{C}} f$, so $\leq_{\mathcal{C}}$ is reflexive. □

Lemma 1.3. *The relation $\leq_{\mathcal{C}}$ is transitive if and only if $\mathcal{C} \circ \mathcal{C} \subseteq \mathcal{C}$.*

Proof. Assume first that $\leq_{\mathcal{C}}$ is transitive, and let $f \in \mathcal{C} \circ \mathcal{C}$. Then $f = g(h_1, \dots, h_n)$ for some $g, h_1, \dots, h_n \in \mathcal{C}$, so $f \leq_{\mathcal{C}} g$. Since $g \in \mathcal{C}$, it is clear that $g \leq_{\mathcal{C}} x_1$. By the transitivity of $\leq_{\mathcal{C}}$, we have that $f \leq_{\mathcal{C}} x_1$, i.e., $f \in \{x_1\} \circ \mathcal{C} = \mathcal{C}$. Thus, $\mathcal{C} \circ \mathcal{C} \subseteq \mathcal{C}$.

Assume then that $\mathcal{C} \circ \mathcal{C} \subseteq \mathcal{C}$, and let $f \leq_{\mathcal{C}} g$ and $g \leq_{\mathcal{C}} h$. By Lemma 1.1,

$$f \in \{g\} \circ \mathcal{C} \subseteq (\{h\} \circ \mathcal{C}) \circ \mathcal{C} \subseteq h \circ (\mathcal{C} \circ \mathcal{C}) \subseteq h \circ \mathcal{C},$$

i.e., $f \leq_{\mathcal{C}} h$. We conclude that $\leq_{\mathcal{C}}$ is transitive. □

Thus, the $\mathcal{C}$-subfunction relation $\leq_{\mathcal{C}}$ is a quasiorder (a reflexive and transitive relation) on $\mathcal{O}_A$ if and only if the defining class $\mathcal{C}$ is a clone. If $\mathcal{C}$ is a clone, then $\equiv_{\mathcal{C}}$ is indeed an equivalence relation, and $\leq_{\mathcal{C}}$ induces a partial order $\preccurlyeq_{\mathcal{C}}$ on the quotient $\mathcal{O}_A/{\equiv_{\mathcal{C}}}$.

The motivating example of $\mathcal{C}$-subfunctions is given by the notion of taking *minors* of Boolean functions: addition of dummy variables, permutation of variables, identification of variables, deletion of inessential variables. These manipulations of variables are easily described in terms of composition of functions from inside with projections. In other words, a minor of f is an $\mathcal{I}_A$-subfunction of f, where $\mathcal{I}_A$ denotes the clone of all projections.

Denote the set of all $\mathcal{C}$-subfunctions of an operation f by $\mathrm{sub}_{\mathcal{C}}(f) = \{g \in \mathcal{O}_A : g \leq_{\mathcal{C}} f\}$. Let $\mathcal{Z} \subseteq \mathcal{O}_A$ be a class of operations on A. The set $\mathrm{FS}_{\mathcal{C}}(\mathcal{Z}) = \{f \in \mathcal{O}_A : \mathrm{sub}_{\mathcal{C}}(f) \cap \mathcal{Z} = \emptyset\}$ is the class of functions induced by the set $\mathcal{Z}$ of *forbidden* $\mathcal{C}$*-subfunctions.*

Classes of Boolean functions have been characterized in terms of forbidden minors [7, 34, 35, 38]. Pippenger [22] generalized the notion of minor for operations on arbitrary finite base sets and developed a Galois theory for classes of operations that are closed under taking minors. Ekin et al. [6] established that the equationally definable classes of Boolean functions are exactly the classes that are closed under taking variable identification minors, i.e., $\mathcal{I}_A$-subfunctions, and the structure of this quasiordering of Boolean functions was studied by Couceiro and Pouzet [4].

Thus, $\mathcal{C}$-subfunctions generalize the notion of minor. Such generalizations have appeared in many areas of mathematics. For example, $\mathcal{O}_A$-subfunctions were studied by Henno [11, 12] in the context of Green's equivalences and quasiorders on Menger systems. Equivalences of Boolean functions under actions of the general linear and affine groups of transformations over the two-element field were studied by Harrison [10], and they correspond to $\mathcal{C}$-equivalences defined by clones of linear Boolean functions, and such linear equivalences have found applications in coding theory and cryptography.

