Introduction to Kleene Algebra	**Lecture 15**
CS786 Spring 2002	**Tuesday, March 26, 2002**

Parikh's Theorem in Commutative KA

Parikh's theorem [6] says that every context-free language is "letter-equivalent" to a regular set; formally, the commutative image of any context-free language is also the commutative image of some regular set, where the *commutative image* of a set A of strings over the finite alphabet $\{a_1, \ldots, a_k\}$ is the set of k-tuples

$$\{(\#a_1(x), \ldots, \#a_k(x)) \in \mathbb{N}^k \mid x \in A\} \quad \subseteq \quad \mathbb{N}^k,$$

where $\#a_i(x)$ is the number of occurrences of a_i in x. The k-tuples in $\mathbb{N}^k$ are often called *Parikh vectors*. For example, the context-free language $\{a^n b^n \mid n \geq 0\}$ is letter-equivalent to the regular set $(ab)^*$; these two sets have a common commutative image $\{(n, n) \mid n \geq 0\}$.

The usual combinatorial proofs of Parikh's theorem (see for example [3]) involve an induction on parse trees of context-free grammars. In this lecture and the next we prove the following general theorem of commutative Kleene algebra, of which Parikh's theorem is a special case:

Theorem 0.1 *Every system of inequalities*

$$f_i(x_1, \ldots, x_n) \quad \leq \quad x_i, \quad 1 \leq i \leq n, \tag{1}$$

where the f_i are polynomials in $K[x_1, \ldots, x_n]$ over a commutative Kleene algebra K, has a unique least solution in K^n; moreover, the components of the solution are given by polynomials in the coefficients of the f_i.

We might take the statement of Theorem 0.1 as a definition of *algebraic closure* in Kleene algebra, in which case the theorem says that any commutative Kleene algebra is algebraically closed.

Pilling [7] proves Theorem 0.1 in the special case of the commutative Kleene algebra $\mathsf{Reg}(\mathbb{N}^k)$, the algebra of regular sets of Parikh vectors, and argues that this is the essential content of Parikh's theorem. Indeed, context-free grammars are just systems of set inequalities, and the context-free languages they generate are the minimal solutions. For example, the context-free grammar $S \to aSb \mid \varepsilon$ is essentially the system consisting of the single inequality $axb + 1 \leq x$ whose least solution in $2^{\{a,b\}^*}$ is the context-free language $\{a^n b^n \mid n \geq 0\}$. Under the assumption of commutativity, the inequality can be rewritten

$abx + 1 \leq x$, whose least solution is the regular set $(ab)^*$. For a somewhat more difficult example, the context-free grammar $S \to \mathtt{[}S\mathtt{]} \mid SS \mid \varepsilon$ generating the set of balanced strings of parentheses is essentially the system consisting of the single inequality $\mathtt{[}x\mathtt{]} + xx + 1 \leq x$. Under the assumption of commutativity, the inequality can be rewritten $\mathtt{[]}x + x^2 + 1 \leq x$, whose least solution is the regular set $(\mathtt{[]})^*$.

Using Theorem **??** of Lecture **??**, one can generalize Pilling's proof to any star-continuous Kleene algebra. However, the proof makes essential use of various infinitary properties such as the continuity of regular operators and the fact that a^* is the supremum of the a^n, $n \geq 0$.

Kuich [5] also gives a generalization of Parikh's theorem that holds for any commutative idempotent ω-continuous semiring. Kuich's result implies Pilling's, since $\mathsf{Reg}(\mathbb{N}^k)$ is embedded in the commutative idempotent ω-continuous semiring $2^{\mathbb{N}^k}$. Conversely, since every commutative idempotent ω-continuous semiring is a commutative Kleene algebra under the usual definition of the * operator

$$a^* \quad = \quad \sum_{n \geq 0} a^n,$$

Pilling's result, suitably generalized to star-continuous Kleene algebras, would imply Kuich's. But again, these proofs depend on the strong infinitary properties of star-continuous algebras.

