The Matrix Ring of a μ-Continuous Chomsky Algebra is μ-Continuous

Hans Leiss

Centrum für Informations- und Sprachverarbeitung,
Ludwig-Maximilians-Universität München,
Oettingenstr. 67, D-80538 München, Germany
leiss@cis.uni-muenchen.de

Abstract

In the course of providing an (infinitary) axiomatization of the equational theory of the class of context-free languages, Grathwohl, Kozen and Henglein (2013) have introduced the class of μ-continuous Chomsky algebras. These are idempotent semirings where least solutions for systems of polynomial inequations (i.e. context-free grammars) can be computed iteratively and where multiplication is continuous with respect to the least fixed point operator μ. We prove that the matrix ring of a μ-continuous Chomsky algebra also is a μ-continuous Chomsky algebra.

1998 ACM Subject Classification F 4.3 Formal languages, F 4.2 Grammars and Other Rewriting Systems, D 3.3 Language Constructs, F.3.3 Studies of Program Constructs

Keywords and phrases context-free language, fixed point operator, idempotent semiring, matrix ring, Chomsky algebra

Digital Object Identifier 10.4230/LIPIcs.CSL.2016.6

1 Introduction

The set of context-free languages over an alphabet X has been characterized algebraically by Gruska [4] as the closure of the finite languages over X under (binary) union $+$, elementwise concatenation $\cdot$, and a least fixed point operator μ. The natural definition of a context-free language is by a simultaneous definition involving auxiliary languages, which suggests using an n-ary fixed point operator to denote the solution of a system $x_i \geq p_i(x_1, \ldots, x_n)$, $1 \leq i \leq n$, of polynomial inequations. Bekić [1], deBakker and Scott [3] noticed that a unary least fixed point operator suffices to name the components of an n-ary least fixed point, provided (i) every countable ascending chain has a supremum in the underlying partial order, and (ii) the functions p_i in the systems $x_i \geq p_i$ are componentwise continuous, i.e. map the sup of an ascending chain to the sup of the image of the chain.

Terms involving a unary fixed point operator μ in addition to semiring operations $+$, $\cdot$, 0, 1, here called μ-terms, are therefore a means to name the context-free languages. Kleene's iteration operator * amounts to a special case of recursion μ, namely to head- or tail recursion, $r^* = \mu x(rx + 1)$. Small fragments of the equational theory of context-free languages using μ-terms had been studied in [9, 10], but only recently a complete (infinitary) axiomatization has been given by Grathwohl, Henglein and Kozen [11]. Essentially, it says that μxt is the sup of all finite iterations of t and that the semiring operations are continuous with respect to definable increasing chains. In the models of this theory, the μ-continuous Chomsky algebras of [11], the partial order derived from $+$ need *not* be complete.

Simple equations between μ-terms relate head- to tail recursion and reflexive transitive closure [9], and most of the context-free grammar normalization algorithms can be derived as equations between μ-terms from minimality assumptions on μ and the semiring properties

25th EACSL Annual Conference on Computer Science Logic (CSL 2016).
Editors: Jean-Marc Talbot and Laurent Regnier; Article No. 6; pp. 6:1–6:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

of +, ·, 0, 1 [10]. In particular, one can express the elimination of a head recursion by a tail recursion as an equation between μ-terms (assuming x is not an initial factor in s)

$$\mu x(xr + s) = \mu x \mu y(yr + s) = \mu x(s \cdot r^*) = \mu x(s \cdot \mu y(ry + 1)), \qquad \text{(with fresh } y\text{)}$$

and then obtain (an efficient algorithm for) the transformation of context-free grammars to Greibach normal form as the matrix version of this equation; an example is given in [10].

It is therefore of some interest to know if a property involving the unary fixed point operator lifts to the n-ary or the matrix (n^2-ary) case. A related question, which we consider here, is whether the matrix algebra of an idempotent semiring that is closed under least fixed points is itself closed under least fixed points.

We will reduce the matrix case to the vector case, and the vector case to the unary one by using the Bekić-Scott equations. However, the algebras we are considering, like the algebra of all context-free languages over X, are *not* closed under countable unions of ascending chains. Therefore, we must take some care to show that all the least fixed points involved in Bekić's reduction exist. We separate arguments about existence of least fixed points from the question of iteratively computing them, restricting the μ-continuity condition of Grathwohl e.a.[11] to the second question. Thus, if M is a Chomsky algebra, i.e. all polynomial inequation systems have least solutions, so is $Mat_{n,n}(M)$, and if M is a μ-continuous Chomsky algebra, $Mat_{n,n}(M)$ also is. In particular, the μ-continuity condition gives rise to its own matrix version. For several conditions on Kleene's iteration * in regular algebra, Conway [2] had noted that the unary case implies the matrix case; likewise, Kozen's axioms for * imply their own matrix versions [8].

2 Park μ-Semirings

▶ **Definition 1.** A **semiring** $(M, +, \cdot, 0, 1)$ is a set M with binary operations $+, \cdot$ on M and elements $0, 1 \in M$ such that $+$ is associative and commutative with neutral element 0, $\cdot$ is associative with neutral element 1 and annihilator 0, and $\cdot$ distributes over $+$ from both sides. An **idempotent semiring** is a semiring where $+$ is idempotent.

▶ **Definition 2.** Let X be a set of variables. The set of **μ-terms over** X is defined by the grammar

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

A term not containing μ will be called **algebraic** or (somewhat unprecise) a **polynomial**. The free occurrences of variables in a term are defined as usual. By ***free***(t) we denote the set of variables having a free occurrence in t; in particular, $free(\mu x t) = free(t) \setminus \{x\}$. By $t(x_1, \ldots, x_n)$ we indicate $free(t) \subseteq \{x_1, \ldots, x_n\}$. In $\mu x t$ all free occurrences of x in t are **bound** by μx. By $t[x/s]$ we denote the result of substituting all free occurrences of x in t by s, renaming bound variables of t to avoid capture of free variables of s by bindings in t. The **μ-depth** of a term is 0 for the terms x, 0, 1, is 1 plus the μ-depth of t for the term $\mu x t$, and is the maximum of the μ-depth of its immediate subterms, otherwise.

▶ **Definition 3.** A **partially ordered μ-semiring** $(M, +, \cdot, 0, 1, \leq)$ is a semiring $(M, +, \cdot, 0, 1)$ with a partial order $\leq$ on M, where every term t defines a function $t^M : (X \to M) \to M$, so that for all variables $x \in X$ and terms s, t we have:

1. for all valuations $g : X \to M$,

$$\begin{array}{rclcrcl} 0^M(g) &=& 0, & & (s+t)^M(g) &=& s^M(g) + t^M(g), \\ 1^M(g) &=& 1, & & (s \cdot t)^M(g) &=& s^M(g) \cdot g^M(g), \\ x^M(g) &=& g(x), & \text{if } s^M \leq t^M, & \text{then} & & \mu x s^M \leq \mu x t^M, \end{array}$$

2. t^M is monotone with respect to the pointwise order on $X \to M$,
3. $t^M(g) = t^M(h)$, for all valuations $g, h : X \to M$ which agree on $\mathit{free}(t)$,
4. $t[x/s]^M(g) = t^M(g[x/s^M(g)])$, for all valuations $g : X \to M$.

When $\mathit{free}(t) \subseteq \{x_1, \ldots, x_n\}$ and $g(x_i) = a_i$ for $1 \le i \le n$, instead of $t^M(g)$ we often write $t^M[x_1/a_1, \ldots, x_n/a_n]$ or just $t^M(a_1, \ldots, a_n)$.

The final two conditions above are called the **coincidence** and **substitution properties**; in the latter, $g[x/a]$ denotes the valuation that agrees with g, except that it assigns a to x. Clearly, the substitution property extends to simultaneous substitutions $[x_1/s_1, \ldots, x_n/s_n]$. A first-order formula built from equations and inequations between μ-terms **holds in** M if it is true for every valuation $g : X \to M$.

▶ **Definition 4.** A **Park μ-semiring** is a partially ordered μ-semiring M where for all terms t and variables x, y, the following hold in M:

$$t[x/\mu xt] \quad \le \quad \mu xt, \tag{1}$$

$$t[x/y] \le y \quad \to \quad \mu xt \le y. \tag{2}$$

It follows easily that $t[x/\mu xt] = \mu xt$ holds in M, as well as $\mu y.t[x/y] = \mu xt$, for $y \notin \mathit{free}(t)$.

Conditions (1) and (2) imply that $\mu xt^M(g)$ is the least solution of $t \le x$ in M, g, i.e. the least $a \in M$ such that $t^M(g[x/a]) \le a$.

3 Chomsky Algebras

▶ **Definition 5** ([11])**.** A **Chomsky algebra** $(M, +, \cdot, 0, 1)$ is an idempotent semiring where every finite system of polynomial inequations

$$\begin{array}{rcl} x_1 & \ge & p_1(x_1, \ldots, x_n, y_1, \ldots, y_m), \\ & \vdots & \\ x_n & \ge & p_n(x_1, \ldots, x_n, y_1, \ldots, y_m), \qquad \text{abbreviated } \bar{x} \ge \bar{p}(\bar{x}, \bar{y}), \end{array} \tag{3}$$

has least solutions, i.e. for all $\bar{b} \in M^m$ there is a least $\bar{a} = a_1, \ldots, a_n \in M^n$ such that $a_i \ge p_i^M(\bar{a}, \bar{b})$ for $i = 1, \ldots, n$, where $\le$ is the natural partial order on M defined by $a \le b$ iff $a + b = b$. Of course, for each $\bar{b}$ the least solution $\bar{a}$ is unique.

