A new proof of Parikh!

1. Algebraic grammars

Let M be a monoid and let Ξ be a finite alphabet (disjoint from M). We define a monoid $M[\Xi]$ as follows. The elements of $M[\Xi]$ are words

(1.1) $$m_0 \xi_1 m_1 \dots m_{k-1} \xi_k m_k$$

with $k \geq 0$, $m_0, \dots, m_k \in M$ and $\xi_1, \dots, \xi_k \in \Xi$. The elements $m_0, \dots, m_k$ are called the <u>coefficients</u> while $\xi_1, \dots, \xi_k$ are called the <u>variables</u>. The product of (1.1) with a word ~~[illegible]~~

$$n_0 \xi_1 n_1 \dots n_{\ell-1} \xi_\ell n_\ell$$

is the ~~[illegible]~~ word

$$m_0 \xi_1 m_1 \dots m_{k-1} \xi_k (m_k n_0) \xi_1 n_1 \dots n_{\ell-1} \xi_\ell n_\ell$$

We shall say that the monoid $M[\Xi]$ is obtained from M by adjoining the set Ξ as a set of free variables.

If $M = \Sigma^*$ then $M[\Xi] = (\Sigma \cup \Xi)^*$. ~~assuming that Σ and Ξ are disjoint.~~

2

An *algebraic* (or *context-free*) *grammar* G in a monoid M consists of the following data:

(1.2) A finite (non-empty) alphabet Ξ disjoint from M. The elements of Ξ are called *variables*.

(1.3) A distinguished element $\xi_0 \in \Xi$ called the *start variable*.

(1.4) A finite subset R of $\Xi \times M[\Xi]$ called the set of *rules*.

For each rule $r \in R$, we denote by $\bar{r} \in \Xi$ its first coordinate and by $\underline{r} \in M[\Xi]$ its second coordinate. Instead of writing $(\bar{r}, \underline{r}) \in R$ we shall frequently write $\bar{r} \longrightarrow \underline{r}$ or $r : \bar{r} \to \underline{r}$

~~and $\underline{r} \notin \Xi$. Thus~~

$$\cancel{\underline{r} \in (\Sigma \cup \Xi)^+ - \underset{\sim}{\Xi}}$$

~~In other words, a positive grammar has no rules of the form~~

$$\cancel{\xi \to 1, \quad \xi \to \xi'.}$$

For each rule $r \in R$ we define a partial function $\cdot r : M[\Xi] \to M[\Xi]$ as follows:

$$w r = w'$$

whenever

$$w = u \bar{r} v \;, \quad u \in M[\Xi] \;, \; v \in M$$
$$w' = u \underline{r} v$$

In all other cases $w r = \emptyset$. We call the readers attention to the condition $v \in M$. It signifies that $\bar{r}$ is rightmost appearan of a variable in w. It is only then that $\bar{r}$ is replaced by $\underline{r}$ (i.e. is rewritten using the rule r).

Given

$$d = r_1 \cdots r_n \in R^*$$

we define

$$wd = (\cdots\{(w r_1) r_2\} \cdots r_n)$$

and $w1 = w$ (if $n = 0$). Thus each $d \in R$ defines a partial function $M[\Xi] \to M[\Xi$

and

$$w(d_1 d_2) = (w d_1) d_2$$

If $wd = w'$ we say that d is a <u>derivation</u> of w into w' and also use the notation $d : w \to w'$ or $w \xrightarrow{d} w'$.

~~So far the start variable~~ ξ_0 ~~plays no role. It does~~ play a role in the following definition.

The <u>language</u> L ~~of G~~ <u>defined by the grammar</u> G is the subset of Σ^*

(1.5) $L = \{ s \mid s \in \Sigma^*, \ d : \xi_0 \to s$ for some $d \in R^*$

~~Since $\xi_0 \notin M$~~ we may replace $d \in R^*$ by

So far the start variable played no role. It does play a role in the following definition. A derivation

$$d : \xi_0 \longrightarrow m \quad , \quad m \in M$$

is called a successful derivation. The set L of all elements $m \in M$ for which such a successful derivation exists is called the language determined by the grammar G ~~and is denoted by GL~~. Thus

$$L = \{ m \mid m \in M,\ d : \xi_0 \to m \text{ for some } d \in R^* \}$$

Since $\xi_0 \notin M$, we may replace $d \in R^*$ by $d \in R^+$.

The above definition does not take into account the number of successful derivations $d : \xi_0 \to m$ that may exist for a particular m. Thus the definition of ~~G~~L as given above ignores multip cies and defines GL as an ordinary or as $\mathcal{B}$-subset of M.

If we wish to pay attention to the phenomenon of multiplicity, then we define mL, to be the number of successful derivations $d: \xi_0 \to m$. Since this number could be infinite we have

$$mL \in \mathcal{N} = N \cup (\infty)$$

where N is the set of all integers $n \geq 0$. Thus L becomes a function

$$(L : M \longrightarrow \mathcal{N}$$

or in other words an $\mathcal{N}$-subset of M.