Our result is a generalization of these results in that it holds in all commutative Kleene algebras. The main difference here is that Kleene algebra as defined in [2] has a *finitary* algebraic axiomatization consisting of finitely many equations and equational implications. Thus one might say that we are replacing the *analytic* arguments of Pilling and Kuich with *algebraic* arguments. The fact that we cannot argue combinatorially in the model $\mathsf{Reg}(\mathbb{N}^k)$ or use the infinitary properties of star-continuous algebras makes the proof more difficult, but also makes the result considerably stronger.

The situation is analogous to the fundamental theorem of algebra, which states that the complex numbers $\mathbb{C}$ are algebraically closed. The most common proof of this theorem, originally due to Gauss, depends on the analytic structure of $\mathbb{C}$ and uses second-order arguments (see e.g. [9]). However, one can give a first-order, purely algebraic proof of the more general result that if R is any real closed field (such as $\mathbb{R}$ or $\mathbb{A}$, the real algebraic numbers), then $R[i]$ is algebraically closed (see e.g. [8]). Like the fundamental theorem of algebra, our result also deals with solutions of polynomial systems, and our proof replaces arguments referring to the analytic or second-order structure of $\mathsf{Reg}(\mathbb{N}^k)$, embodied in the star-continuity axiom, with first-order equational arguments referring only to the finitary algebraic structure of commutative Kleene algebras.

Our development involves the definition of differential operators $\frac{\partial}{\partial x}$ on commutative Kleene algebras of polynomials and a version of Taylor's theorem:

$$f(x + d) \quad = \quad f(x) + f'(x + d) \cdot d.$$

Differential operators allow us to define the *Jacobian matrix* of a system of inequalities, which we use to give a closed form solution.

The results of this lecture and the next are from [1].

Commutative Kleene Algebra

Recall that a Kleene algebra is *commutative* if it satisfies the property $xy = yx$ for all x, y. Commutativity assumptions also arise in practice [4].

We have observed previously that the following is a theorem of commutative Kleene algebra that does not hold in Kleene algebra in general:

$$(p+q)^* \quad = \quad p^*q^*. \tag{2}$$

We have also observed that one can prove a normal form theorem that says that every expression is equivalent to a sum $y_1 + \cdots + y_n$, where each y_i is a product of atomic symbols and expressions of the form $(a_1 \cdots a_k)^*$, where the a_i are atomic symbols. For example,

$$(((ab)^*c)^* + d)^* \quad = \quad d^* + (ab)^*c^*cd^*.$$

This normal form was observed by Pilling [7] in the context of $\mathsf{Reg}(\mathbb{N}^k)$, but using (2) it is easily shown to hold in all commutative Kleene algebras.

Polynomials

If K is a commutative Kleene algebra, we denote by $K[\mathbf{x}]$ the commutative Kleene algebra of polynomials in indeterminates $\mathbf{x}$ over K. These are very much like polynomials over a ring or field. We can think of a polynomial as a regular expression over K and $\mathbf{x}$ reduced modulo the axioms of commutative Kleene algebra and the diagram of K (the set of ground identities that hold in K). Typical examples of polynomials are

$$\begin{gathered}(ax+by)^* \\ 1 + (ax^*b^*)^* + bx + cy \\ a + xy(bxy)^*,\end{gathered}$$

where x, y are indeterminates and $a, b, c \in K$.

Formally, $K[\mathbf{x}]$ is defined to be the direct sum (coproduct) of K with the free commutative Kleene algebra on generators $\mathbf{x}$ in the category of commutative Kleene algebras. The most significant property of polynomials is that any pair of maps h, h', where $h : K \to L$ is a Kleene algebra homomorphism and $h' : \mathbf{x} \to L$ is a set function, extend simultaneously and uniquely to a Kleene algebra homomorphism $\widehat{h} : K[\mathbf{x}] \to L$. When h is the identity on K, the map $\widehat{h}$ is just polynomial evaluation; intuitively, applying $\widehat{h}$ can be regarded as substituting the values $h'(x)$ for the indeterminates $x \in \mathbf{x}$ and then evaluating the resulting expression.

If $\mathbf{x} = x_1, \ldots, x_n$ and $\mathbf{a} = a_1, \ldots, a_n$, we write $f(\mathbf{a})$ or $f(\mathbf{x}) \mid_{\mathbf{x}=\mathbf{a}}$ for the value of f evaluated at $x_i \mapsto a_i$, $1 \leq i \leq n$.