Representation of classes of Boolean functions by forbidden subfunctions played a key role in Zverovich's [38] proof of Post's theorem. Since any approach to Post's theorem is potentially a good candidate for proving Post-like results for large sublattices or sections of the lattice of clones on larger base sets, we believe that $\mathcal{C}$-subfuntions could prove useful in clone theory, as well as in other areas of mathematics.

In this thesis, we analyze various $\mathcal{C}$-subfunction partial orders defined by different clones on finite base sets. We are mostly concerned about the descending chain condition and the largest antichains contained in $\preccurlyeq_{\mathcal{C}}$, because of the fact that representation of classes of operations by minimal sets of forbidden subfunctions is possible if the corresponding $\mathcal{C}$-subfunction partial order satisfies the descending chain condition, and these minimal sets are guaranteed to be finite if the partial order contains only finite antichains. On the other hand, we pay little attention to the closely related ascending chain condition, because it plays no role in the forbidden subfunction characterization.

1.3 Labeled posets

> *The chief forms of beauty are order and symmetry and definiteness, which the mathematical sciences demonstrate in a special degree.*
>
> ARISTOTLE (384–322 B.C.)

For an integer $k \geq 1$, a *k-labeled partially ordered set* (a *k-poset*) is an object $((P, \leq), c)$ where $(P, \leq)$ is a partially ordered set and $c : P \to \{0, \ldots, k-1\}$ is a *labeling function.* If the underlying poset $(P, \leq)$ is a lattice, chain, forest, etc., then we speak of *k-lattices, k-chains, k-forests,* etc. An *alternating chain* is a k-chain $((P, \leq), c)$ satisfying the condition that $c(a) \neq c(b)$ whenever a covers b in $(P, \leq)$.

A *homomorphism* of a k-poset $((P, \leq), c)$ to a k-poset $((P', \leq'), c')$ is a mapping $h : P \to P'$ that preserves both the ordering and the labels, i.e., $h(a) \leq h(b)$ in P' whenever $a \leq b$ in P and $c = c' \circ h$. If there exists a homomorphism of $((P, \leq), c)$ to $((P', \leq'), c')$, we say that $((P, \leq), c)$ is *homomorphic* to $((P', \leq'), c')$. Two k-posets are *homomorphically equivalent* if they are homomorphic to each other. We define a quasiorder $\leq$ on the set of all k-posets by the existence of a homomorphism: $((P, \leq), c) \leq ((P', \leq'), c')$ if and only if there is a homomorphism of $((P, \leq), c)$ to $((P', \leq'), c')$.

Labeled posets are also known as *partially ordered multisets* (*pomsets*) or *partial words.* They have been used as a model for parallel processes (see Pratt [24]). Algebraic properties of labeled posets have been studied by Grabowski [9], Gischer [8], Bloom and Ésik [1], and Rensink [27]. To the best of our knowledge, the homomorphicity order of finite k-posets was first studied by Kosub and Wagner [18] in the context of Boolean hierarchies of partitions, followed by works by Kosub [16, 17] and Selivanov [31]. Kosub and Wagner were mostly concerned with k-lattices, whereas Selivanov studied k-forests. Kuske [20] and Kudinov and Selivanov [19] have studied the undecidability of the first-order theory of the homomorphicity order of k-posets.

In this thesis, homomorphisms of k-posets are applied in the analysis of subfunction relations defined by clones of monotone functions in Publication 3. We also study the homomorphicity order of finite k-posets on its own right in Publication 5.

Chapter 2

Author's contribution

The mathematician does not study pure mathematics because it is useful; he studies it because he delights in it and he delights in it because it is beautiful.

HENRI POINCARÉ (1854–1912)

We present our main results in this chapter. Each of the following sections summarizes one of the research papers that are part of this thesis, focusing on the key results that are essentially due to the author of this thesis. We also indicate some open problems and possible directions for further research.

2.1 Publication 1

Composition of Post classes and normal forms of Boolean functions

In this paper, we consider the class compositions of clones of Boolean functions. The composition $\mathcal{C}_1 \circ \mathcal{C}_2$ of Post classes $\mathcal{C}_1$ and $\mathcal{C}_2$ is either the join $\mathcal{C}_1 \vee \mathcal{C}_2$ in the Post lattice or it is not a clone. All pairs of clones $\mathcal{C}_1$, $\mathcal{C}_2$ are classified accordingly in a sequence of forty-two propositions that are proved using various techniques. In this way we obtain a class composition table (see Table 1 in Publication 1, summarized in Theorems 2 and 3 in Publication 1).