An $\mathcal{N}$-subset or a $\mathcal{B}$-subset of M is called <u>algebraic</u> if it is defined by some algebraic grammar in M. We denote the class of all algebraic $\mathcal{N}$-subsets of M by

$$M\,Alg_{UV}$$

while the class of all algebraic B-subsets is denoted by

$$M\,Alg_{SN}.$$

Proposition 1.1. Every derivation

$$d : w_2 w_1 \to m \quad , \quad m \in M$$

admits a unique factorization

$$d = d_1 d_2 , \quad \cancel{d_1 : w_1 \to m_1 , \quad d_2 : w_2 \to m_2}$$

such that

(1.5) $\quad d_1 : w_1 \to m_1 , \quad d_2 : w_2 \to m_2 , \quad m_2 m_1 = m$

Conversely given (1.5) we have

$d_1 d_2 : w_2 w_1 \to m_2 m_1$.

Proof. We first show the uniqueness of the factorization $d = d_1 d_2$. Indeed assume that $d = e_1 e_2$ is another such factorization. Assume $|d_1| < |e_1|$. Then $e_1 = d_1 c$. Since $w_1 d_1 = m \in M$ we have $w_1 e_1 = mc$. This is impossible since $mc = \emptyset$.

The proof of the existence of the factorization $d = d_1 d_2$ proceeded by induction on the length $|d|$ of the derivation d. If $|d| = 0$, then $w_1 w_2 = m$ and $d_1 = 1 = d_2$. Now assume that $|d| = 1 + n$ and that the assertion holds for derivations of length n.

We also observe that if $w_1 \in M$ then $(w_2 w_1)d = (w_2 d)w_1$ and thus $d_1 = d$ and $d_2 = 1$ give the required factorization.

Let $d = rd'$. Then $w_2 w_1 = u\bar{r}p$ with $p \in M$. Since we assume that $w_1 \notin M$ it follows that $w_1 = v\bar{r}p$, $u = w_2 v$. Then $w_2 w_1 r = w_2 v \underline{r} p$. Since $d' : w_2(v\underline{r}p) \to m$ we have a factorization $d' = d_1' d_2$ such that $(v\underline{r}p)d_1 = m_1$, $w_2 d_2 = m_2$ and $m = m_2 m_1$. Then $d = (r d_1')d_2$ is the required factorization of d ∎

Exercise 1.1. Show that for each algebraic K-subset A of M (with $K = \mathbb{B}$ or $K = \mathbb{N}$) there exists a finitely generated submonoid M' of M such that A is an algebraic K-subset of M'.

2. Examples

Example 2.1. Let A be a finite subset of monoid M. Consider the grammar with a single variable ξ (which is then necessarily the start variable) and with rules

$$\xi \longrightarrow a$$

one for each $a \in A$. Clearly A is the language ~~con~~ defined by this grammar. Thus A is an algebraic subset of M.

This example also shows that if we allow the set R of rules to be infinite then any subset of M could be shown to be algebraic (see however Theorem IV, Example 2.2. Let $M = \{\sigma, \tau\}^*$. Consider the grammar

$$\xi \to \sigma \xi \tau \quad , \quad \xi \to 1$$

Since there is only one variable ξ it is automatically the start variable. Let r_1 and r_2 designate the two displayed rules. The only derivations $d : \xi \to w$ possible are

$$r_1^n : \xi \to \sigma^n \xi \tau^n \quad , \quad n \geq 0$$

$$r_2 r_1^n : \xi \to \sigma^n \tau^n \quad , \quad n \geq 0$$

The derivations $r_2 r_1^n$ are the only successful ones and consequently the grammar defines the algebraic set

$$A = \{\sigma^n \tau^n \mid n \geq 0\}$$

If in this example we replace M by the commutative monoid freely generated by σ and τ, we obtain the subset $A = (\sigma\tau)^*$.

Example 2.3. Let $M = \{\sigma, \tau\}^*$, Consider the grammar with variables ξ (start variable) and η.

$$r_1 : \xi \to \sigma \xi \tau, \quad r_2 : \xi \to \tau \eta \sigma, \quad r_3 \; \xi \to 1$$
$$r_4 \; \eta \to \tau \eta \sigma, \quad r_5 : \eta \to 1$$

Show that the grammar computes the subset

$$A = \{ \sigma^p \tau^q \sigma^q \tau^p \mid p \geq 0, q \geq 0 \}$$

Show that the only successful derivations are

$$r_1^p r_2 r_4^{q-1} r_5 : \xi \to \sigma^p \tau^q \sigma^q \tau^p \quad \text{if } q > 0$$
$$r_1^p r_3 \quad : \xi \to \sigma^p \tau^p \quad \text{if } q = 0$$

Example 2.4. Given any alphabet Σ consider grammar with rules

$$r_\sigma : \xi \to \sigma \xi \sigma, \quad r'_\sigma : \xi \to \sigma, \quad r''_\sigma : \xi \to 1$$

Show that this grammar defines the set of all <u>palindromes</u>

$$A = \{ s \mid s \in \Sigma^*, s = s^g \}$$

where s^g is the reversal of s.

