Infinitary Axiomatization of the Equational Theory of Context-Free Languages

Niels Bjørn Bugge Grathwohl
Department of Computer Science (DIKU)
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen, Denmark
bugge@diku.dk

Fritz Henglein
Department of Computer Science (DIKU)
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen, Denmark
henglein@diku.dk

Dexter Kozen
Department of Computer Science
Cornell University
Ithaca, NY 14853-7501, USA
kozen@cs.cornell.edu

We give a natural complete infinitary axiomatization of the equational theory of the context-free languages, answering a question of Leiß (1992).

1 Introduction

Algebraic reasoning about programming language constructs has been a popular research topic for many years. At the propositional level, the theory of flowchart programs and linear recursion are well handled by such systems as Kleene algebra and iteration theories, systems that characterize the equational theory of the regular sets. To handle more general forms of recursion including procedures with recursive calls, one must extend to the context-free languages, and here the situation is less well understood. One reason for this is that, unlike the equational theory of the regular sets, the equational theory of the context-free languages is not recursively enumerable. This has led some researchers to declare its complete axiomatization an insurmountable task [13].

Whereas linear recursion can be characterized with the star operator * of Kleene algebra or the dagger operation † of iteration theories, the theory of context-free languages requires a more general fixpoint operator μ. The characterization of the context-free languages as least solutions of algebraic inequalities involving μ goes back to a 1971 paper of Gruska [7]. More recently, several researchers have given equational axioms for semirings with μ and have developed fragments of the equational theory of context-free languages [3, 5, 6, 8, 9, 13].

In this paper we consider another class of models satisfying a condition called *μ-continuity* analogous to the star-continuity condition of Kleene algebra:

$$a(\mu x.p)b = \sum_{n\geq 0} a(nx.p)b,$$

where the summation symbol denotes supremum with respect to the natural order in the semiring, and

$$0x.p = 0 \qquad\qquad (n+1)x.p = p[x/nx.p].$$

This infinitary axiom combines the assertions that $\mu x.p$ is the supremum of its finite approximants $nx.p$ and that multiplication in the semiring is continuous with respect to these suprema. Analogous to a

D. Baelde and A. Carayol (Eds.): Fixed Points
in Computer Science 2013 (FICS 2013)
EPTCS 126, 2013, pp. 44–55, doi:10.4204/EPTCS.126.4

similar result for star-continuous Kleene algebra, we show that all context-free languages over a μ-continuous idempotent semiring have suprema. Our main result is that the μ-continuity condition, along with the axioms of idempotent semirings, completely axiomatize the equational theory of the context-free languages. This is the first completeness result for the equational theory of the context-free languages, answering a question of Leiß [13].

1.1 Related Work

Courcelle [3] investigates *regular systems*, finite systems of fixpoint equations over first-order terms over a ranked alphabet with a designated symbol $+$ denoting set union, thereby restricting algebras to power set algebras. He stages their interpretation by first interpreting recursion over first-order terms as infinite trees, essentially as the final object in the corresponding coalgebra, then interpreting the signature symbols in ω-complete algebras. He provides soundness and completeness for transforming regular systems that preserve all solutions and soundness, but not completeness for preserving their least solutions. Courcelle's approach is syntactic since it employs unfolding of terms in fixpoint equations.

Leiß [13] investigates three classes of idempotent semirings with a syntactic least fixpoint operator μ. The three classes are called KAF, KAR, and KAG in increasing order of specificity. All these classes are assumed to satisfy the fundamental *Park axioms*

$$p[x/\mu x.p] \leq \mu x.p \qquad\qquad p \leq x \;\Rightarrow\; \mu x.p \leq x,$$

which say that $\mu x.p$ is the least solution of the inequality $p \leq x$. The classes KAR and KAG further assume

$$\mu x.(b+ax) = \mu x.(1+xa)\cdot b \qquad\qquad \mu x.(b+xa) = b\cdot \mu x.(1+ax)$$

and

$$\mu x.(s+rx) = \mu x.(\mu y.(1+yr)\cdot s) \qquad\qquad \mu x.(s+xr) = \mu x.(s\cdot \mu y.(1+ry)),$$

respectively. These axioms can be viewed as imposing continuity properties of the semiring operators with respect to μ. All standard interpretations, including the context-free languages over an alphabet X, are continuous and satisfy the KAG axioms. Ésik and Leiß [5, 6] show that conversion to Greibach normal form can be performed purely algebraically under these assumptions.