Differential Operators

A map $D : K \to K$ on a commutative Kleene algebra K is called a *differential operator* if for all $x, y \in K$,

$$\begin{aligned} D(x+y) &= Dx + Dy \\ D(xy) &= xDy + yDx \\ D(x^*) &= x^*Dx \\ D0 &= D1 = 0. \end{aligned} \tag{3}$$

For example, in $\mathsf{Reg}(\mathbb{N}^k)$, for every $1 \leq i \leq k$, the map

$$\begin{aligned} A \ \mapsto \ & \{(a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_k) \mid \\ & (a_1, \ldots, a_k) \in A, \ a_i > 0\} \end{aligned}$$

is a differential operator.

Theorem 0.2 *Any differential operator $D : K \to K$ and set function $D : \mathbf{x} \to K$ have a unique joint extension to a differential operator $D : K[\mathbf{x}] \to K[\mathbf{x}]$.*

Proof. The given maps D can be extended by induction to $D : K[\mathbf{x}] \to K[\mathbf{x}]$ using (3); but we must take care that the extended D is well-defined on equivalence classes modulo the axioms of commutative Kleene algebra and the diagram of K. That the extended D respects the diagram of K follows from the fact that the given $D : K \to K$ is a differential operator on K. To prove that D respects the commutative Kleene algebra axioms requires a case for each axiom. We argue the cases $a^* = 1 + aa^*$ and $ab \leq b \to a^*b \leq b$ explicitly.

For the case $a^* = 1 + aa^*$,

$$\begin{aligned} D(1 + aa^*) &= D1 + D(aa^*) \\ &= 0 + aD(a^*) + a^*Da \\ &= aa^*Da + a^*Da \\ &= a^*Da \\ &= D(a^*). \end{aligned}$$

For the case $ab \leq b \to a^*b \leq b$, suppose $ab \leq b$. By the induction hypothesis, $D(ab) \leq Db$, and we wish to show that $D(a^*b) \leq Db$. From $D(ab) \leq Db$ we have that $aDb + bDa \leq Db$, thus by Kleene algebra we have that $a^*Db \leq Db$ and $a^*bDa \leq Db$. Therefore

$$\begin{aligned} D(a^*b) &= a^*Db + bD(a^*) \\ &= a^*Db + ba^*Da \\ &\leq Db. \end{aligned}$$

□

In particular, for $x \in \mathbf{x}$, we define a certain differential operator $\frac{\partial}{\partial x} : K[\mathbf{x}] \to K[\mathbf{x}]$ as follows. The value of $\frac{\partial}{\partial x}$ applied to $f \in K[\mathbf{x}]$ is denoted $\frac{\partial f}{\partial x}$ or $\frac{\partial f}{\partial x}(\mathbf{x})$. We define $\frac{\partial}{\partial x}$ to be the unique differential operator such that $\frac{\partial x}{\partial x} = 1$, $\frac{\partial y}{\partial x} = 0$ for $y \in \mathbf{x} - \{x\}$, and $\frac{\partial a}{\partial x} = 0$ for $a \in K$.

For univariate polynomials $f, e \in K[x]$, we sometimes write f' for $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial x}(e)$ or $f'(e)$ for the result of evaluating the polynomial $\frac{\partial f}{\partial x}$ at $x \mapsto e$.

For example, if $f(x) = axb + x^2 + 1$, then $f'(x) = ab + x$.

Note that $\frac{\partial}{\partial x}(f(e))$ and $\frac{\partial f}{\partial x}(e)$ are different in general. The former refers to the result of evaluating $f(x)$ at $x \mapsto e$ first, then applying the differential operator $\frac{\partial}{\partial x}$ to $f(e)$; whereas the latter refers to the result of applying the differential operator $\frac{\partial}{\partial x}$ to $f(x)$ first, then evaluating the resulting polynomial $\frac{\partial f}{\partial x}(x)$ at $x \mapsto e$. These two expressions are related by the *chain rule*:

Theorem 0.3 (chain rule) *For $f, e \in K[x]$,*

$$\frac{\partial}{\partial x}(f(e)) = \frac{\partial f}{\partial x}(e) \cdot \frac{\partial e}{\partial x},$$

or in more conventional notation,

$$f(e(x))' = f'(e(x)) \cdot e'(x).$$

Proof. This is a straightforward induction on the structure of f. We argue the cases