▶ **Example 6.** Let $(X^*, \cdot, \epsilon)$ be the monoid of all finite words of elements of X, with concatenation $\cdot$ as product and the empty string ϵ as unit. Its power set $\mathcal{P}X^*$, the set of all languages over X, is an idempotent semiring $(\mathcal{P}X^*, +, \cdot, 0, 1)$, where $0 := \emptyset$, $1 := \{\epsilon\}$, and for $A, B \subseteq X^*$, $A + B := A \cup B$ is set union and $A \cdot B := \{a \cdot b \mid a \in A, b \in B\}$ the elementwise concatenation. A finite system of polynomial inequations (3) is a **context-free grammar** with nonterminals $x_1, \ldots, x_n$ and terminals $y_1, \ldots, y_m$. For a vector $\bar{B}$ of m languages, it leads to an increasing sequence $\bar{A}_k = (A_{k,1}, \ldots, A_{k,n})$ of language vectors by

$$A_{0,i} := \emptyset, \quad A_{k+1,i} := p_i^{\mathcal{P}X^*}(\bar{A}_k, \bar{B}), \quad i = 1, \ldots, n.$$

Clearly, any solution $\bar{A}$ of $\bar{x} \ge \bar{p}^{\mathcal{P}X^*}[\bar{y}/\bar{B}]$ must satisfy $\bar{A} \supseteq \bigcup\{\bar{A}_k \mid k \in \mathbb{N}\}$, where union and subsumption are meant componentwise. The least solution of the inequation system, relative to $\bar{B}$, is $\bar{A} := \bigcup\{\bar{A}_k \mid k \in \mathbb{N}\}$, since $+$ and $\cdot$ are compatible with arbitrary unions, i.e. for languages $A, B \subseteq X^*$ and $\emptyset \ne \mathcal{C} \subseteq \mathcal{P}X^*$,

$$A + \bigcup \mathcal{C} \quad = \quad \bigcup\{A + C \mid C \in \mathcal{C}\},$$

$$A \cdot \bigcup \mathcal{C} \cdot B \quad = \quad \bigcup\{A \cdot C \cdot B \mid C \in \mathcal{C}\},$$

which implies $\bar{A} \supseteq \bar{p}^{\mathcal{P}X^*}(\bar{A}, \bar{B})$. Therefore, $(\mathcal{P}X^*, \cup, \cdot, \emptyset, \{\epsilon\})$ is a Chomsky algebra.

Another example is the set $Rel\,(S)$ of all binary relations over a set S with empty relation as 0, the identity relation as 1, union as + and relation product as $\cdot$. The standard example of a Chomsky algebra is the algebra of context-free languages over the alphabet X.

▶ **Example 7.** The set $\mathcal{C}X^*$ of **context-free languages** over X is the smallest set $\mathcal{L} \subseteq \mathcal{P}X^*$ such that (i) each finite subset of $X \cup \{\epsilon\}$ is in $\mathcal{L}$ and (ii) if $\bar{x} \geq \bar{p}(\bar{x}, \bar{y})$ is a polynomial system, and $\bar{B} = B_1, \ldots, B_m \in \mathcal{L}$, then the components A_i of the least $\bar{A} = A_1, \ldots, A_n \in \mathcal{P}X^*$ with $\bar{A} \supseteq \bar{p}^{\mathcal{P}X^*}(\bar{A}, \bar{B})$ belong to $\mathcal{L}$. With the operations inherited from $\mathcal{P}X^*$, $(\mathcal{C}X^*, +, \cdot, 0, 1)$ is a Chomsky algebra. For example, $\{a^n b^n \mid n \in \mathbb{N}\}$ is a context-free language over $X \supseteq \{a, b\}$; it is the least solution of $x \geq axb + 1$ relative to the standard valuation $g(a) = \{a\}, g(b) = \{b\}$.

Of course, the regular languages over X do not form a Chomsky algebra, as they don't have solutions for inequations like $axb + 1 \leq x$.

The next lemma is a slight improvement of Lemma 2.1 in [11] in that it cares about the well-definedness of $\mu x t^M(g)$ and the partial order.

▶ **Lemma 8.** *Every Chomsky algebra M is an idempotent, partially ordered μ-semiring, if for all terms t, variables x and valuations $g : X \to M$ we take*

$$\mu x t^M(g) := \text{the least } a \in M \text{ such that } t^M(g[x/a]) \leq a. \tag{4}$$

Moreover, every inequation system $\bar{t}(\bar{x}, \bar{y}) \leq \bar{x}$ with μ-terms $\bar{t}(\bar{x}, \bar{y})$ has least solutions in M, i.e. for all parameters $\bar{b}$ from M there is a least tuple $\bar{a}$ in M such that $\bar{t}^M(\bar{a}, \bar{b}) \leq \bar{a}$.

Proof. We define term functions $t^M : (X \to M) \to M$ by induction on the μ-depth of t. The cases for terms of the forms 0, 1, x, $(s + t)$, and $(s \cdot t)$ are obvious; for $\mu x t$ we need to show that $t \leq x$ has least solutions in M, so that the definition of $\mu x t^M(g)$ via (4) is well-defined. More generally, we simultaneously prove by induction on k:

1. every inequation system $\bar{t} \leq \bar{x}$ has least solutions in M, for terms $\bar{t}$ of μ-depth $\leq k$,
2. term functions t^M for terms t of μ-depth $\leq k$ satisfy the properties of term functions in partially ordered μ-semirings (cf. Definition 3).

Let $\bar{t}(\bar{x}, \bar{y}) \leq \bar{x}$ be a system of term inequations

$$\begin{array}{rcl} t_1(x_1, \ldots, x_n, y_1, \ldots, y_m) & \leq & x_1 \\ & \vdots & \\ t_n(x_1, \ldots, x_n, y_1, \ldots, y_m) & \leq & x_n \end{array} \tag{5}$$

where each t_i has μ-depth $\leq k$. We may assume that bound variables in $\bar{t}(\bar{x}, \bar{y}) \leq \bar{x}$ are pairwise distinct and distinct from $\bar{x}, \bar{y}$. If $k = 0$, $\bar{t} \leq \bar{x}$ is a system of polynomial inequations and therefore has least solutions in M. If $k > 0$, we may assume that $t_n \leq x_n$ is an inequation where t_n is of maximal μ-depth. Then t_n can be written as $t_n = t'_n[x_{n+1}/\mu x_{n+1} t_{n+1}]$ where $\mu x_{n+1} t_{n+1}$ is of maximal μ-depth. Consider the inequation system $\bar{t}' \leq \bar{x}'$ obtained from $\bar{t} \leq \bar{x}$ by replacing $t_n \leq x_n$ by the two inequations $t'_n \leq x_n$ and $t_{n+1} \leq x_{n+1}$. Its maximal μ-depth, or its number of terms of maximal μ-depth, is less than that of $\bar{t} \leq \bar{x}$. Hence, by induction, $\bar{t}' \leq \bar{x}'$ has least solutions in M. Fix parameters $\bar{b}$ and let $(\bar{a}, a_{n+1})$ be the least elements of M such that $\bar{t}'^M(\bar{a}, a_{n+1}, \bar{b}) \leq (\bar{a}, a_{n+1})$. Then, by definition, $a_{n+1} = \mu x_{n+1} t_{n+1}{}^M(\bar{a}, \bar{b})$, hence by the substitution property for terms of μ-depth $\leq k$,

$$t_n^M(\bar{a}, \bar{b}) = t'_n[x_{n+1}/\mu x_{n+1} t_{n+1}]^M(\bar{a}, \bar{b}) = t_n'^M(\bar{a}, a_{n+1}, \bar{b}) \leq a_n.$$

This gives $\bar{t}^M(\bar{a}, \bar{b}) \leq \bar{a}$, and so $\bar{a}$ is a solution of $\bar{t} \leq x$ relative to $\bar{b}$. If $\bar{c}$ is another one, take $c_{n+1} := \mu x_{n+1} t_{n+1}{}^M(\bar{c}, \bar{b})$. Then $t_{n+1}^M(\bar{c}, c_{n+1}, \bar{b}) \leq c_{n+1}$ and, by the substitution property, $t'^M_n(\bar{c}, c_{n+1}, \bar{b}) = t^M_n(\bar{c}, \bar{b}) \leq c_n$. Therefore, $\bar{t}'^M(\bar{c}, c_{n+1}, \bar{b}) \leq (\bar{c}, c_{n+1})$, whence $(\bar{a}, a_{n+1}) \leq (\bar{c}, c_{n+1})$ and $\bar{a} \leq \bar{c}$. Thus, $\bar{a}$ is the least solution of $\bar{t} \leq \bar{x}$ in M relative to $\bar{b}$.