Ésik and Kuich [4] introduce *continuous semirings*, which are required to have suprema for all directed sets, and they employ domain theory to solve polynomial fixpoint equations. Idempotent continuous semirings are μ-continuous Chomsky algebras as defined here, but not conversely. As we shall prove, the family of context-free languages over any alphabet constitutes a μ-continuous Chomsky algebra. It is not a continuous semiring, however, since the union of context-free languages is not necessarily context-free.

2 Chomsky Algebras

2.1 Polynomials

Let $(C, +, \cdot, 0, 1)$ be an idempotent semiring and X a fixed set of variables. A *polynomial over indeterminates X with coefficients in C* is an element of $C[X]$, where $C[X]$ is the coproduct (direct sum) of C and

the free idempotent semiring on generators X in the category of idempotent semirings. For example, if $a,b,c \in C$ and $x,y \in X$, then the following are polynomials:

$$0 \qquad a \qquad axbycx+1 \qquad ax^2byx+by^2xc \qquad 1+x+x^2+x^3$$

The elements of $C[X]$ are not purely syntactic, as they satisfy all the equations of idempotent semirings and identities of C. For example, if $a^2 = b^2 = 1$ in C, then

$$(axa+byb)^2 = ax^2a+axabyb+bybaxa+by^2b.$$

Every polynomial can be written as a finite sum of monomials of the form

$$a_0x_0a_1x_1\cdots a_{n-1}x_{n-1}a_n,$$

where each $a_i \in C-\{0\}$ and $x_i \in X$. The *free variables* of such an expression p are the elements of X appearing in it and are denoted $\mathsf{FV}(p)$. The representation is unique up to associativity of multiplication and associativity, commutativity, and idempotence of addition.

2.2 Polynomial Functions and Evaluation

Let $C[X]$ be the semiring of polynomials over indeterminates X and let D be an idempotent semiring containing C as a subalgebra. By general considerations of universal algebra, any valuation $\sigma : X \to D$ extends uniquely to a semiring homomorphism $\hat{\sigma} : C[X] \to D$ preserving C pointwise. Formally, the functor $X \mapsto C[X]$ is left adjoint to a forgetful functor that takes an idempotent semiring D to its underlying set. Intuitively, $\hat{\sigma}$ is the *evaluation morphism* that evaluates a polynomial at the point $\sigma \in D^X$. Thus each polynomial $p \in C[X]$ determines a *polynomial function* $[\![p]\!] : D^X \to D$, where $[\![p]\!](\sigma) = \hat{\sigma}(p)$.

The set of all functions $D^X \to D$ with the pointwise semiring operations is itself an idempotent semiring with C as an embedded subalgebra under the embedding $c \mapsto \lambda\sigma.c$. The map $[\![\cdot]\!] : C[X] \to (D^X \to D)$ is actually $\hat{\tau}$, where $\tau(x) = \lambda f.f(x)$.

For the remainder of the paper, we write σ for $\hat{\sigma}$, as there is no longer any need to distinguish them.

2.3 Algebraic Closure and Chomsky Algebras

A *system of polynomial inequalities over* C is a set

$$p_1 \le x_1,\ p_2 \le x_2,\ \ldots,\ p_n \le x_n \tag{1}$$

where $x_i \in X$ and $p_i \in C[X]$, $1 \le i \le n$. A *solution* of (1) in C is a valuation $\sigma : X \to C$ such that $\sigma(p_i) \le \sigma(x_i)$, $1 \le i \le n$. The solution σ is a *least solution* if $\sigma \le \tau$ pointwise for any other solution τ. If a least solution exists, then it is unique.

An idempotent semiring C is said to be *algebraically closed* if every finite system of polynomial inequalities over C has a least solution in C.

The category of *Chomsky algebras* consists of algebraically closed idempotent semirings along with semiring homomorphisms that preserve least solutions of systems of polynomial inequalities.