$f = gh$ and $f = g^*$ explicitly.

$$\begin{aligned}
(g(e)h(e))' &= g(e)h(e)' + h(e)g(e)' \\
&= g(e)h'(e)e' + h(e)g'(e)e' \\
&= (g(e)h'(e) + h(e)g'(e))e' \\
&= (gh' + hg')(e)e' \\
&= (gh)'(e)e'.
\end{aligned}$$

$$\begin{aligned}
(g(e)^*)' &= g(e)^* g(e)' \\
&= g(e)^* g'(e)e' \\
&= (g^* g')(e)e' \\
&= (g^*)'(e)e'.
\end{aligned}$$

□

For example, if $f(x) = ax + x^* + 1$, then

$$\begin{aligned}
f'(x) &= a + x^* \\
f(f(x)) &= a(ax + x^* + 1) + (ax + x^* + 1)^* + 1 \\
&= a^2x + ax^* + (ax)^*x^* + 1 \\
f(f(x))' &= (a^2x + ax^* + (ax)^*x^* + 1)' \\
&= a^2 + (a+1)x^*(ax)^* \\
f'(f(x))f'(x) &= (a + (ax + x^* + 1)^*)(a + x^*) \\
&= a^2 + (a+1)x^*(ax)^*.
\end{aligned}$$

We also have the following version of Taylor's theorem in commutative Kleene algebra.

Theorem 0.4 (Taylor's theorem) *For $f, d \in K[x]$,*

$$f(x+d) = f(x) + f'(x+d) \cdot d.$$

In particular, evaluating at $x \mapsto 0$,

$$f(d) = f(0) + f'(d) \cdot d.$$

Proof. This is again a straightforward induction on the structure of f. As before, we argue the cases $f = gh$ and $f = g^*$ explicitly. For the case $f = gh$,

$$\begin{aligned}
& gh(x+d) \\
&= g(x+d)h(x+d) \\
&= g(x+d)h(x) + g(x+d)h'(x+d)d \\
&= g(x)h(x) + g'(x+d)h(x)d \\
&\quad + g(x+d)h'(x+d)d,
\end{aligned}$$

and by symmetry,

$$\begin{aligned}
& gh(x+d) \\
& \quad = \quad g(x)h(x) + g'(x+d)h(x+d)d \\
& \qquad\qquad + g(x)h'(x+d)d,
\end{aligned}$$

therefore by monotonicity,

$$\begin{aligned}
& gh(x+d) \\
& \quad = \quad g(x)h(x) + g'(x+d)h(x)d + g(x+d)h'(x+d)d \\
& \qquad\quad + g'(x+d)h(x+d)d + g(x)h'(x+d)d \\
& \quad = \quad g(x)h(x) + g(x+d)h'(x+d)d \\
& \qquad\quad + g'(x+d)h(x+d)d \\
& \quad = \quad gh(x) + (gh)'(x+d)d.
\end{aligned}$$

For the case $f = g^*$,

$$\begin{aligned}
& g(x+d)^* \\
& \quad = \quad (g(x) + g'(x+d)d)^* \\
& \quad = \quad g(x)^*(g'(x+d)d)^* \qquad \text{by (2)} \\
& \quad = \quad g(x)^* + g(x)^* g'(x+d)d(g'(x+d)d)^* \\
& \quad = \quad g(x)^* + g'(x+d)dg(x)^*(g'(x+d)d)^* \\
& \quad = \quad g(x)^* + g'(x+d)dg(x+d)^* \\
& \quad = \quad g(x)^* + g(x+d)^* g'(x+d)d \\
& \quad = \quad g(x)^* + (g^*)'(x+d)d.
\end{aligned}$$

□

Continuing with the example $f(x) = ax + x^* + 1$ above,

$$\begin{aligned}
f(x+d) &= a(x+d) + (x+d)^* + 1 \\
&= ax + ad + x^*d^* + 1 \\
f'(x+d) &= a + (x+d)^* \\
&= a + x^*d^* \\
f(x) + f'(x+d)d &= ax + x^* + 1 + ad + x^*d^*d \\
&= ax + ad + x^*d^* + 1.
\end{aligned}$$