We leave it to the reader to check that the properties of term functions in partially ordered μ-semirings hold in M for terms of μ-depth $\leq k$. ◀

▶ **Corollary 9.** *Under the interpretation of μ-terms in (4), every Chomsky algebra M is a Park μ-semiring. In particular, the context-free languages $\mathcal{C}X^*$ form a Park μ-semiring.*

Proof. Since $\mu x t^M(g)$ is the least $a \in M$ with $t^M(g[x/a]) \leq a$ and the substitution property holds in M, we have

$$t[x/\mu x t]^M(g) = t^M(g[x/\mu x t^M(g)]) \leq \mu x t^M(g).$$

Suppose $t[x/y]^M(g) \leq y^M(g) = g(y)$ for some $g : X \to M$ and some variable y. Then by the substitution property,

$$t^M(g[x/g(y)]) = t[x/y]^M(g) \leq g(y) = x^M(g[x/g(y)]),$$

so $g(y)$ is a solution of $t \leq x$ in M relative to g, and by comparison with the least solution, $\mu x t^M(g) \leq g(y)$. ◀

▶ **Lemma 10.** *If $g : X \to \mathcal{C}X^*$, then $t^{\mathcal{C}X^*}(g) = t^{\mathcal{P}X^*}(g)$ for each μ-term t.*

Proof. By induction on the structure of t. For $\mu x t$, let $\bar{x} \geq \bar{p}_t(\bar{x}, \bar{y})$ be the polynomial system obtained from $t(x, \bar{y})$ in the proof of Lemma 8 such that

$$\mu x t^{\mathcal{P}X^*}(g) = \text{the least } A \subseteq X^* \text{ such that } t^{\mathcal{P}X^*}(g[x/A]) \subseteq A \tag{6}$$

is a component of the least $\bar{A}$ in $\mathcal{P}X^*$ with $\bar{A} \supseteq \bar{p}_t^{\mathcal{P}X^*}(\bar{A}, \bar{B})$ and $\bar{B} = g(\bar{y})$. By definition of $\mathcal{C}X^*$, $\bar{A}$ and hence $A := \mu x t^{\mathcal{P}X^*}(g)$ belong to $\mathcal{C}X^*$. By induction,

$$t^{\mathcal{C}X^*}(g[x/A]) = t^{\mathcal{P}X^*}(g[x/A]) \subseteq A,$$

so by Park's axioms, $\mu x t^{\mathcal{C}X^*}(g) \subseteq A = \mu x t^{\mathcal{P}X^*}(g)$. The converse $\mu x t^{\mathcal{P}X^*}(g) \subseteq \mu x t^{\mathcal{C}X^*}(g)$ follows from the induction hypothesis and the fact that $\mathcal{C}X^* \subseteq \mathcal{P}X^*$. ◀

4 Vector Versions

To denote the least simultaneous fixed point of an inequation system $\bar{x} \geq \bar{t}$, one might introduce terms $\mu\bar{x}\bar{t}$ where μ is an n-ary fixed-point operator, accompanied by projection functions π^n_i to get the i-th component of the fixed point. As has been observed by Bekić[1], de Bakker, Scott around 1970, it is sufficient to have a unary fixed point operator. Components of a simultaneous fixed point can be named by terms using the unary μ in a nested fashion.

▶ **Theorem 11** (Bekić [1]). *Let $(M, \leq)$ be a partially ordered set in which every countable increasing chain $\{a_i \mid i \in \mathbb{N}\}$ has a least upper bound, $\sum_{i\in\mathbb{N}} a_i$. Suppose $f, g : M^2 \to M$ are* **continuous** *in each component, i.e. map the least upper bound of countable increasing chains to the least upper bound of the images of the chain elements. Then the least solution of the system $(x, y) \geq (f(x, y), g(x, y))$ can be obtained by taking the least solution of each inequation separately, plugging it into the other equation and taking the least solutions of the resulting equations separately. I.e., the binary and unary fixed point operators are related by*

$$\mu(x, y)(f(x, y), g(x, y)) = (\mu x.f(x, \mu y.g(x, y)), \mu y.g(\mu x.f(x, y), y). \tag{7}$$

For an n-dimensional inequation system $\bar{x} \geq \bar{t}$, we define an n-tuple $\mu\bar{x}\bar{t}$ of μ-terms by recursively using Bekić's equations (7).

▶ **Definition 12.** ([10]) For vectors $\bar{t} = t_1, \ldots, t_n$ of terms and $\bar{x} = x_1, \ldots, x_n$ of pairwise different variables, we define the term vector $\mu\bar{x}\bar{t}$ as follows. If $n = 1$, then $\mu\bar{x}\bar{t} := \mu x_1 t_1$. If $n > 1$, $\bar{x} = (\bar{y}, \bar{z})$ and $\bar{t} = (\bar{r}, \bar{s})$ with term vectors $\bar{r}, \bar{s}$ of lengths $|\bar{y}|, |\bar{z}| < n$, then $\mu\bar{x}\bar{t}$ is[1]

$$\mu(\bar{y}, \bar{z})(\bar{r}, \bar{s}) := (\mu\bar{y}.\bar{r}[\bar{z}/\mu\bar{z}\bar{s}], \mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]). \tag{8}$$

A Chomsky algebra need not be closed under unions of countable increasing chains. For example, the set $\mathcal{C}X^*$ of all context-free languages over X has increasing chains of finite languages whose unions are not context-free. So one cannot apply Bekić's theorem literally to prove that $\mu\bar{x}\bar{t}$ denotes the least solution of $\bar{x} \geq \bar{t}$ in Chomsky algebras. Instead, we prove the Park axioms for term vectors by induction on the dimension.

For term vectors $\bar{s}$, $\bar{t}$ of the same dimension, let $\bar{s} = \bar{t}$ resp. $\bar{s} \leq \bar{t}$ be the conjunction of the equations resp. inequations of corresponding components. For $\bar{t} = (t_1, \ldots, t_n)$ we write $\bar{t}^M(g)$ for $(t_1^M(g), \ldots, t_n^M(g))$. The following property can be shown by induction on $|\bar{x}|$.

▶ **Lemma 13.** *If none of the variables in $\bar{x}$ occurs free in the terms of $\bar{s}$, then*

$$\mu\bar{x}\bar{t}[\bar{y}/\bar{s}] = \mu\bar{x}.\bar{t}[\bar{y}/\bar{s}],$$

A proof of this and the next lemma is given in the appendix.

▶ **Lemma 14.** *Let M be a Park μ-semiring. For all vectors $\bar{t}$ of terms and vectors $\bar{x}, \bar{y}$ of variables of the same dimension, the vector versions of (1) and (2),*

$$\bar{t}[\bar{x}/\mu\bar{x}\bar{t}] \quad \leq \quad \mu\bar{x}\bar{t}, \tag{9}$$

$$\bar{t}[\bar{x}/\bar{y}] \leq \bar{y} \quad \to \quad \mu\bar{x}\bar{t} \leq \bar{y}, \tag{10}$$

hold in M. Hence, for any Chomsky algebra M and valuation $g : X \to M$, $\mu\bar{x}\bar{t}^M(g)$ is the least $\bar{a}$ such that $\bar{t}^M(g[\bar{x}/\bar{a}]) \leq \bar{a}$.

One shows (9) and (10) simultaneously with a vector version of the substitution property, using lemma 13: in any Park μ-semininiring M and for any $g : X \to M$,

$$\mu\bar{x}\bar{t}[\bar{y}/\bar{s}]^M(g) = \mu\bar{x}\bar{t}^M(g[\bar{y}/\bar{s}^M(g)]), \qquad \text{if no variable of } \bar{x} \text{ is free in the terms } \bar{s}. \tag{11}$$

▶ **Corollary 15.** *If M is a Park μ-semiring, the vector version of the μ-rule holds: for vectors $\bar{s}$, $\bar{t}$ of terms and $\bar{x}$ of different variables, all of the same dimension, if $\bar{s}^M \leq \bar{t}^M$, then $\mu\bar{x}\bar{s}^M \leq \mu\bar{x}\bar{t}^M$.*

5 μ-Continuity

The usual way to compute the simultaneous least fixed point $\bar{A}$ of $\bar{x} \geq \bar{t}(\bar{x}, \bar{y})$ in $\mathcal{P}X^*$ relative to $\bar{B}$ is to approximate it by its finite stages $\bar{A}_m$ and (componentwise) take their union, i.e.

$$\bar{A} := \bigcup_{m \in \mathbb{N}} \bar{A}_m, \qquad \text{where} \quad \bar{A}_0 := \bar{\emptyset}, \qquad \bar{A}_{m+1} := \bar{t}^{\mathcal{P}X^*}(\bar{A}_m, \bar{B}).$$

As shown below, the continuity of $+$ and $\cdot$ in $\mathcal{P}X^*$ imply that $\bar{A}$ equals $\mu\bar{x}\bar{t}^{\mathcal{P}X^*}(\bar{B})$.

[1] To distinguish $\mu xt[y/s]$ from $\mu x(t[y/s])$, we write $\mu x.t[y/s]$ for the latter, using . to save the brackets of the metalanguage. We have no . in the object language's μ-terms, preferring $\mu x(t + s)$ over $\mu x.t + s$.