Having done this, we are able to decompose clones into proper subclones. In particular, we obtain all decompositions of the clone Ω of all Boolean functions into prime clones (clones that cannot be further decomposed into proper subclones). Such decompositions correspond to certain normal form representations of Boolean functions. These include the well-known disjunctive and conjunctive normal form representations (see [2]) and the Zhegalkin (or Reed–Muller) polynomial representation [21, 26, 37], as well as the dual of Zhegalkin polynomial representation, which relates to the Zhegalkin polynomial representation in a similar way as DNF relates to CNF. We also discover a new normal form of Boolean functions which we call the *median normal form,* corresponding to the decomposition $\Omega = SM \circ \Omega(1)$ where SM denotes the clone of self-dual monotone functions and $\Omega(1)$ denotes the clone of all essentially at most unary functions. This decomposition of Ω implies that every Boolean function can be represented as an iterated composition of the ternary majority function with itself, with possible substitution of negated variables or Boolean constants for variables.

Finally, we make a comparison between the efficiency of these normal form systems, and it turns out that the median normal form system is in a certain sense more efficient that the other above-mentioned normal form systems. Given a normal form system $\mathbf{N}$, the $\mathbf{N}$-complexity of f, denoted $C_{\mathbf{F}}(f)$, is the length of the shortest formula representing f that is in the given normal form. For two normal form systems $\mathbf{N}$ and $\mathbf{M}$, we say that $\mathbf{N}$ *provides polynomially more efficient representations* than $\mathbf{M}$, if there is a polynomial p such that for all functions f, $C_{\mathbf{N}}(f) \leq p(C_{\mathbf{M}}(f))$. We show that the DNF, CNF, Zhegalkin polynomial and dual Zhegalkin polynomial normal form systems provide representations of pairwise incomparable efficiency, while the median normal form system provides polynomially more efficient representations than the other four normal form systems.

For further research

In the current framework, we only considered factorizations of clones into two proper subclones at a time. Could we obtain different results, if we allowed factorizations into more than two subclones?

We described algorithms for converting DNF, CNF, and Zhegalkin polynomial representations of Boolean functions into median normal form. But the median normal form representations so obtained are not necessarily the shortest possible. Is there an efficient algorithm for finding the shortest median normal form representations?

2.2 Publication 2

On the effect of variable identification on the essential arity of functions on finite sets

Let $1 \leq i \leq n$. We say that the i-th variable is *essential* in an n-ary operation f on A if there are points $\mathbf{a}, \mathbf{b} \in A^n$ such that $a_j = b_j$ for all $j \neq i$, $a_i \neq b_i$, and $f(\mathbf{a}) \neq f(\mathbf{b})$. Otherwise the i-th variable is said to be *inessential* in f. The *essential arity* of f, denoted $\operatorname{ess} f$, is the number of essential variables in f.

This publication deals with $\mathcal{I}_A$-subfunctions, where $\mathcal{I}_A$ denotes the clone of projections on A. Substitution of projections to the arguments of a function amounts to permutation of variables, identification of variables, and addition and deletion of inessential variables. It is easy to see that the essential arity of a function cannot be increased by these operations, and the only one of these operations that may decrease essential arity is the identification of variables. Thus, if the i-th and j-th variables are essential in f, we call the function $f_{i \leftarrow j} = f(x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_n)$ a *variable identification minor* of f, obtained by identifying x_i with x_j. We define the *arity gap* of f, denoted $\operatorname{gap} f$, by

$$\operatorname{gap} f = \min_{i \neq j} (\operatorname{ess} f - \operatorname{ess} f_{i \leftarrow j})$$

where i and j range over the set of indices of the essential variables of f. It is clear that $\operatorname{gap} f \geq 1$ for every operation f with at least two essential variables.

In this paper, we consider the effect of variable identification on the number of essential variables of operations on finite sets. Salomaa [30] showed that every Boolean function f with $n \geq 3$ essential variables has a variable identification minor with at least $n - 2$ essential variables, i.e., $\operatorname{gap} f \leq 2$.