We often wish to differentiate simultaneously with respect to a sequence of indeterminates $\mathbf{y} = y_1, \ldots, y_k$. We define an operator $\frac{\partial}{\partial \mathbf{y}}$ that when applied to an element $f \in K[\mathbf{x}]$ produces a row vector of length k whose i^{th} component is $\frac{\partial f}{\partial y_i}$. More generally, $\frac{\partial}{\partial \mathbf{y}}$ applied to a column vector consisting of m elements $f_1, \ldots, f_m \in K[\mathbf{x}]$ produces an $m \times k$ matrix whose i, j^{th} element is $\frac{\partial f_i}{\partial y_j}$.

By iterating Theorem 0.4, one can show that for $f \in K[\mathbf{x}]$ and $\mathbf{e} = e_1, \ldots, e_n$,

$$\begin{aligned} f(\mathbf{e}) &= f(0, \ldots, 0) + \frac{\partial f}{\partial x_1}(\mathbf{e})e_1 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{e})e_n \\ &= f(\mathbf{0}) + \frac{\partial f}{\partial \mathbf{x}}(\mathbf{e}) \cdot \mathbf{e}, \end{aligned}$$

where $\cdot$ denotes dot product of vectors. The same holds for a column vector $\mathbf{f} = f_1, \ldots, f_m$ of elements of $K[\mathbf{x}]$; here $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x})$ is an $m \times n$ matrix whose i, j^{th} element is $\frac{\partial f_i}{\partial x_j}$. This matrix is called the *Jacobian* of $\mathbf{f}$. We thus have

Theorem 0.5

$$\mathbf{f}(\mathbf{e}) = \mathbf{f}(\mathbf{0}) + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{e}) \cdot \mathbf{e},$$

where $\cdot$ denotes matrix-vector multiplication.

We also have the following vector-vector and matrix-vector versions of the chain rule, Theorem 0.3:

Theorem 0.6

$$\begin{aligned} \frac{\partial}{\partial z}(f(\mathbf{e})) &= \frac{\partial f}{\partial x_1}(\mathbf{e})\frac{\partial e_1}{\partial z} + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{e})\frac{\partial e_n}{\partial z} \\ &= \frac{\partial f}{\partial \mathbf{x}}(\mathbf{e}) \cdot \frac{\partial \mathbf{e}}{\partial z} \\ \frac{\partial}{\partial z}(\mathbf{f}(\mathbf{e})) &= \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{e}) \cdot \frac{\partial \mathbf{e}}{\partial z}. \end{aligned}$$

The proof of these theorems are straightforward generalizations of their one-dimensional versions.

In the next lecture we will apply these general results to obtain an algebraic version of Parikh's theorem.

References

[1] Mark Hopkins and Dexter Kozen. Parikh's theorem in commutative Kleene algebra. In *Proc. Conf. Logic in Computer Science (LICS'99)*, pages 394–401. IEEE, July 1999.

[2] Dexter Kozen. A completeness theorem for Kleene algebras and the algebra of regular events. *Infor. and Comput.*, 110(2):366–390, May 1994.

[3] Dexter Kozen. *Automata and Computability.* Springer-Verlag, New York, 1997.

[4] Dexter Kozen. Kleene algebra with tests. *Transactions on Programming Languages and Systems*, 19(3):427–443, May 1997.

[5] Werner Kuich. The Kleene and Parikh theorem in complete semirings. In T. Ottmann, editor, *Proc. 14th Colloq. Automata, Languages, and Programming*, volume 267 of *Lecture Notes in Computer Science*, pages 212–225, New York, 1987. EATCS, Springer-Verlag.

[6] R. J. Parikh. On context-free languages. *J. Assoc. Comput. Mach.*, 13(4):570–581, 1966.

[7] D. L. Pilling. Commutative regular equations and Parikh's theorem. *J. London Math. Soc.*, 6(4):663–666, June 1973.

[8] Ian Stewart. *Galois Theory.* Chapman and Hall, London, 1973.

[9] B. L. van der Waerden. *Algebra*, volume 1. Frederick Ungar, 1970.