The canonical example of a Chomsky algebra is the family of context-free languages $\mathsf{CF}X$ over an alphabet X. A system of polynomial inequalities (1) can be regarded as context-free grammar, and the least solution of the system is the context-free language generated by the grammar. For example, the set of strings in $\{a,b\}^\star$ with equally many a's and b's is generated by the grammar

$$S \to \varepsilon \mid aB \mid bA \qquad A \to aS \mid bAA \qquad B \to bS \mid aBB, \tag{2}$$

which corresponds to the system

$$1 + aB + bA \leq S \qquad aS + bAA \leq A \qquad bS + aBB \leq B, \tag{3}$$

where the symbols a, b are interpreted as the singleton sets $\{a\}, \{b\}$, the symbols S, A, B are variables ranging over sets of strings, and the semiring operations $+$, $\cdot$, 0, and 1 are interpreted as set union, set product $AB = \{xy \mid x \in A,\ y \in B\}$, $\emptyset$, and $\{\varepsilon\}$, respectively.

2.4 μ-Expressions

Let X be a set of indeterminates. Leiß [13] and Ésik and Leiß [5, 6] consider *μ-expressions* defined by the grammar

$$t ::= x \mid t + t \mid t \cdot t \mid 0 \mid 1 \mid \mu x.t$$

where $x \in X$. These expressions provide a syntax with which least solutions of polynomial systems can be named. Scope, bound and free occurrences of variables, α-conversion, and safe substitution are defined as usual (see e.g. [1]). We denote by $t[x/u]$ the result of substituting u for all free occurrences of x in t, renaming bound variables as necessary to avoid capture. Let $\mathsf{T}X$ denote the set of μ-expressions over indeterminates X.

Let C be a Chomsky algebra and X a set of indeterminates. An *interpretation* over C is a map $\sigma : \mathsf{T}X \to C$ that is a homomorphism with respect to the semiring operations and such that

$$\sigma(\mu x.t) = \text{the least } a \in C \text{ such that } \sigma[x/a](t) \leq a, \tag{4}$$

where $\sigma[x/a]$ denotes σ with x rebound to a. The element a exists and is unique: Informally, each μ-expression t can be associated with a system of polynomial inequalities such that $\sigma(t)$ is a designated component of its least solution, which exists by algebraic closure.

Every set map $\sigma : X \to C$ extends uniquely to such a homomorphism. An interpretation σ *satisfies* the equation $s = t$ if $\sigma(s) = \sigma(t)$ and satisfies the inequality $s \leq t$ if $\sigma(s) \leq \sigma(t)$. All interpretations over Chomsky algebras satisfy the axioms of idempotent semirings, α-conversion (renaming of bound variables), and the *Park axioms*

$$t[x/\mu x.t] \leq \mu x.t \qquad t \leq x \ \Rightarrow\ \mu x.t \leq x. \tag{5}$$

The Park axioms say intuitively that $\mu x.t$ is the least solution of the single inequality $t \leq x$. It follows easily that

$$t[x/\mu x.t] = \mu x.t. \tag{6}$$

Thus Chomsky algebras are essentially the ordered Park μ-semirings of [6] with the additional restriction that $+$ is idempotent and the order is the natural order $x \leq y \Leftrightarrow x + y = y$.

2.5 Bekić's Theorem

It is well known that the ability to name least solutions of single inequalities with μ gives the ability to name least solutions of all finite systems of inequalities. This is known as Bekić's theorem [2]. The construction is analogous to the definition of $M^\star$ for a matrix M over a Kleene algebra.

Bekić's theorem can be proved by regarding a system of inequalities as a single inequality on a Cartesian product, partitioning into two systems of smaller dimension, then applying the result for the 2×2 case inductively. The 2×2 system

$$p(x,y) \le x \qquad\qquad q(x,y) \le y$$

has least solution a_0, b_0, where

$$a(y) = \mu x.p(x,y) \qquad\qquad b_0 = \mu y.q(a(y),y) \qquad\qquad a_0 = a(b_0),$$

as can be shown using the Park axioms (5); see [14] or [6] for a comprehensive treatment.

For example, in the context-free languages, the set of strings in $\{a,b\}^\star$ with equally many a's and b's is represented by the term

$$\mu S.(1 + a \cdot \mu B.(bS + aBB) + b \cdot \mu A.(aS + bAA)) \tag{7}$$

obtained from the system (2) by this construction.