We generalize Salomaa's result to operations on arbitrary finite sets: every operation on a k-element set ($k \geq 2$) with at least $k + 1$ essential variables has a variable identification minor with at least $n - k$ essential variables, i.e., $\operatorname{gap} f \leq k$.

Furthermore, we strengthen Salomaa's theorem on Boolean functions by classifying all Boolean functions according to whether their arity gap is one or two. It turns out that there are only a few exceptional cases where the arity gap is two, namely the functions of the following forms:

- $x_{i_1} + x_{i_2} + \cdots + x_{i_n} + c$,
- $x_i x_j + x_i + c$,
- $x_i x_j + x_i x_k + x_j x_k + c$,

- $x_i x_j + x_i x_k + x_j x_k + x_i + x_j + c$,

where $c \in \{0, 1\}$. The proof makes good use of the Zhegalkin polynomial representations of Boolean functions and it separates cases in a clever and efficient way.

For further research

The problem on the arity gap of operations on finite sets with more than two elements still remains quite open. We do not know whether the bound $\operatorname{gap} f \leq k$ is sharp for operations on a k-element set with at least $k+1$ essential variables. In fact, we do not even know whether there exists any such operation of arity gap at least three.

Our framework could be generalized a little bit, and we could consider functions $f : A^n \to B$. We note that the codomain set plays no role in the proof of the upper bound $\operatorname{gap} f \leq k$ (insofar as there are at least two distinct elements in B to have essential variables to begin with), and hence this bound still holds for any sets A and B with $|A| = k$ and $|B| \geq 2$. Could we find, for example, a classification of pseudo-Boolean functions (mappings $\{0,1\}^n \to \mathbb{R}$) according to whether their arity gap is one or two?

2.3 Publication 3

Descending chains and antichains of the unary, linear, and monotone subfunction relations

In this paper, we study the $\mathcal{C}$-subfunction relations defined by certain clones on a finite base set A, namely

- the clone $\mathcal{O}_A$ of all operations on A,
- the clones containing only essentially at most unary operations on A,
- the clone of linear functions on a finite field A,
- the clone of monotone functions with respect to a partial order on A.

For the clone $\mathcal{O}_A$ of all operations on A, we observe that the notions of $\mathcal{O}_A$-equivalence and $\mathcal{O}_A$-subfunction correspond to Green's equivalence $\mathcal{L}$ and Green's quasiorder $\leq_L$ on full function systems, which were studied by Henno [11, 12]. A very simple criterion on function range holds here: $f \leq_{\mathcal{O}_A} g$ if and only if $\operatorname{Im} f \subseteq \operatorname{Im} g$. Hence, there are only a finite number of $\mathcal{O}_A$-equivalence classes on $\mathcal{O}_A$ for any finite base set A.

For the clones containing only essentially at most unary operations, the descending chain condition is seen to hold by an easy argument on the essential arity. As regards antichains, we generalize the infinite antichain construction presented by Pippenger [22, Proposition 3.4], and we rephrase his proof in the language of subfunctions and functional composition.

An operation f on a finite field A is *linear* if it has the form $f = a_1x_1 + \cdots + a_nx_n + c$ for some $a_1, \ldots, a_n, c \in A$. The class L of linear functions is a clone on A, and it is a maximal clone according to Rosenberg's classification [29]. We show that the L-subfunction partial order satisfies the descending chain condition but contains infinite antichains. The same applies to the partial order defined by the clone L_0 of linear functions with constant part $c = 0$.

For any partial order $\leq$ on A, the class $M_{\leq}$ of monotone functions with respect to $\leq$ is a clone; furthermore, by Rosenberg's classification [29], it is a maximal clone if $\leq$ has a greatest and a smallest element. Our analysis of $M_{\leq}$-subfunctions makes use of homomorphisms between k-posets.

We associate with each n-ary function f on A the k-poset $P(f, \leq) = ((A, \leq)^n, f)$. It holds that $f \leq_{M_{\leq}} g$ if and only if $P(f, \leq) \leq P(g, \leq)$. Then we make use of a family of k-lattices constructed by Kosub and Wagner [18] and show that if $|A| \geq 3$ and the partial order $\leq$ on A is not an antichain, then $\preccurlyeq_{M_{\leq}}$ contains both infinite descending chains and infinite antichains. In fact, using the results obtained in Publication 5, we conclude that $\preccurlyeq_{M_{\leq}}$ is a universal partial order in the sense that every countable poset can be embedded into it.