▶ **Lemma 16.** *In $M = \mathcal{P}X^*$ or $M = \mathcal{C}X^*$, all term functions are* **continuous***: for any term t, valuation $g : X \to M$ and increasing chain $\{A_i \mid i \in \mathbb{N}\}$ of elements with union in M,*

$$t^M(g[y/\bigcup_{i\in\mathbb{N}} A_i]) = \bigcup_{i\in\mathbb{N}} t^M(g[y/A_i]). \tag{12}$$

Proof. This is clear when t is one of $x, y, 0, 1$. For $(t_1 + t_2)$ and $(t_1 \cdot t_2)$ use induction, the continuity of $+$ and $\cdot$ in M, and the fact that the A_i form an increasing chain. For μxt, by the monotonicity of μxt^M it is sufficient to show

$$\mu xt^M(g[y/\bigcup_{i\in\mathbb{N}} A_i]) \subseteq \bigcup_{i\in\mathbb{N}} \mu xt^M(g[y/A_i]). \tag{13}$$

Let $B_i := \mu xt^M(g[y/A_i])$ for $i \in \mathbb{N}$, and let A and B be the unions of the increasing chains $\{A_i \mid i \in \mathbb{N}\}$ and $\{B_i \mid i \in \mathbb{N}\}$, respectively. If $M = \mathcal{P}X^*$, then $B \in M$ and

$$\begin{array}{rcll} t^M(g[y/A][x/B]) & = & \bigcup_{i,j\in\mathbb{N}} t^M(g[y/A_i][x/B_j]) & \text{(by induction)} \\ & = & \bigcup_{i\in\mathbb{N}} t^M(g[y/A_i][x/B_i]) & \text{(by monotonicity, increasing chains)} \\ & = & \bigcup_{i\in\mathbb{N}} t[x/\mu xt]^M(g[y/A_i]) & \text{(by the substitution property)} \\ & \subseteq & \bigcup_{i\in\mathbb{N}} \mu xt^M(g[y/A_i]) & \text{(by Park's inequation)} \\ & = & \bigcup_{i\in\mathbb{N}} B_i = B. & \end{array}$$

By the Park rule, $\mu xt^M(g[y/A]) \subseteq B$ follows, which proves (13) for $M = \mathcal{P}X^*$. By lemma 10, we can transfer (12) from $M = \mathcal{P}X^*$ to $M = \mathcal{C}X^*$. ◀

The power set semiring $(\mathcal{P}M, \cup, \cdot, \emptyset, \{1\})$ of a monoid $(M, \cdot^M, 1)$, a generalization of Example 6, and the semiring *Rel* (S) of all binary relations on a set S are **continuous**: their operations $+$ and $\cdot$ are continuous and the partial order $\subseteq$ is complete (all directed subsets have suprema). The semiring $\mathcal{C}X^*$ of context-free languages over X is *not* complete.

▶ **Definition 17** ([11]). A Chomsky algebra M is **μ-continuous**, if for all μ-terms $t(x, \bar{y})$ and all $a, b, \bar{c} \in M$,

$$a \cdot \mu xt^M[\bar{y}/\bar{c}] \cdot b = \sum\{a \cdot (mxt)^M[\bar{y}/\bar{c}] \cdot b \mid m \in \mathbb{N}\}, \tag{14}$$

where the term mxt, the m-fold iteration of t in x, is defined by $0xt := 0$, $(m+1)xt := t[x/mxt]$. In particular, the supremum on the right hand side of (14) must exist.

As has been mentioned in [11] without proof, we have:

▶ **Theorem 18.** *The Chomsky algebra $\mathcal{C}X^*$ of all context-free languages over X is μ-continuous. In particular, for all terms t and valuations $g : X \to \mathcal{C}X^*$,*

$$\mu xt^{\mathcal{C}X^*}(g) = \bigcup\{mxt^{\mathcal{C}X^*}(g) \mid m \in \mathbb{N}\}.$$

Proof. We first consider $M = \mathcal{P}X^*$ and then transfer the result to $\mathcal{C}X^*$ by Lemma 10. From $0xt^M(g) \subseteq 1xt^M(g)$ and $0xt^M(g) \subseteq \mu xt^M(g)$, by induction we get $mxt^M(g) \subseteq (m+1)xt^M(g) \subseteq \mu xt^M(g)$ for all m, using monotonicity, the substitution property, and the Park inequation. Therefore, $\bigcup\{mxt^M(g) \mid m \in \mathbb{N}\} \subseteq \mu xt^M(g)$. For the converse,

$$\begin{array}{rcll} t^M(g[x/\bigcup\{(mxt)^M(g) \mid m \in \mathbb{N}\}]) & = & \bigcup\{t^M(g[x/(mxt)^M(g)]) \mid m \in \mathbb{N}\} & \text{(by (12))} \\ & = & \bigcup\{t[x/mxt]^M(g) \mid m \in \mathbb{N}\} & \\ & = & \bigcup\{((m+1)xt)^M(g) \mid m \in \mathbb{N}\} & \\ & = & \bigcup\{(mxt)^M(g) \mid m \in \mathbb{N}\}. & \end{array}$$

Hence, by the Park rule, $\mu xt^M(g) \subseteq \bigcup\{(mxt)^M(g) \mid m \in \mathbb{N}\}$. It follows that for $A, B \in M$, $A \cdot \mu xt^M(g) \cdot B = \bigcup\{A \cdot mxt^M(g) \cdot B \mid m \in \mathbb{N}\}$. ◀

The iterative computation of least fixed points is used in [11] to interpret μ-terms in the semiring $\mathcal{C}X^*$:

▶ **Example 19.** ([11]) The **canonical interpretation** L of μ-terms over X in $\mathcal{C}X^*$ is

$$\begin{array}{lcl@{\qquad}rcl} L(x) &=& \{x\} & L(s+t) &=& L(s)\cup L(t)\\ L(0) &=& \emptyset & L(s\cdot t) &=& \{uv \mid u\in L(s), v\in L(t)\}\\ L(1) &=& \{\epsilon\} & L(\mu xt) &=& \bigcup\{L(mxt)\mid m\in\mathbb{N}\}. \end{array}$$

By Theorem 18, the canonical interpretation L of μ-terms in the semiring $\mathcal{C}X^*$ coincides with the interpretation in the Chomsky algebra $\mathcal{C}X^*$ under the valuation L, i.e. $L(t) = t^{\mathcal{C}X^*}(L)$ for all terms t. This can be proven by induction on the well-founded relation $\prec$ of terms from [7] were $t_i \prec (t_1+t_2)$, $t_i \prec (t_1\cdot t_2)$, and $mxt \prec \mu xt$.

▶ Remark. Grathwohl e.a. [11], Theorem 3.2, prove that an idempotent semiring M with an interpretation of μ-terms satisfying the μ-continuity condition also satisfies the Park axioms, hence makes $\mu xt^M(g)$ the least fixed point of $x \geq t$ in M relative to g.

For a vector of terms $\bar{t} = (t_1,\dots,t_n)$, put $L(\bar{t}) := (L(t_1),\dots,L(t_n))$. The m-th iteration of $\bar{t}$ in $\bar{x}$ can be expressed syntactically by a term vector $m\bar{x}\bar{t}$, where $|\bar{x}| = |\bar{t}|$ and

$$0\bar{x}\bar{t} := \bar{0}, \qquad (m+1)\bar{x}\bar{t} := \bar{t}[\bar{x}/m\bar{x}\bar{t}]. \tag{15}$$

Vector versions of substitution and other properties of term functions in $\mathcal{C}X^*$ will be used for L, based on (11) and theorem 18.

We now prove a vector version of (the main part of) the μ-continuity condition for $\mathcal{C}X^*$, i.e. that the least fixed point of a system $\bar{x} \geq \bar{t}$, as determined by the Bekić-Scott equations embodied in $\mu\bar{x}\bar{t}$, is the supremum of its approximations by iterations $m\bar{x}\bar{t}$. In particular, Bekić's reduction works in $\mathcal{C}X^*$, where not every ascending chain has a supremum.

▶ **Lemma 20.** *For all vectors $\bar{x},\bar{y},\bar{z}$ of pairwise distinct variables and vectors $\bar{t},\bar{r},\bar{s}$ of μ-terms such that $|\bar{x}| = |\bar{t}|$, $|\bar{y}| = |\bar{r}|$ and $|\bar{z}| = |\bar{s}|$, we have:*
1. $L(\mu\bar{x}\bar{t}) = \bigcup\{L(m\bar{x}\bar{t}) \mid m\in\mathbb{N}\}$,
2. $L(\bar{s}[\bar{x}/\mu\bar{x}\bar{t}]) = \bigcup\{L(\bar{s}[\bar{x}/m\bar{x}\bar{t}]) \mid m\in\mathbb{N}\}$,
3. $L(\mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]) = \bigcup\{L(m\bar{z}.\bar{s}[\bar{y}/k\bar{y}\bar{r}]) \mid m,k\in\mathbb{N}\}$.

Proof. Proof by simultaneous induction on the vector length of $|\bar{x}| = |\bar{y}|+|\bar{z}|$ with $|\bar{y}|,|\bar{z}| < |\bar{x}|$.
1. For $|\bar{x}| = 1$, the claim holds by definition of L, so suppose $|\bar{x}| > 1$.