2.6 μ-Continuity

Let $nx.t$ be an abbreviation for the n-fold composition of t applied to 0, defined inductively by

$$0x.t = 0 \qquad\qquad (n+1)x.t = t[x/nx.t].$$

A Chomsky algebra is called *μ-continuous* if it satisfies the *μ-continuity axiom*:

$$a(\mu x.t)b = \sum_{n \ge 0} a(nx.t)b, \tag{8}$$

where the summation symbol denotes supremum with respect to the natural order $x \le y \Leftrightarrow x + y = y$. Note that the supremum of a and b is $a + b$.

The family $\mathsf{CF}X$ of context-free languages over an alphabet X forms a μ-continuous Chomsky algebra. The *canonical interpretation* over this algebra is $L_X : \mathsf{T}X \to \mathsf{CF}X$, where

$$\begin{aligned} L_X(x) &= \{x\} & L_X(t+u) &= L_X(t) \cup L_X(u) \\ L_X(0) &= \emptyset & L_X(tu) &= \{xy \mid x \in L_X(t),\ y \in L_X(u)\} \\ L_X(1) &= \{\varepsilon\} & L_X(\mu x.t) &= \bigcup_{n \ge 0} L_X(nx.t). \end{aligned} \tag{9}$$

Under L_X, every term in $\mathsf{T}X$ represents a context-free language over its free variables (note that x is not free in $nx.t$). In the example (7) of §2.5, the free variables are a, b and the bound variables are S, A, B, corresponding to the terminal and nonterminal symbols, respectively, of the grammar (2) of §2.3.

2.7 Relation to Other Axiomatizations

In this section we show that the various axiomatizations considered in [5, 6, 13] are valid in all μ-continuous Chomsky algebras.

A *μ-semiring* [6] is a semiring $(A, +, \cdot, 0, 1)$ satisfying the *μ-congruence* and *substitution* properties:

$$t = u \Rightarrow \mu x.t = \mu x.u \qquad\qquad \sigma(t[y/u]) = \sigma[y/\sigma(u)](t).$$

Idempotence is not assumed.

Lemma 2.1. *Every Chomsky algebra is a μ-semiring.*

Proof. The μ-congruence property is immediate from the definition of the μ operation (4). The substitution property is a general property of systems with variable bindings; see [1, Lemma 5.1.5]. It can be proved by induction. For the case of $\mu x.t$, we assume without loss of generality that $y \neq x$ (otherwise there is nothing to prove) and that x is not free in u.

$$\begin{aligned}\sigma((\mu x.t)[y/u]) &= \sigma(\mu x.(t[y/u])) \\ &= \text{least } a \text{ such that } \sigma[x/a](t[y/u]) \leq a \\ &= \text{least } a \text{ such that } \sigma[x/a][y/\sigma(u)](t) \leq a \\ &= \text{least } a \text{ such that } \sigma[y/\sigma(u)][x/a](t) \leq a \\ &= \sigma[y/\sigma(u)](\mu x.t).\end{aligned}$$

□

We now consider various axioms proposed in [13].

Lemma 2.2. *In all μ-continuous Chomsky algebras,*

$$\mu x.(1+ax) = \mu x.(1+xa), \quad x \notin \mathsf{FV}(a).$$

Proof. By μ-continuity, it suffices to show that $nx.(1+ax) = nx.(1+xa)$ for all n. We show by induction that for all n, $nx.(1+ax) = nx.(1+xa) = \sum_{i=0}^{n} a^i$. The basis $n = 0$ is trivial. For the inductive case,

$$(n+1)x.(1+ax) = 1+a(nx.(1+ax)) = 1+a(\textstyle\sum_{i=0}^{n} a^i) = \sum_{i=0}^{n+1} a^i,$$

and this is equal to $(n+1)x.(1+xa)$ by a symmetric argument. □

Lemma 2.3. *The following two equations hold in all μ-continuous Chomsky algebras:*

$$a(\mu x.(1+xb)) = \mu x.(a+xb) \qquad (\mu x.(1+bx))a = \mu x.(a+bx).$$

Proof. We show the first equation only; the second follows from a symmetric argument. By μ-continuity, we need only show that the equation holds for any n. The basis $n = 0$ is trivial. For the inductive case,

$$\begin{aligned}a((n+1)x.(1+xb)) &= a + a(nx.(1+xb))b \\ &= a + (nx.(a+xb))b \\ &= (n+1)x.(a+xb),\end{aligned}$$

where the induction hypothesis has been used in the second step. □

These properties also show that μ-continuous Chomsky algebras are algebraically complete semirings in the sense of [5, 6].