On the contrary, in the case that $|A| = 2$ the clone M of monotone Boolean functions defines a subfunction partial order that is far from being universal. Namely, $\preccurlyeq_M$ is isomorphic to the homomorphicity order of 2-lattices. As pointed out by Kosub and Wagner [18], every 2-lattice is homomorphically equivalent to its longest alternating chain. An alternating 2-chain is completely determined by its length and the label of its smallest element. Denoting by $C(n, b)$ the alternating 2-chain of length n with its smallest element labeled by b, we have that $C(n, b)$ is homomorphic to $C(n', b')$ if and only if either $n = n'$ and $b = b'$, or $n < n'$. It is then easy to see that this partial order is as presented in Figure 4 of Publication 5, having ascending chains but no infinite descending chains and its largest antichain contains just two elements.

For further research

It could be possible to obtain explicit description of the $\mathcal{C}$-subfunction relations for some clones $\mathcal{C}$. In particular, it could be possible to describe the $\mathcal{C}$-subfunction relations for all clones $\mathcal{C}$ of Boolean functions with the help of the Post Lattice. For clones on larger base sets, such a complete description would not be possible, since the lattice of clones on finite sets with more than two elements is largely unknown.

It would be interesting to determine the clones $\mathcal{C}$ on A for which the corresponding $\mathcal{C}$-equivalence relation on $\mathcal{O}_A$ has only a finite number on equivalence classes. Such clones form an order filter on the lattice of clones on A, so it would be sufficient to determine the minimal elements of the filter.

Yet another goal is to classify some natural classes of functions by forbidden $\mathcal{C}$-subfunctions.

2.4 Publication 4

An infinite descending chain of Boolean subfunctions consisting of threshold functions

For $b \in \{0, 1\}$, an n-ary Boolean function f is called *b-separating,* if there is an i ($1 \leq i \leq n$) such that for all $\mathbf{a} \in \{0, 1\}^n$ with $f(\mathbf{a}) = b$ we have that $a_i = b$. The classes U_∞ and W_∞ of 1-separating and 0-separating functions, repsectively, are clones on $\{0, 1\}$.

An n-ary function is a *threshold function,* if there are *weights* $w_1, \ldots, w_n \in \mathbb{R}$ and a *threshold* $w_0 \in \mathbb{R}$ such that $f(\mathbf{a}) = 1$ if and only if $\sum_{i=1}^{n} w_i a_i \geq w_0$.

In this paper, we show that there is an infinite descending chain of U_∞-subfunctions. The proof is based on an explicit construction consisting entirely of threshold functions. Analogous results hold for the clone W_∞ of 0-separating Boolean functions.

2.5 Publication 5

Labeled posets are universal

This paper is a study of the homomorphicity order of finite k-posets. For $k \geq 1$, denote by $\mathcal{P}_k$ the set of finite k-posets, and consider the quasiorder on $\mathcal{P}_k$ defined by the existence of a homomorphism: $(P, c) \leq (P', c')$ if and only if there is a homomorphism of (P, c) to (P', c'). Denote by $\mathcal{P}'_k$ the partial order induced by the homomorphicity quasiorder on the homomorphic equivalence classes of $\mathcal{P}_k$.

We show that the homomorphicity partial order $\mathcal{P}'_k$ of finite k-posets is a distributive lattice whenever $k \geq 2$. The lattice operations can be described as follows. The join of (the equivalence classes of) k-posets (P, c) and (P', c') is (the equivalence class of) the disjoint union $(P, c) \cup (P', c')$, and their meet is (the equivalence class of) the k-poset (Q, d), where

$$Q = \{(a, a') \in P \times P' : c(a) = c'(a')\},$$

$d(a, a') = c(a) = c'(a')$, and $(a, a') \leq (b, b')$ in Q if and only if $a \leq b$ in P and $a' \leq b'$ in P'.

The lattice $\mathcal{P}'_k$ has a smallest element, namely (the class of) the empty k-poset, but it has no maximal elements. Furthermore, it is a universal partial order in the sense that every countable poset can be embedded into it. We also prove that the homomorphicity order of finite k-lattices ($k \geq 3$) is universal. This is shown by constructing an embedding of a poset that is known to be universal in each of the homomorphicity orders mentioned above. Such a universal poset is provided by Hubička and Nešetřil [13, 14], and it comprises of a particular order relation imposed on the set of finite sets of finite sequences of natural numbers.