   (a) $\bigcup_m L(m\bar{x}\bar{t}) \subseteq L(\mu\bar{x}\bar{t})$: Clearly, $L(0\bar{x}\bar{t}) = \bar{\emptyset} \subseteq L(\mu\bar{x}\bar{t})$. If $L(m\bar{x}\bar{t}) \subseteq L(\mu\bar{x}\bar{t})$, then

$$\begin{array}{rcll} L((m+1)\bar{x}\bar{t}) &=& L(\bar{t}[\bar{x}/m\bar{x}\bar{t}]) & ((15))\\ &=& L[\bar{x}/L(m\bar{x}\bar{t})](\bar{t}) & \text{(substitution property)}\\ &\subseteq& L[\bar{x}/L(\mu\bar{x}\bar{t})](\bar{t}) & \text{(induction hypothesis)}\\ &=& L(\bar{t}[\bar{x}/\mu\bar{x}\bar{t}]) & \text{(substitution property)}\\ &\subseteq& L(\mu\bar{x}\bar{t}) & \text{(Lemma 14, (9))}. \end{array}$$

   (b) $L(\mu\bar{x}\bar{t}) \subseteq \bigcup\{L(m\bar{x}\bar{t}) \mid m\in\mathbb{N}\}$. By induction, claim 3. holds with $|\bar{y}|,|\bar{z}| < |\bar{x}|$, so

$$\begin{array}{rcll} L(\mu\bar{x}\bar{t}) &=& L(\mu(\bar{y},\bar{z})(\bar{r},\bar{s})) & (\bar{x} = (\bar{y},\bar{z}), \bar{t} = (\bar{r},\bar{s}))\\ &=& L(\mu\bar{y}.\bar{r}[\bar{z}/\mu\bar{z}\bar{s}], \mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]) & \text{(definition of } \mu\bar{x}\bar{t})\\ &=& (\bigcup_{m,k} L(m\bar{y}.\bar{r}[\bar{z}/k\bar{z}\bar{s}]), \bigcup_{m,k} L(m\bar{z}.\bar{s}[\bar{y}/k\bar{y}\bar{r}])) & \text{(induction hypothesis 3.)}\\ &=& \bigcup_{m,k} L(m\bar{y}.\bar{r}[\bar{z}/k\bar{z}\bar{s}], m\bar{z}.\bar{s}[\bar{y}/k\bar{y}\bar{r}]) & \text{(monotonicity)}. \end{array}$$

It is therefore sufficient to show

$$L(m\bar{y}.\bar{r}[\bar{z}/k\bar{z}\bar{s}], m\bar{z}.\bar{s}[\bar{y}/k\bar{y}\bar{r}]) \subseteq L(m_k(\bar{y},\bar{z})(\bar{r},\bar{s})), \quad \text{for } m_k = m(k+1). \tag{16}$$

Put $(\bar{A}_n, \bar{B}_n) := L(n\bar{x}\bar{t}) = L(n(\bar{y},\bar{z})(\bar{r},\bar{s}))$. By induction on k one obtains

$$L[\bar{y}/\bar{A}_n](k\bar{z}\bar{s}) \subseteq \bar{B}_{n+k} \quad \text{and} \quad L[\bar{z}/\bar{B}_n](k\bar{y}\bar{r}) \subseteq \bar{A}_{n+k}, \tag{17}$$

using monotonicity of $\bar{s}^{\mathcal{C}X^*}$ and $\bar{r}^{\mathcal{C}X^*}$ and the substitution property. To see (16), put $\bar{r}_k(\bar{y}) := \bar{r}[\bar{z}/k\bar{z}\bar{s}]$ and $\bar{s}_k(\bar{z}) := \bar{s}[\bar{y}/k\bar{y}\bar{r}]$. Suppose that for some m we have

$$L(m\bar{y}\bar{r}_k, m\bar{z}\bar{s}_k) \subseteq (\bar{A}_n, \bar{B}_n) \qquad \text{with } n = m(k+1). \tag{18}$$

which is clearly true for $m = 0$. Then, using the monotonicity of $\bar{r}^{\mathcal{C}X^*}$ and $\bar{s}^{\mathcal{C}X^*}$,

$$\begin{array}{lll} L((m+1)\bar{y}\bar{r}_k, (m+1)\bar{z}\bar{s}_k) & & \\ \quad = L(\bar{r}_k[\bar{y}/m\bar{y}\bar{r}_k], \bar{s}_k[\bar{z}/m\bar{z}\bar{s}_k]) & & \text{(definition)} \\ \quad \subseteq (L[\bar{y}/\bar{A}_n](\bar{r}_k), L[\bar{z}/\bar{B}_n](\bar{s}_k)) & & \text{(substitution, (18))} \\ \quad \subseteq (L[\bar{y}/\bar{A}_n, \bar{z}/\bar{B}_{n+k}](\bar{r}), L[\bar{y}/\bar{A}_{n+k}, \bar{z}/\bar{B}_n](\bar{s})) & & \text{((17))} \\ \quad \subseteq (L[\bar{y}/\bar{A}_{n+k}, \bar{z}/\bar{B}_{n+k}](\bar{r}), L[\bar{y}/\bar{A}_{n+k}, \bar{z}/\bar{B}_{n+k}](\bar{s})) & & \text{(increasing chains)} \\ \quad = (\bar{A}_{n+k+1}, \bar{B}_{n+k+1}) & & \\ \quad = (\bar{A}_{(m+1)(k+1)}, \bar{B}_{(m+1)(k+1)}). & & \end{array}$$

Thus, we have (16), and hence $L(\mu\bar{x}\bar{t}) \subseteq \bigcup\{L(m\bar{x}\bar{t}) \mid m \in \mathbb{N}\}$.

2. This is a vector version of lemma 2.5 of [11]. Since $\bigcup$ on $(\mathcal{C}X^*)^{|\bar{t}|}$, substitution $\bar{s}[\bar{x}/\mu\bar{x}\bar{t}]$ and evaluation $L(\bar{s})$ is done componentwise, it is sufficient to consider $|\bar{s}| = 1$.
By 1., $L(\mu\bar{x}\bar{t}) = \bigcup\{\bar{A}_m \mid m \in \mathbb{N}\}$, where $\bar{A}_m = L(m\bar{x}\bar{t})$ is a vector of (definable, hence) context-free languages. By induction, $\bar{A}_m \subseteq \bar{A}_{m+1}$ for all m, componentwise. Then

$$\begin{array}{rcll} L(s[\bar{x}/\mu\bar{x}\bar{t}]) & = & L[\bar{x}/L(\mu\bar{x}\bar{t})](s) & \text{(substitution property)} \\ & = & L[\bar{x}/\bigcup\{L(m\bar{x}\bar{t}) \mid m \in \mathbb{N}\}](s) & \text{(claim 1.)} \\ & = & \bigcup\{L[\bar{x}/L(m\bar{x}\bar{t})](s) \mid m \in \mathbb{N}\} & \text{(Lemma 16)} \\ & = & \bigcup\{L(s[\bar{x}/m\bar{x}\bar{t}]) \mid m \in \mathbb{N}\} & \text{(substitution property)} \end{array}$$

3. To prove claim 1. for $|\bar{x}| > 1$, we used claim 3. with $|\bar{y}| + |\bar{z}| = |\bar{x}|$ and $0 < |\bar{y}|, |\bar{z}| < |\bar{x}|$. Hence, to prove claim 3. for dimension $|\bar{y}| + |\bar{z}|$, we have claim 1. for smaller dimensions as induction hypothesis. Therefore, an application of 1. (with $\mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]$ as $\mu\bar{x}\bar{t}$) gives

$$L(\mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]) = \bigcup_m L(m\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]).$$

To finish the proof of claim 3., it is now sufficient to show for all m:

$$L(m\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]) = \bigcup_k L(m\bar{z}.\bar{s}[\bar{y}/k\bar{y}\bar{r}]). \tag{19}$$

Equation (19) is clear for $m = 0$. Assume it is true for m. Then we proceed as follows, using abbreviations $\bar{s}_\omega := \bar{s}[\bar{y}/\mu\bar{y}\bar{r}]$ and $\bar{s}_k := \bar{s}[\bar{y}/k\bar{y}\bar{r}]$:

$$\begin{array}{rcll} L((m+1)\bar{z}\bar{s}_\omega) & = & L(\bar{s}_\omega[\bar{z}/m\bar{z}\bar{s}_\omega]) & \text{(definition)} \\ & = & L[\bar{z}/L(m\bar{z}\bar{s}_\omega)](\bar{s}_\omega) & \text{(substitution property)} \\ & = & \bigcup_k L[\bar{z}/L(m\bar{z}\bar{s}_\omega)](\bar{s}_k) & \text{(induction hypothesis 2. for } |\bar{y}|) \\ & = & \bigcup_k L[\bar{z}/\bigcup_{k'} L(m\bar{z}\bar{s}_{k'})](\bar{s}_k) & \text{(induction hypothesis (19))} \\ & = & \bigcup_k \bigcup_{k'} L[\bar{z}/L(m\bar{z}\bar{s}_{k'})](\bar{s}_k) & \text{(Lemma 16, componentwise)} \\ & = & \bigcup_k L[\bar{z}/L(m\bar{z}\bar{s}_k)](\bar{s}_k) & \text{(monotonicity)} \\ & = & \bigcup_k L(\bar{s}_k[\bar{z}/m\bar{z}\bar{s}_k]) & \text{(substitution property)} \\ & = & \bigcup_k L((m+1)\bar{z}\bar{s}_k) & \text{(definition).} \end{array}$$

Hence we have (19) for all m. ◀

By the following lemma from [11] we can transfer the vector version of the μ-continuity condition from L resp. $\mathcal{C}X^*$ to an arbitrary μ-continuous Chomsky algebra.