Lemma 2.4. *The* Greibach inequalities

$$\mu x.s(\mu y.(1+ry)) \leq \mu x.(s+xr) \qquad \mu x.(\mu y.(1+yr))s \leq \mu x.(s+rx)$$

of KAG [13] *hold in all μ-continuous Chomsky algebras.*

Proof. For the left-hand inequality, let $u = \mu x.(s + xr)$. By the Park axioms, it suffices to show that $s(\mu y.(1+ry))[x/u] \leq u$. But

$$\begin{aligned} s(\mu y.(1+ry))[x/u] &= s[x/u](\mu y.(1+r[x/u]y)) \\ &= s[x/u](\mu y.(1+yr[x/u])) \\ &= \mu y.(s[x/u]+yr[x/u]) \\ &= \mu x.(s+xr), \end{aligned}$$

where Lemmas 2.2 and 2.3 have been used.

The right-hand ineuuality can be proved by a symmetric argument. □

Various other axioms of [5, 6, 13] follow from the Park axioms.

The μ-continuity condition (8) implies the Park axioms (5), but we must defer the proof of this fact until §3. For now we just observe a related property of the canonical interpretation L_X.

Lemma 2.5. *For any $s,t \in \mathsf{T}X$ and $y \in X$,*

$$L_X(s[y/\mu y.t]) = \bigcup_{n\geq 0} L_X(s[y/ny.t]).$$

Proof. We proceed by induction on the structure of s. The cases for $+$ and $\cdot$ are quite easy, using the facts that for chains of sets of strings $A_0 \subseteq A_1 \subseteq A_2 \subseteq \cdots$ and $B_0 \subseteq B_1 \subseteq B_2 \subseteq \cdots$,

$$\bigcup_m A_m \cup \bigcup_n B_n = \bigcup_n A_n \cup B_n \qquad\qquad \bigcup_m A_m \cdot \bigcup_n B_n = \bigcup_n A_n B_n.$$

The base cases are also straightforward. For $\mu x.s$, assume without loss of generality that $y \neq x$ and x is not free in t.

$$\begin{aligned} L_X((\mu x.s)[y/\mu y.t]) &= \bigcup_m L_X((mx.s)[y/\mu y.t]) \\ &= \bigcup_m \bigcup_n L_X((mx.s)[y/ny.t]) \\ &= \bigcup_n \bigcup_m L_X((mx.s)[y/ny.t]) \\ &= \bigcup_n L_X((\mu x.s)[y/ny.t]). \end{aligned}$$

□

3 Main Results

Our main result depends on an analog of a result of [10] (see [12]). It asserts that the supremum of a context-free language over a μ-continuous Chomsky algebra K exists, interpreting strings over K as products in K. Moreover, multiplication is continuous with respect to suprema of context-free languages.

Lemma 3.1. *Let $\sigma : \mathsf{T}X \to K$ be any interpretation over a μ-continuous Chomsky algebra K. Let $\tau : \mathsf{T}X \to \mathsf{CF}X$ be any interpretation over the context-free languages $\mathsf{CF}X$ such that for all $x \in X$ and $s,u \in \mathsf{T}X$,*

$$\sigma(sxu) = \sum_{y\in\tau(x)} \sigma(syu).$$

Then for any $s,t,u \in \mathsf{T}X$,

$$\sigma(stu) = \sum_{y \in \tau(t)} \sigma(syu).$$

In particular,

$$\sigma(stu) = \sum_{y \in L_X(t)} \sigma(syu), \tag{10}$$

where L_X is the canonical interpretation defined in §2.6.

Remark 1. Note carefully that the lemma does not assume *a priori* knowledge of the existence of the suprema. The equations should be interpreted as asserting that the supremum on the right-hand side exists and is equal to the expression on the left-hand side.

Proof. The proof is by induction on the structure of t, that is by induction on the subexpression relation $t+u \succ t, t+u \succ u, t \cdot u \succ t, t \cdot u \succ u, \mu x.t \succ nx.t$, which is well-founded [11].

All cases are similar to the proof in [12, Lemma 7.1] for star-continuous Kleene algebra, with the exception of the case $t = \mu x.p$.