Theorem 4.6 and Proposition 6.1 of Publication 3 imply the universality of $M_{\leq}$-subfunction partial orders. As explained in the proof of Theorem 6.2 of Publication 3, for every k-lattice L, there exists an integer n and a function $f \colon A^n \to A$ such that the k-poset $P(f, \leq)$ corresponding to f and the partial order $\leq$ on A retracts to L and hence is homomorphically equivalent to it. It then follows from Proposition 6.1 of Publication 3 that the homomorphicity order $(\mathcal{L}'_k, \leq)$ of finite k-lattices embeds into the $M_{\leq}$-subfunction partial order $\preceq_{M_{\leq}}$. Since $(\mathcal{L}'_k, \leq)$ is shown to be universal when $k \geq 3$, it follows that the subfunction partial order defined by the clone $M_{\leq}$ of monotone functions with respect to a non-antichain partial order $\leq$ on a base set with at least three elements is also universal.

Furthermore, we represent k-posets by directed graphs and establish a categorical isomorphism between k-posets and their digraph representations. This yields a new proof for the well-known fact that the homomorphicity order of graphs is universal.

For further research

What are the cores, i.e., the k-posets that are not homomorphically equivalent to any k-poset of smaller cardinality?

Establish representations of graphs by k-posets such that there is a categorical isomorphism between graphs and their k-poset representations. Can this be done with k-lattices? What is the smallest k for which this is possible?

Complexity-theoretical aspects of k-posets could also be studied. Certain decision problems related to existence of homomorphisms between k-posets are likely to be NP-complete.

Bibliography

[1] S. L. Bloom, Z. Ésik, Free shuffle algebras in language varieties, *Theoret. Comput. Sci.* **163** (1996) 55–98.

[2] F. M. Brown, *Boolean Reasoning: The Logic of Boolean Equations,* second ed., Dover, Mineola, NY, 2003.

[3] M. Couceiro, S. Foldes, Function class composition, relational constraints and stability under compositions with clones, RUTCOR Research Report 22-2004, Rutgers University, 2004, `http://rutcor.rutgers.edu/~rrr/`.

[4] M. Couceiro, M. Pouzet, On a quasi-ordering on Boolean functions, manuscript, arXiv:math.CO/0601218.

[5] B. Csákány, Minimal clones—a minicourse, *Algebra Universalis* **54** (2005) 73–89.

[6] O. Ekin, S. Foldes, P. L. Hammer, L. Hellerstein, Equational characterizations of Boolean function classes, *Discrete Math.* **211** (2000) 27–51.

[7] A. Feigelson, L. Hellerstein, The forbidden projections of unate functions, *Discrete Appl. Math.* **77** (1997) 221–236.

[8] J. L. Gischer, The equational theory of pomsets, *Theoret. Comput. Sci.* **61** (1988) 199–224.

[9] J. Grabowski, On partial languages, *Ann. Soc. Math. Polon. Ser. IV Fund. Inform.* **4**(2) (1981) 427–498.

[10] M. A. Harrison, On the classification of Boolean functions by the general linear and affine groups, *J. Soc. Indust. Appl. Math.* **12**(2) (1964) 285–299.

[11] J. Henno, Green relations on Menger systems, *Tartu Riikliku Ülik. Toim.* **277** (1971) 37–46 (in Russian, with Estonian and English summaries).

[12] J. Henno, On idempotents and Green relations in the algebras of many-placed functions, *Ann. Acad. Sci. Fenn., Ser. A 1 Math.* **3** (1977) 169–178.

[13] J. Hubička, J. Nešetřil, Universal partial order represented by means of oriented trees and other simple graphs, *European J. Combin.* **26** (2005) 765–778.

[14] J. Hubička, J. Nešetřil, Finite paths are universal, *Order* **22** (2005) 21–40.

[15] A. Hulanicki, S. Świerczkowski, Number of algebras with a given set of elements, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **8** (1960) 283–284.

[16] S. Kosub, Complexity and Partition, Doctoral Dissertation, Bayerische Julius-Maximilians-Universität Würzburg, 2000. Available from the Ph.D. Theses Series of *Electronic Colloquium on Computational Complexity.*

[17] S. Kosub, NP-partitions over posets with an application to reducing the set of solutions on NP problems, *Theory Comput. Systems* **38** (2005) 83–113.