▶ **Lemma 21** ([11], Lemma 3.1)**.** *Let M be a μ-continuous Chomsky algebra, $g : X \to M$ a valuation of terms in M, $h : X \to \mathcal{C}X^*$ be a valuation in the algebra $\mathcal{C}X^*$ of context-free languages over X, such that, for all $x \in X$ and μ-terms s, u,*

$$(sxu)^M(g) = \sum\{(syu)^M(g) \mid y \in h(x)\}.$$

Then, for all μ-terms s, t, u,

$$(stu)^M(g) = \sum\{(syu)^M(g) \mid y \in t^{\mathcal{C}X^*}(h)\}.$$

Notice that the canonical interpretation $L : X \to \mathcal{C}X^*$ satisfies the assumptions on the valuation h of the lemma.

▶ **Corollary 22.** *For any vectors $\bar{s}, \bar{t}, \bar{u}$ of μ-terms of equal length, any μ-continuous Chomsky-algebra M and valuation $g : X \to M$, under componentwise multiplication and supremum,*

$$(\bar{s} \cdot \bar{t} \cdot \bar{u})^M(g) = \sum\{(\bar{s} \cdot \bar{w} \cdot \bar{u})^M(g) \mid \bar{w} \in L(\bar{t})\}, \tag{20}$$

where, for $x_1 \cdots x_k \in X^*$, $(x_1 \cdots x_k)^M(g) := g(x_1) \cdot^M \ldots \cdot^M g(x_k)$.

This is clear since for each component, the equation follows from the lemma. The vector-version of the μ-continuity condition follows:

▶ **Corollary 23.** *Let M be a μ-continuous Chomsky algebra and $g : X \to M$. Then*

$$\bar{a} \cdot \mu\bar{x}\bar{t}^M(g) \cdot \bar{b} = \sum\{\bar{a} \cdot m\bar{x}\bar{t}^M(g) \cdot \bar{b} \mid m \in \mathbb{N}\},$$

for any term vector $\bar{t}$ and vectors $\bar{a}, \bar{b}$ of elements of M of the same length as $\bar{t}$.

Proof. We may assume that there are suitable vectors $\bar{s}, \bar{u}$ of terms such that $\bar{a} = \bar{s}^M(g)$ and $\bar{b} = \bar{u}^M(g)$. Then

$$\begin{array}{rcll}
\bar{a} \cdot \mu\bar{x}\bar{t}^M(g) \cdot \bar{b} & = & (\bar{s} \cdot \mu\bar{x}\bar{t} \cdot \bar{u})^M(g) & \\
& = & \sum\{(\bar{s} \cdot \bar{w} \cdot \bar{u})^M(g) \mid \bar{w} \in L(\mu\bar{x}\bar{t})\} & \text{(equation (20))} \\
& = & \sum\{(\bar{s} \cdot \bar{w} \cdot \bar{u})^M(g) \mid \bar{w} \in \bigcup\{L(m\bar{x}\bar{t}) \mid m \in \mathbb{N}\}\} & \text{(Lemma 20, 1.)} \\
& = & \sum\bigcup\{\{(\bar{s} \cdot \bar{w} \cdot \bar{u})^M(g) \mid \bar{w} \in L(m\bar{x}\bar{t})\} \mid m \in \mathbb{N}\} & \\
& = & \sum\{\sum\{(\bar{s} \cdot \bar{w} \cdot \bar{u})^M(g) \mid \bar{w} \in L(m\bar{x}\bar{t})\} \mid m \in \mathbb{N}\} & \\
& = & \sum\{(\bar{s} \cdot m\bar{x}\bar{t} \cdot \bar{u})^M(g) \mid m \in \mathbb{N}\} & \text{(equation (20))} \\
& = & \sum\{\bar{a} \cdot m\bar{x}\bar{t}^M(g) \cdot \bar{b} \mid m \in \mathbb{N}\}. &
\end{array}$$

◀

In particular, μ-continuity of M implies that the least solution $\mu\bar{x}\bar{t}^M(g)$ of a system $\bar{t} \leq \bar{x}$ relative to g is the supremum of the finite iterations $m\bar{x}\bar{t}^M(g)$:

▶ **Corollary 24.** *Let M be a μ-continuous Chomsky algebra and $g : X \to M$. Then*

$$\mu\bar{x}\bar{t}^M(g) = \sum\{m\bar{x}\bar{t}^M(g) \mid m \in \mathbb{N}\}.$$

6 Closure under Matrix Rings

To prove the main result, that the square matrices over a μ-continuous Chomsky algebra form a μ-continuous Chomsky algebra, we recall the case of Park μ-semirings:

▶ **Theorem 25** ([10], Theorem 7.6). *If M is a Park μ-semiring, so is $Mat_{n,n}(M)$.*

Proof. Let M be a Park μ-semiring and $N = Mat_{n,n}(M)$ the set of $n \times n$ matrices of elements of M. Equipped with the usual matrix operations, $(N, +, \cdot, 0, 1)$ is a semiring, because M is. To define the term functions $t^N : (X \to N) \to N$, for each variable $x \in X$ we fix n^2 distinct variables $x_{i,j}$, $1 \leq i, j \leq n$, which also have to be distinct from all $y_{i,j}$ for variables $y \neq x$. For each term t, define a vector t' of n^2 terms recursively by

$$\begin{array}{rcl} x' &:=& (x_{i,j}), \\ 0' &:=& 0_{n,n}, \\ 1' &:=& 1_{n,n}, \end{array} \qquad \begin{array}{rcl} (s+t)' &:=& s' + t', \\ (s \cdot t)' &:=& s' \cdot t', \\ (\mu x t)' &:=& \mu x' t', \end{array}$$

where $0_{n,n}$ and $1_{n,n}$ are the zero and unit matrices, $+$ and $\cdot$ are the usual matrix addition and multiplication operations applied to matrices of terms, and $\mu x' t'$ is the term vector $\mu \bar{x} \bar{t}$ defined recursively by Bekić's equation (8) from x' and t'.

Any valuation $g : X \to N$ is obtained from a valuation $\hat{g} : X \to M$ by $g(x) = (a_{i,j})$, where $a_{i,j} = \hat{g}(x_{i,j})$ for $1 \leq i, j \leq n$. Define the term function t^N by

$$t^N(g) := ({t'_{i,j}}^M(\hat{g})), \qquad \text{where } t'_{i,j} \text{ is the } (i,j)\text{-th entry of } t'. \tag{21}$$

Concerning the properties for partially ordered μ-semirings, those for 0^N, 1^N, x^N, $(s+t)^N$ and $(s \cdot t)^N$ are immediate from the definition. The μ-rule holds in N, since, by corollary 15, the vector version of the μ-rule holds in M.

The monotonicity of the term function t^N follows from the monotonicity of the term functions ${t'}^M_{i,j}$. Likewise for the coincidence property. By Lemma 13 and Lemma 14, the vector versions of the substitution property and Park axioms hold in M; these imply that the substitution property and the Park axioms hold in N. ◀

▶ **Theorem 26.** *If M is a (μ-continuous) Chomsky algebra, so is $Mat_{n,n}(M)$.*

Proof. Let M be a Chomsky algebra and $N := Mat_{n,n}(M)$. By Corollary 9, M is an idempotent Park μ-semiring. Hence, by Theorem 25, N also is, and thus, by Lemma 14, N satisfies the vector versions of Park's axioms. In particular, every system $\bar{x} \geq \bar{p}(\bar{x}, \bar{y})$ of polynomial inequations does have, for any parameters $\bar{B} \in N$, a least solution in N, $(\mu \bar{x} \bar{p})^N(\bar{B})$. Hence, N is a Chomsky algebra.

Suppose, in addition, M is μ-continuous. To show that N is μ-continuous, we assume $n > 1$, as $Mat_{1,1}(M)$ is isomorphic to M. Let $A = (a_{i,j}), B = (b_{i,j}) \in N$, $t(x, \bar{y})$ a μ-term and $g : X \to N$ coming from a valuation $\hat{g} : X \to M$ as in (21). In order to show

$$A \cdot \mu x t^N(g) \cdot B = \sum \{A \cdot m x t^N(g) \cdot B \mid m \in \mathbb{N}\}, \tag{22}$$

we first show

$$\mu x t^N(g) = \sum \{m x t^N(g) \mid m \in \mathbb{N}\}. \tag{23}$$

By definition, $\mu x t^N(g) = ((\mu x t)'^M_{i,j}(\hat{g}))$, where $(\mu x t)' = \mu x' t'$ is obtained from the matrices

$$x' = \begin{pmatrix} x_{1,1} & \dots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \dots & x_{n,n} \end{pmatrix} \quad \text{and} \quad t' = \begin{pmatrix} t_{1,1} & \dots & t_{1,n} \\ \vdots & \ddots & \vdots \\ t_{n,1} & \dots & t_{n,n} \end{pmatrix}$$

of pairwise different variables $x_{i,j}$ and of μ-terms $t_{i,j}(x', \bar{y}')$ (with variables $y' = (y_{i,j})$ for parameters y in $\bar{y}$) according to Definition 12 for the inequation system $t' \leq x'$ of size n^2. Accordingly, we have a square term matrix $(mxt)' = mx't'$ for the m-th iteration of t' in x'. By Corollary 24, we get (23):

$$\mu xt^N(g) = ((\mu xt)'^M_{i,j}(\hat{g})) = \sum\{((mxt)'_{i,j}{}^M(\hat{g})) \mid m \in \mathbb{N}\} = \sum\{mxt^N(g) \mid m \in \mathbb{N}\}.$$