For variables $t = x \in X$, the desired property holds by assumption. For the constants $t = 0$ and $t = 1$,

$$\sigma(s0u) = 0 = \sum \emptyset = \sum_{y \in \emptyset} \sigma(syu) = \sum_{y \in \tau(0)} \sigma(syu)$$
$$\sigma(s1u) = \sigma(su) = \sum_{y \in \{\varepsilon\}} \sigma(syu) = \sum_{y \in \tau(1)} \sigma(syu).$$

For sums $t = p+q$,

$$\begin{aligned}
\sigma(s(p+q)u) &= \sigma(spu) + \sigma(squ) \\
&= \sum_{x \in \tau(p)} \sigma(sxu) + \sum_{y \in \tau(q)} \sigma(syu) && (11) \\
&= \sum_{z \in \tau(p) \cup \tau(q)} \sigma(szu) && (12) \\
&= \sum_{z \in \tau(p+q)} \sigma(szu). && (13)
\end{aligned}$$

Equation (11) is by two applications of the induction hypothesis. Equation (12) is by the properties of supremum. Equation (13) is by the definition of sum in $\mathsf{CF}X$.

For products $t = pq$,

$$\begin{aligned}
\sigma(spqu) &= \sum_{x \in \tau(p)} \sum_{y \in \tau(q)} \sigma(sxyu) && (14) \\
&= \sum_{z \in \tau(p) \cdot \tau(q)} \sigma(szu) && (15) \\
&= \sum_{z \in \tau(pq)} \sigma(szu). && (16)
\end{aligned}$$

Equation (14) is by two applications of the induction hypothesis. Equations (15) and (16) are by the definition of product in $\mathsf{CF}X$.

Finally, for $t = \mu x.p$,

$$\sigma(s(\mu x.p)u) = \sum_n \sigma(s(nx.p)u) \tag{17}$$
$$= \sum_n \sum_{y \in \tau(nx.p)} \sigma(syu) \tag{18}$$
$$= \sum_{y \in \bigcup_n \tau(nx.p)} \sigma(syu) \tag{19}$$
$$= \sum_{y \in \tau(\mu x.p)} \sigma(syu). \tag{20}$$

Equation (17) is just the μ-continuity property (8). Equation (18) is by the induction hypothesis, observing that $\mu x.p \succ nx.p$. Equation (19) is a basic property of suprema. Finally, equation (20) is by the definition of $\tau(\mu x.p)$ in $\mathsf{CF}X$.

The result (10) for the special case of $\tau = L_X$ is immediate, observing that L_X satisfies the assumption of the lemma: for $x \in X$,

$$\sigma(sxu) = \sum_{y \in \{x\}} \sigma(syu) = \sum_{y \in L_X(x)} \sigma(syu).$$

□

At this point we can show that the μ-continuity condition implies the Park axioms.

Theorem 3.2. *The μ-continuity condition* (8) *implies the Park axioms* (5).

Proof. We first show $p \le x \Rightarrow \mu x.p \le x$ in any idempotent semiring satisfying the μ-continuity condition. Let σ be a valuation such that $\sigma(\mu x.p) = \sum_n \sigma(nx.p)$. Suppose that $\sigma(p) \le \sigma(x)$. We show by induction that for all $n \ge 0$, $\sigma(nx.p) \le \sigma(x)$. This is certainly true for $0x.p = 0$. Now suppose it is true for $nx.p$. Using monotonicity,

$$\sigma((n{+}1)x.p) = \sigma(p[x/nx.p]) \le \sigma(p[x/x]) = \sigma(p) \le \sigma(x).$$

By μ-continuity, $\sigma(\mu x.p) = \sum_n \sigma(nx.p) \le \sigma(x)$.

Now we show that $p[x/\mu x.p] \le \mu x.p$. This requires the stronger property that a μ-expression is chain-continuous with respect to suprema of context-free languages as a function of its free variables. Using Lemmas 2.5 and 3.1,

$$\begin{aligned}
\sigma(p[x/\mu x.p]) &= \sum \{\sigma(y) \mid y \in L_X(p[x/\mu x.p])\} \\
&= \sum \left\{ \sigma(y) \mid y \in \bigcup_n L_X(p[x/nx.p]) \right\} \\
&= \sum_n \sum \{\sigma(y) \mid y \in L_X(p[x/nx.p])\} \\
&= \sum_n \sigma(p[x/nx.p]) \\
&= \sum_n \sigma((n{+}1)x.p) \\
&= \sigma(\mu x.p).
\end{aligned}$$

□

The following is our main theorem.