[18] S. Kosub, K. W. Wagner, The Boolean hierarchy of NP-partitions, in: H. Reichel, S. Tison (eds.), *STACS 2000, 17th Annual Symposium on Theoretical Aspects of Computer Science,* Lecture Notes in Comput. Sci. **1770**, Springer, Berlin, Heidelberg, 2000, pp. 157–168. (An expanded version is available as Technical Report TUM-U0209, Institut für Informatik, Technische Universität München, München, 2002.)

[19] O. V. Kudinov, V. L. Selivanov, Undecidability in the homomorphic quasiorder of finite labeled forests, in: A. Beckmann, U. Berger, B. Löwe, J. V. Tucker (eds.), *Logical Approaches to Computational Barriers,* Lecture Notes in Comput. Sci. **3988**, Springer, Berlin, Heidelberg, 2006, pp. 289–296.

[20] D. Kuske, Theories of orders on the set of words, *RAIRO—Inf. Theor. Appl.* **40** (2006) 53–74.

[21] D. E. Muller, Applications of Boolean algebra to switching circuit design and to error correction, *IRE Trans. Electron. Comput.* **3**(3) (1954) 6–12.

[22] N. Pippenger, Galois theory for minors of finite functions, *Discrete Math.* **254** (2002) 405–419.

[23] E. L. Post, *The Two-Valued Iterative Systems of Mathematical Logic,* Annals of Mathematical Studies **5**, Princeton University Press, Princeton, 1941.

[24] V. R. Pratt, Modelling concurrency with partial orders, *Internat. J. Parallel Programming* **15** (1987) 33–71.

[25] R. W. Quackenbush, A survey of minimal clones, *Aequationes Math.* **50** (1995) 3–16.

[26] I. S. Reed, A class of multiple-error-correcting codes and the decoding scheme, *IRE Trans. Inf. Theory* **4**(4) (1954) 38–49.

[27] A. Rensink, Algebra and theory of order-deterministic pomsets, *Notre Dame J. Formal Logic* **37** (1996) 283–320.

[28] M. Reschke, K. Denecke, Ein neuer Beweis für die Ergebnisse von E. L. Post über abgeschlossene Klassen Boolescher Funktionen, *J. Inf. Process. Cybern. EIK* **25**(7) (1989) 361–380.

[29] I. Rosenberg, Über die funktionale Vollständigkeit in den mehrwertigen Logiken, *Rozpravy Československé Akad. Věd, Řada Mat. Přírod. Věd* **80** (1970) 3–93.

[30] A. Salomaa, On essential variables of functions, especially in the algebra of logic, *Ann. Acad. Sci. Fenn. Ser. A I. Math.* **339** (1963) 3–11.

[31] V. L. Selivanov, Boolean hierarchies of partitions over a reducible base, *Algebra Logic* **43** (2004) 44–61. Translated from *Algebra Logika* **43** (2004) 77–109.

[32] Á. Szendrei, *Clones in Universal Algebra,* Séminaire de mathématiques supérieures **99**, Les Presses de l'Université de Montréal, Montréal, 1986.

[33] A. B. Ugol'nikov, Closed Post classes, *Izv. Vyssh. Uchebn. Zaved. Mat.* **32**(7) (1988) 79–88 (in Russian); translated in *Sov. Math.* **32**(7) (1988) 131–142.

[34] C. Wang, Boolean minors, *Discrete Math.* **141** (1991) 237–258.

[35] C. Wang, A. C. Williams, The threshold order of a Boolean function, *Discrete Appl. Math.* **31** (1991) 51–69.

[36] Yu. I. Yanov, A. A. Muchnik, On existence of k-valued closed classes without a finite basis, *Dokl. Akad. Nauk SSSR* **127** (1959) 44–46 (in Russian).

[37] I. I. Zhegalkin, On the calculation of propositions in symbolic logic, *Mat. Sb.* **34** (1927) 9–28 (in Russian).

[38] I. Zverovich, Characterizations of closed classes of Boolean functions in terms of forbidden subfunctions and Post classes, *Discrete Appl. Math.* **149** (2005) 200–218.

Errata

If a mistake is not a stepping stone, it is a mistake.