Concerning (22), it is sufficient to consider the (i,j)-th entry and use Corollary 23:

$$\begin{aligned}
&(A \cdot \mu xt^N(g) \cdot B)_{i,j} \\
&= \sum_{k,l \leq n} a_{i,k} \cdot^M (\mu xt^N(g))_{k,l} \cdot^M b_{l,j} \\
&= \sum_{k,l \leq n} a_{i,k} \cdot^M (\mu xt)'^M_{k,l}(\hat{g}) \cdot^M b_{l,j} \\
&= \sum_{k,l \leq n} \sum\{a_{i,k} \cdot^M (mxt)'^M_{k,l}(\hat{g}) \cdot^M b_{l,j} \mid m \in \mathbb{N}\} \quad \text{(Corollary 23)} \\
&= \sum\{\sum_{k,l \leq n} a_{i,k} \cdot^M (mxt)'^M_{k,l}(\hat{g}) \cdot^M b_{l,j} \mid m \in \mathbb{N}\} \\
&= \sum\{\sum_{k,l \leq n} a_{i,k} \cdot^M ((mxt)^N(g))_{k,l} \cdot^M b_{l,j} \mid m \in \mathbb{N}\} \\
&= \sum\{(A \cdot mxt^N(g) \cdot B)_{i,j} \mid m \in \mathbb{N}\}
\end{aligned}$$

Hence, N also is μ-continuous. ◀

7 Open Questions

In analogy to $\mathcal{C}X^*$, we can define the semiring of context-free subsets $\mathcal{C}M$ of an arbitrary monoid M by closing the semiring $\mathcal{F}M$ of finite subsets of M under least solutions in $\mathcal{P}M$ of polynomial inequations $\bar{x} \geq \bar{p}(\bar{x}, \bar{y})$ with parameters. Hopkins [5], [6] gives an elegant algebraic generalization of formal language theory, where he defines, for each monoid M, a dioid (= idempotent semiring) $\mathcal{C}M$ of context-free subsets of M in a different way. An open question is whether the two definitions of $\mathcal{C}M$ agree, and whether μ-continuous Chomsky algebras coincide with Hopkins' $\mathcal{C}$-dioids, just as *-continuous Kleene algebras coincide with his $\mathcal{R}$-dioids [6]. Hopkins left open whether the class of $\mathcal{C}$-dioids is closed under the matrix ring construction.

Another open question is whether $\mathcal{C}M$ can be constructed as an ideal-closure for a suitable notion of $\mathcal{C}$- or μ-ideal of M, if M is an idempotent semiring or a Kleene algebra.

Acknowledgements. I thank Mark Hopkins for comments and questions on drafts of this paper, for providing unpublished material of his extending [5], [6], and for email discussions on the subject. Thanks also to Dexter Kozen for making me aware of [11] and reviving my interest in the subject.

References

1 Hans Bekić. Definable operations in general algebras, and the theory of automata and flowcharts. In C. B. Jones, editor, *Programming Languages and Their Definition*, volume 177 of *LNCS*, pages 30–55. Springer-Verlag, Berlin Heidelberg, 1984.
2 J. H. Conway. *Regular Algebra and Finite Machines.* Chapman and Hall, London, 1971.

3 J.W. de Bakker and D. Scott. A theory of programs. IBM Seminar, Vienna, 1969.

4 J. Gruska. A characterization of context-free languages. *Journal of Computer and System Sciences*, 5:353–364, 1971.

5 Mark Hopkins. The algebraic approach I: The algebraization of the Chomsky hierarchy. In *10th Int. Conf. on Relational Methods in Computer Science and 5th Int. Conf. on Applications of Kleene Algebra*, volume 4988 of *LNCS*, pages 155–172. Springer-Verlag, Berlin Heidelberg, 2008.

6 Mark Hopkins. The algebraic approach II: The algebraization of the Chomsky hierarchy. In *10th Int. Conf. on Relational Methods in Computer Science and 5th Int. Conf. on Applications of Kleene Algebra*, volume 4988 of *LNCS*, pages 173–190. Springer-Verlag, Berlin Heidelberg, 2008.

7 Dexter Kozen. Results in the propositional μ-calculus. *Theoretical Computer Science*, 27(3):333–345, 1983.

8 Dexter Kozen. A completeness theorem for Kleene algebras and the algebra of regular events. In *6th Annual Symposium on Logic in Computer Science, July 15–18, 1991, Amsterdam*, Los Alamitos, CA, 1991. Computer Society Press.

9 H. Leiß. Towards Kleene Algebra with Recursion. In E. Börger, G. Jäger, H. Kleine-Büning, and M. M. Richter, editors, *CSL'91. 5th Workshop on Computer Science Logic. Berne, Switzerland, October 7–11, 1991*, pages 242–256. Springer LNCS 626, 1991.

10 Hans Leiß and Zoltan Ésik. Algebraically complete semirings and Greibach normal form. *Annals of Pure and Applied Logic*, 133:173–203, 2005.

11 Niels Bjørn Bugge Grathwohl, Fritz Henglein, and Dexter Kozen. Infinitary axiomatization of the equational theory of context-free languages. In D. Baelde and A. Carayol, editors, *Fixed Points in Computer Science (FICS 2013)*, volume 126 of *EPTCS*, pages 44–55, 2013. doi:10.4204/EPTCS.126.4.

A Appendix

Proof of claims following the definition of Park μ-semiring. For $t[x/\mu xt] = \mu xt$, we only need the $\geq$-inequation:

$$\begin{array}{rrcll} & t[x/\mu xt] & \leq & \mu xt & \text{(by (1))} \\ \Rightarrow & t[x/t[x/\mu xt]] & \leq & t[x/\mu xt] & \text{(by monotonicity)} \\ \Rightarrow & \mu xt & \leq & t[x/\mu xt] & \text{(by (2)).} \end{array}$$

For $\mu y.t[x/y] = \mu xt$ with $y \notin \mathit{free}(t)$,

$$\begin{array}{rrcll} & t[x/\mu y.t[x/y]] & = & t[x/y][y/\mu y.t[x/y]] & \\ & & \leq & \mu y.t[x/y] & \text{(by (1))} \\ \Rightarrow & \mu xt & \leq & \mu y.t[x/y] & \text{(by (2)).} \end{array}$$

By symmetry, we have $\mu y.t[x/y] \leq \mu x.t[x/y][y/x] \equiv \mu xt$. ◀

Proof of Lemma 8, properties of term functions. It remains to be seen that the properties of term functions in partially ordered μ-semirings hold in M for terms of μ-depth $\leq k$. We consider the substitution property. Suppose $r[x/s]$ has μ-depth $\leq k$, and the substitution property holds for terms of μ-depth $< k$. If x does not occur freely in r, then $r[x/s] = r$ and g agrees with $g[x/s^M(g)]$ on $\mathit{free}(r)$, hence $r[x/s]^M(g) = r^M(g) = r^M(g[x/s^M(g)])$. Otherwise, if r is one of $0, 1, y, (r_1 + r_2), (r_1 \cdot r_2)$, the claim holds by induction. So suppose $r = \mu yt$ with x different from y. We may assume that $y \notin \mathit{free}(s)$. Then $r[x/s] = \mu y.t[x/s]$, and by induction, $t[x/s]^M(h) = t^M(h[x/s^M(h)])$ for all $h : X \to M$. Therefore, the least $a \in M$ with $t[x/s]^M(h[y/a]) \leq a$, is the least a with

$$t^M(h[x/s^M(h)][y/a]) = t^M(h[y/a][x/s^M(h[y/a])]) = t[x/s]^M(h[y/a]) \leq a,$$

which is $\mu y t^M(h[x/s^M(h)])$. Hence, $r[x/s]^M(h) = r^M(h[x/s^M(h)])$.

The other properties of term functions are shown similarly. Since M is an idempotent semiring, $+$ and $\cdot$ are monotone, hence t^M is monotone for any algebraic term t. By induction on the μ-depth, $\mu x t^M$ is monotone, and if $s^M \leq t^M$, we get $\mu x s^M \leq \mu x t^M$, since any a with $t^M(g[x/a]) \leq a$ satisfies $s^M(g[x/a]) \leq a$. It is also clear that $t^M(g) = t^M(h)$ if g agrees with h on *free*(t). Hence, M is a partially ordered μ-semiring. ◀