Theorem 3.3. *Let X be an arbitrary set and let $s,t \in \mathsf{T}X$. The following are equivalent:*

(i) *The equation $s = t$ holds in all μ-continuous Chomsky algebras; that is, $s = t$ is a logical consequence of the axioms of idempotent semirings and the μ-continuity condition*

$$a(\mu x.t)b = \sum_{n\geq 0} a(nx.t)b, \tag{21}$$

or equivalently, the universal formulas

$$a(nx.t)b \leq a(\mu x.t)b, \quad n \geq 0 \tag{22}$$

$$\left(\bigwedge_{n\geq 0} (a(nx.t)b \leq w)\right) \;\Rightarrow\; a(\mu x.t)b \leq w. \tag{23}$$

(ii) *The equation $s = t$ holds in the semiring of context-free languages $\mathsf{CF}Y$ over any set Y.*

(iii) *$L_X(s) = L_X(t)$, where $L_X : \mathsf{T}X \to \mathsf{CF}X$ is the standard interpretation mapping a μ-expression to a context-free language of strings over its free variables.*

Thus the axioms of idempotent semirings and μ-continuity are sound and complete for the equational theory of the context-free languages.

Proof. The implication (i) $\Rightarrow$ (ii) holds since $\mathsf{CF}Y$ is a μ-continuous Chomsky algebra, and (iii) is a special case of (ii). Finally, if (iii) holds, then by two applications of Lemma 3.1, for any interpretation $\sigma : \mathsf{T}X \to K$ over a μ-continuous Chomsky algebra K,

$$\sigma(s) = \sum_{x\in L_K(s)} \sigma(x) = \sum_{x\in L_K(t)} \sigma(x) = \sigma(t),$$

which proves (i). □

Theorem 3.4. *The context-free languages over the alphabet X form the free μ-continuous Chomsky algebra on generators X.*

Proof. Let K be a μ-continuous Chomsky algebra. Any map $\sigma : X \to K$ extends uniquely to an interpretation $\sigma : \mathsf{T}X \to K$. By Lemma 3.1, this decomposes as

$$\sigma = \textstyle\sum \circ \mathsf{CF}\,\sigma \circ L_X,$$

where $L_X : \mathsf{T}X \to \mathsf{CF}X$ is the canonical interpretation in the context-free languages over X, $\mathsf{CF}\,\sigma : \mathsf{CF}X \to \mathsf{CF}K$ is the map $\mathsf{CF}\,\sigma(A) = \{\sigma(x) \mid x \in A\}$, and $\sum : \mathsf{CF}K \to K$ takes the supremum of a context-free language over K, which is guaranteed to exist by Lemma 3.1. The unique morphism $\mathsf{CF}X \to K$ corresponding to σ is $\sum \circ \mathsf{CF}\,\sigma$. Thus CF is left adjoint to the forgetful functor from μ-continuous Chomsky algebras to Set. The maps $x \mapsto \{x\} : X \to \mathsf{CF}X$ and $\sum : \mathsf{CF}K \to K$ are the unit and counit, respectively, of the adjunction. □

4 Conclusion

We have given a natural complete infinitary axiomatization of the equational theory of the context-free languages. Leiß [13] states as an open problem:

> Are there natural equations between μ-regular expressions that are valid in all continuous models of KAF, but go beyond KAG?

We have identified such a system in this paper, thereby answering Leiß's question. He does not state axiomatization as an open problem, but observes that the set of pairs of equivalent context-free grammars is not recursively enumerable, then goes on to state:

> Since there is an effective translation between context-free grammars and μ–regular expressions ..., the equational theory of context-free languages in terms of μ-regular expressions is not axiomatizable at all.

Nevertheless, we have given an axiomatization. How do we reconcile these two views? Leiß is apparently using "axiomatization" in the sense of "recursive axiomatization." But observe that the axiom (23) is an infinitary Horn formula. To use it as a rule of inference, one would need to establish infinitely many premises of the form $x(ny.p)z \leq w$. But this in itself is a Π^0_1-complete problem. One can show that it is Π^0_1-complete to determine whether a given context-free grammar G over a two-letter alphabet generates all strings. By coding G as a μ-expression w, the problem becomes $\mu x.(1+ax+bx) \leq w$, which by (21) is equivalent to showing that $nx.(1+ax+bx) \leq w$ for all n.