Eli Siegel (1902–1978)

Publication 1. "Composition of Post classes and normal forms of Boolean functions"

- Page 3224, line 4: Remove ")" after "[16,18,22]".

Publication 2. "On the effect of variable identification on the essential arity of functions on finite sets"

- Page 5, line 12 (Equation (11)): "$\mathfrak{p}_{l \leftarrow i} = x_i x_j + x_i x_k + x_j x_k + x_i \mathfrak{a}' + \mathfrak{a}''$" should read "$\mathfrak{p}_{l \leftarrow i} = x_i x_j + x_i x_k + x_j x_k + x_i \mathfrak{a}_i + x_j \mathfrak{a}_j + x_k \mathfrak{a}_k + x_i \mathfrak{a}' + \mathfrak{a}''$" (three terms are missing).
- Page 5, line 13: Remove "because no terms cancel,".
- Page 6, lines 5–11: Subcase 1.2.3 is empty and can be removed.
- Page 7, line −19: Insert "or $x_i + x_k$" after "$x_i x_k$".
- Page 8, line −11: "$\mathfrak{r}_{m \leftarrow l}$" should read "$\mathfrak{p}_{m \leftarrow l}$".
- Page 9, line −6 (Equation (34)): "x_l" should read "x_t".

Publication 3. "Descending chains and antichains of the unary, linear, and monotone subfunction relations"

- Page 135, line 12: "h_m" should read "h_n" (2 occurrences).
- Page 138, line −21 (second paragraph of the proof of Proposition 5.9): "$\mathbf{v}$" should read "$\mathbf{d}$".

- Page 139, lines 16–17: The passage

  > If both S_1 and S_2 are nonempty, let $r \in S_1$, $s \in S_2$ and $t \notin \operatorname{Im}\sigma$. Then $v_{\sigma(r)}$, $v_{\sigma(s)}$ and v_t are distinct elements of A, and so $\mathbf{v} \notin \{0,1\}^m$.

  should be replaced by the following:

  > For a nonnegative integer q, denote by $q \cdot 1$ the sum
  >
  > $$\underbrace{1 + \cdots + 1}_{q \text{ terms}}.$$
  >
  > If both S_1 and S_2 are nonempty, then $v_{\sigma(i)} = 1 + |S_2| \cdot 1 = (|S_2| + 1) \cdot 1$ for all $i \in S_1$, $v_{\sigma(i)} = (|S_2| - 1) \cdot 1$ for all $i \in S_2$, and $v_j = |S_2| \cdot 1$ for $j = \sigma(1)$ and for all $j \notin \operatorname{Im}\sigma$. If A has characteristic greater than 2, then $(|S_2| - 1) \cdot 1$, $|S_2| \cdot 1$, $(|S_2| + 1) \cdot 1$ are pairwise distinct elements of A, and so $\mathbf{v} \notin \{0,1\}^m$. If A has characteristic 2, then $(|S_2| - 1) \cdot 1 = (|S_2| + 1) \cdot 1 \neq |S_2| \cdot 1$, and hence $w(\mathbf{v})$ equals either $n - 1$ or $m - n + 1$ and so $w(\mathbf{v}) \notin \{1, m - 1\}$.

Publication 5. "Labeled posets are universal"

- Page 2, line -1: "$\mathcal{L}/{\equiv}$" should read "$\mathcal{L}_k/{\equiv}$".

Publication 1

M. Couceiro, S. Foldes, E. Lehtonen, Composition of Post classes and normal forms of Boolean functions, *Discrete Math.* **306** (2006) 3223–3243.

Publication 2

M. Couceiro, E. Lehtonen, On the effect of variable identification on the essential arity of functions on finite sets, *Int. J. Found. Comput. Sci.,* to appear.

Publication 3

E. Lehtonen, Descending chains and antichains of the unary, linear, and monotone subfunction relations, *Order* **23** (2006) 129–142.

Publication 4

E. Lehtonen, An infinite descending chain of Boolean subfunctions consisting of threshold functions, *Contributions to General Algebra* **17**, Proceedings of the Vienna Conference 2005 (AAA 70), Verlag Johannes Heyn, Klagenfurt, 2006, pp. 145–148.

Publication 5

E. Lehtonen, Labeled posets are universal, *European J. Combin.* (2007)
doi:10.1016/j.ejc.2007.02.005.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O. Box 527
FIN-33101 Tampere, Finland