Proof of Lemma 13. By induction on $|\bar{x}|$. For $|\bar{x}| = 1$, we have $\mu x t[\bar{y}/\bar{s}] = \mu x.t[\bar{y}/\bar{s}]$ by the definition of $[\bar{y}/\bar{s}]$. For $|\bar{x}| > 1$, assume $\bar{x} = (y, z)$, $\bar{t} = (r, s)$, and write $\bar{u}$ for $\bar{y}$, $\bar{v}$ for $\bar{s}$. Then, since $y, z \notin$ *free*(u),

$$\begin{aligned}
\mu\bar{x}\bar{t}[\bar{v}/\bar{u}] &= \mu(y,z)(r,s)[\bar{v}/\bar{u}]\\
&= (\mu y.r[z/\mu z s], \mu z.s[y/\mu y r])[\bar{v}/\bar{u}]\\
&= ((\mu y.r[z/\mu z s])[\bar{v}/\bar{u}], (\mu z.s[y/\mu y r])[\bar{v}/\bar{u}])\\
&= (\mu y.r[z/\mu z s][\bar{v}/\bar{u}], \mu z.s[y/\mu y r][\bar{v}/\bar{u}])\\
&= (\mu y.r[\bar{v}/\bar{u}][z/\mu z s[\bar{v}/\bar{u}]], \mu z.s[\bar{v}/\bar{u}][y/\mu y r[\bar{v}/\bar{u}]])\\
&= (\mu y.r[\bar{v}/\bar{u}][z/\mu z.s[\bar{v}/\bar{u}]], \mu z.s[\bar{v}/\bar{u}][y/\mu y.r[\bar{v}/\bar{u}]])\\
&= \mu(y,z)(r[\bar{v}/\bar{u}], s[\bar{v}/\bar{u}])\\
&= \mu\bar{x}.\bar{t}[\bar{v}/\bar{u}].
\end{aligned}$$

◀

Proof of Lemma 14. By induction on the dimension $|\bar{t}|$, we prove (10), (9) and, moreover, the substitution property for term vectors, for any $g : X \to M$, and $\bar{s}$ such that $\bar{x} \notin$ *free*$(\bar{s})$,

$$\mu\bar{x}\bar{t}[\bar{y}/\bar{s}]^M(g) = \mu\bar{x}\bar{t}^M(g[\bar{y}/\bar{s}^M(g)]).$$

We suppress parameters in the notation and consider $\bar{x} = (\bar{y}, \bar{z})$, $\bar{t} = (\bar{r}(\bar{y}, \bar{z}), \bar{s}(\bar{y}, \bar{z}))$ with $|\bar{y}|, |\bar{z}| < |\bar{x}|$. For (10), suppose tuples $\bar{a}, \bar{b} \in A$ satisfy $(\bar{r}, \bar{s})^M(\bar{a}, \bar{b}) \leq (\bar{a}, \bar{b})$. By induction,

$$\mu\bar{y}\bar{r}^M[\bar{z}/\bar{b}] \leq \bar{a} \quad \text{and} \quad \mu\bar{z}\bar{s}^M[\bar{y}/\bar{a}] \leq \bar{b}.$$

By the substitution property and monotonicity of term vectors of dimension $< |\bar{t}|$ via (11),

$$\begin{aligned}
\bar{r}[\bar{z}/\mu\bar{z}\bar{s}]^M[\bar{y}/\bar{a}] &= \bar{r}^M[\bar{y}/\bar{a}, \bar{z}/\mu\bar{z}\bar{s}^M[\bar{y}/\bar{a}])\\
&\leq \bar{r}^M[\bar{y}/\bar{a}, \bar{z}/\bar{b}]\\
&\leq \bar{a}.
\end{aligned}$$

Using (10) inductively, $\mu\bar{y}.\bar{r}[\bar{z}/\mu\bar{z}\bar{s}]^M \leq \bar{a}$. Likewise, $\mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]^M \leq \bar{b}$, and so we obtain

$$\mu(\bar{y}, \bar{z})(\bar{r}, \bar{s})^M = (\mu\bar{y}.\bar{r}[\bar{z}/\mu\bar{z}\bar{s}]^M, \mu\bar{z}.\bar{s}[\bar{y}/\mu\bar{y}\bar{r}]^M) \leq (\bar{a}, \bar{b}).$$

To show (9), we improve readability by writing substitutions in place, i.e.

$$\bar{t}[\bar{x}/\mu\bar{x}\bar{t}] = (\bar{r}(\mu\bar{y}.\bar{r}(\bar{y}, \mu\bar{z}\bar{s}), \mu\bar{z}.\bar{s}(\mu\bar{y}\bar{r}, \bar{z})), \bar{s}(\mu\bar{y}.\bar{r}(\bar{y}, \mu\bar{z}\bar{s}), \mu\bar{z}.\bar{s}(\mu\bar{y}\bar{r}, \bar{z}))).$$

As both components of $\bar{t}[\bar{x}/\mu\bar{x}\bar{t}]^M \leq \mu\bar{x}\bar{t}^M$ can be treated alike, we only show the first one,

$$\bar{r}(\mu\bar{y}.\bar{r}(\bar{y}, \mu\bar{z}\bar{s}), \mu\bar{z}.\bar{s}(\mu\bar{y}\bar{r}, \bar{z}))^M \leq \mu\bar{y}.\bar{r}(\bar{y}, \mu\bar{z}\bar{s})^M. \tag{24}$$

By induction, $\bar{r}(\bar{y},\mu\bar{z}\bar{s})[\bar{y}/\mu\bar{y}.\bar{r}(\bar{y},\mu\bar{z}\bar{s})]^M \leq \mu\bar{y}.\bar{r}(\bar{y},\mu\bar{z}\bar{s})^M$, so for (24) it is sufficient to prove

$$\bar{r}(\mu\bar{y}.\bar{r}(\bar{y},\mu\bar{z}\bar{s}),\mu\bar{z}.\bar{s}(\mu\bar{y}\bar{r},\bar{z}))^M \leq \bar{r}(\bar{y},\mu\bar{z}\bar{s})[\bar{y}/\mu\bar{y}.\bar{r}(\bar{y},\mu\bar{z}\bar{s})]^M,$$

for which, by monotonicity of $\bar{r}^M$, it's sufficient to show

$$\mu\bar{z}.\bar{s}(\mu\bar{y}\bar{r},\bar{z})^M \leq \mu\bar{z}.\bar{s}(\mu\bar{y}.\bar{r}(\bar{y},\mu\bar{z}\bar{s}),z)^M. \tag{25}$$

Let $\bar{a}=\mu\bar{y}.\bar{r}(\bar{y},\mu\bar{z}\bar{s})^M$ and $\bar{b}$ the least $\bar{b}'$ with $\bar{s}^M(\bar{a},\bar{b}')\leq\bar{b}'$, i.e. $\bar{b}=\mu\bar{z}\bar{s}^M[\bar{y}/\bar{a}]$. Then

$$\begin{array}{rcll} \bar{r}^M(\bar{a},\bar{b}) & = & \bar{r}[z/\mu\bar{z}\bar{s}]^M(\bar{a}) & \text{(by (11), inductively)} \\ & = & \bar{r}(\bar{y},\mu\bar{z}\bar{s})^M[y/\bar{a}] & \\ & \leq & \bar{a} & \text{(by (9), (11), inductively),} \end{array}$$

hence $\mu\bar{y}\bar{r}^M[z/b]\leq\bar{a}$ by an application of (10). By monotonicity, it follows that

$$\begin{array}{rcll} \bar{s}(\mu\bar{y}.\bar{r}(\bar{y},\bar{z}),\bar{z})^M[\bar{z}/\bar{b}] & \leq & \bar{s}^M(\bar{a},\bar{b}) & \\ & \leq & \bar{b} & \text{(by the choice of } \bar{b}). \end{array}$$

An application of (10) to this gives $\mu\bar{z}.\bar{s}(\mu\bar{y}.\bar{r}(\bar{y},\bar{z}),\bar{z})^M\leq\bar{b}$, which is (25).

To show (11) for $|\bar{t}|>1$, by induction it follows from Lemma 13 that

$$\begin{array}{rcl} \mu\bar{x}\bar{t}[\bar{y}/\bar{s}]^M(g) & = & (\mu\bar{x}.\bar{t}[\bar{y}/\bar{s}])^M(g) \\ & = & \text{the least } \bar{a} \text{ such that } \bar{t}[\bar{y}/\bar{s}]^M(g[\bar{x}/\bar{a}])\leq\bar{a} \\ & = & \text{the least } \bar{a} \text{ such that } \bar{t}^M(g[\bar{x}/\bar{a}][\bar{y}/\bar{s}^M(g[\bar{x}/\bar{a}])])\leq\bar{a} \\ & = & \text{the least } \bar{a} \text{ such that } \bar{t}^M(g[\bar{y}/\bar{s}^M(g)][\bar{x}/\bar{a}])\leq\bar{a} \\ & = & \mu\bar{x}\bar{t}^M(g[\bar{y}/\bar{s}^M(g)]). \end{array}$$

The case $|\bar{t}|=1$ is an instance of the substitution property for M. ◀

Omitted bits in the proof of Theorem 25. The properties of 0^N, 1^N, x^N, $(s+t)^N$ and $(s\cdot t)^N$ are immediate from the definition. For example,

$$(s+t)^N(g)=((s+t)'^M_{i,j}(\hat{g}))=(s'^M_{i,j}(\hat{g})+^M t'^M_{i,j}(\hat{g}))=s'^N(g)+t'^N(g).$$

The μ-rule holds in N, since, by corollary 15, the vector version of the μ-rule holds in M. Namely, suppose $s^N\leq t^N$, and $g:X\to N$. Then $\mu xt^N(g)=\mu x't'^M(\hat{g})$ is the least $\bar{a}$ such that $t'^M(\hat{g}[x'/\bar{a}])\leq\bar{a}$ by lemma 14. From $s^N\leq t^N$ we get $s'^M(\hat{g}[x'/\bar{a}])\leq t'^M(\hat{g}[x'/\bar{a}])\leq\bar{a}$, which implies $\mu xs^N(g)\leq\mu xt^N(g)$. ◀