Acknowledgments

We thank Zoltán Ésik, Hans Leiß, and the anonymous referees for helpful comments. The DIKU-affiliated authors express their thanks to the Department of Computer Science at Cornell University for hosting them in the Spring 2013 and to the Danish Council for Independent Research for financial support for this work under Project 11-106278, "Kleene Meets Church (KMC): Regular Expressions and Types".

References

[1] Henk Barendregt (1984): *The Lambda Calculus: Its Syntax and Semantics*. *Studies in Logic and the Foundations of Mathematics* 103, North-Holland.

[2] Hans Bekić (1984): *Definable operations in general algebras, and the theory of automata and flowcharts*. In C.B. Jones, editor: *Programming Languages and Their Definition, Lecture Notes in Computer Science* 177, Springer Berlin Heidelberg, pp. 30–55, doi:10.1007/BFb0048939.

[3] Bruno Courcelle (1986): *Equivalences and Transformations of Regular Systems – Applications to Recursive Program Schemes and Grammars*. *Theoretical Computer Science* 42, pp. 1–122, doi:10.1016/0304-3975(86)90050-2.

[4] Zoltán Ésik & Werner Kuich (2007): *Modern automata theory*. Unpublished manuscript.

[5] Zoltán Ésik & Hans Leiß (2002): *Greibach Normal Form in Algebraically Complete Semirings*. In: *CSL '02: Proceedings of the 16th International Workshop and 11th Annual Conference of the EACSL on Computer Science Logic*, Springer-Verlag, London, UK, pp. 135–150, doi:10.1007/3-540-45793-3_10.

[6] Zoltán Ésik & Hans Leiß (2005): *Algebraically Complete Semirings and Greibach Normal Form*. *Annals of Pure and Applied Logic* 133, pp. 173–203, doi:10.1016/j.apal.2004.10.008.

[7] Jozef Gruska (1971): *A characterization of context-free languages*. *J. Comput. Syst. Sci.* 5(4), pp. 353–364, doi:10.1016/S0022-0000(71)80023-5.

[8] Mark Hopkins (2008): *The Algebraic Approach I: The Algebraization of the Chomsky Hierarchy*. In R. Berghammer, B. Möller & G. Struth, editors: *Proc. 10th Int. Conf. Relational Methods in Computer Science and 5th Int. Conf. Applications of Kleene Algebra (RelMiCS/AKA 2008)*, *Lecture Notes in Computer Science* 4988, Springer-Verlag, Berlin Heidelberg, pp. 155–172, doi:10.1007/978-3-540-78913-0_13.

[9] Mark Hopkins (2008): *The Algebraic Approach II: Dioids, Quantales and Monads*. In R. Berghammer, B. Möller & G. Struth, editors: *Proc. 10th Int. Conf. Relational Methods in Computer Science and 5th Int. Conf. Applications of Kleene Algebra (RelMiCS/AKA 2008)*, *Lecture Notes in Computer Science* 4988, Springer-Verlag, Berlin Heidelberg, pp. 173–190, doi:10.1007/978-3-540-78913-0_14.

[10] Dexter Kozen (1981): *On Induction vs. *-Continuity*. In: *Proc. Logics of Programs*, *Lecture Notes in Computer Science (LNCS)* 131, Springer, pp. 167–176, doi:10.1007/BFb0025769.

[11] Dexter Kozen (1983): *Results on the propositional [mu]-calculus*. *Theoretical Computer Science* 27(3), pp. 333 – 354, doi:10.1016/0304-3975(82)90125-6.

[12] Dexter Kozen (1991): *The Design and Analysis of Algorithms*. Springer-Verlag, New York, doi:10.1007/978-1-4612-4400-4.

[13] Hans Leiß (1992): *Towards Kleene Algebra with Recursion*. In: *CSL '91: Proceedings of the 5th Workshop on Computer Science Logic*, Springer-Verlag, London, UK, pp. 242–256, doi:10.1007/BFb0023771.

[14] Glynn Winskel (1993): *The Formal Semantics of Programming Languages*. MIT Press.