Example 2.5. Let $M = \sigma^*$. Consider the grammar

$$\xi_0 \to \sigma, \quad \xi \to \xi', \quad \xi' \to \sigma$$

The grammar defines the set consisting of σ alone. However note that we have two distinct derivations $d : \xi_0 \to \sigma$, one of length 1 and another one of length 2. Thus the $\mathbb{N}$-subset A of σ^* defined by this grammar is 2σ.

3. Weights.

Let K be a commutative semiring, and let M be a monoid. An (M, K)-grammar G is an algebraic grammar (Ξ, ξ_0, R) in M together with a weight function

$$\mu : R \to K$$

Thus each rule r has a weight $r\mu \in K$. This function μ is extended to R^* by setting

$$d\mu = (r_1\mu) \cdots (r_n\mu)$$

for $d = r_1 \cdots r_n$ (Note that $d\mu = 1$ if $n = 0$).

Using these weight we now modify the definition of the language L defined by the grammar. L is now to become a K-subset of M i.e. a function

$$L : M \to K$$

defined by

$$(3.1) \qquad mL = \sum_{\xi_0 d = m} d\mu$$

There is something drastically wrong with this definition. The summation in (3.1) extends over all derivations

$$d : \xi_0 \longrightarrow m \tag{3.2}$$

and there may be infinitely many of them. Thus the sum (3.1) may be undefined if K is just a semiring.

One way to remedy this difficulty is to use a complete semiring K rather than just a semiring (see A, In such a semiring arbitrary summations can be carried out and thus (3.1) is well defined. The semirings $\mathcal{B}$ and $\mathcal{N}$ are examples of complete semirings. Actually the assumption that K is complete is too strong. In I\III.3 we shall introduce the notion of a <u>continuous semiring</u>. In such a semiring countable sums can be performed and this suffices to give meaning to (3.1). The class of algebraic K-subsets of M thus defined will be denoted by

(3.3) $$M\mathcal{A}lg_K$$

There is a second important case in which the summation (3.1) is legitimate. If we replace M by a semigroup S which is <u>loose</u> (see next section for a definition) and if we assume that the grammar is <u>positive</u> (see next section for a definition), then it can be shown that for each $s \in S$, the derivations $d : \xi_0 \to s$ form a finite set. Thus (3.1) is again well defined (for any commutative semiring K. The class of

positive
algebraic K-subsets of S thus defined will be denoted by

(3.4) $\qquad S\,\mathcal{A}lg_K$

There is no conflict between the notation (3.3) and (3.4) since in (3.3) M is a monoid, while in (3.4) S is a loose semigroup and such a semigroup cannot have a unit element.

There is a slight conflict of notation involving $M\mathcal{A}lg_{\mathcal{B}}$ and $M\mathcal{A}lg_{\mathcal{N}}$ since these we introduced in section 1 without weights (or, what is equivalent, with all ~~weights~~ the rules given weight 1). In the case of $M\mathcal{A}lg_{\mathcal{B}}$ the conflict is resolved trivially. Since $\mathcal{B} = \{0,1\}$ all weights are 0 or 1. Rules of weight 0 may be discarded so that only rules of weight 1 remain. In the case of $M\mathcal{A}lg_N$ the conflict is resolved by

Proposition 3.1. For any $(M, \mathcal{N})$-gramm. G there exists an $(M, \mathcal{N})$ grammar G' will all rules having weight 1 and such that G and G' define the same $\mathcal{N}$-subset of M.

Proof. Let $G = (\Xi, \xi_0, R)$ and let $\mu : R \to \mathcal{N}$ be the weight function. The grammar is modified in the following manner. For each rule $r : \xi \to w$ in R the following four cases are considered

(i) $r\mu = 0$. The rule r is discarded

(ii) $r\mu = 1$. The rule r is left unchanged

(iii) $r\mu = k$ with $1 < k < \infty$. New variables $\eta_1, \dots, \eta_k$ are introduced and the rule r is replaced by the $2k$ rules

$$\xi \to \eta_i \, , \quad \eta_i \to w \, , \qquad i = 1, \dots, k$$

(iv) $r\mu = \infty$. A new variable η is introduced and the rule r is replaced by the three rules

$$\xi \to \eta \,,\ \eta \to \eta \,,\ \eta \to w$$

All the new rules introduced in (iii) and (iv) receive the weight 1. The verification that the new grammar defines the same $\mathcal{N}$-subset of M as G is immediate ∎

The following notational device is useful in handling grammars with weights. If $d : \xi \to w$ is a derivation with weight $d\mu = k$, then we shall write $d : \xi \to kw$. In particular $r : \xi \to kw$ signifies that the rule r has weight k.

4. Loose semigroups and positive grammars

A semigroup S is said to be <u>loose</u> if each $s \in S$ admits only a finite number of factorizations

(4.1) $$s = s_1 \dots s_n, \quad s_1, \dots, s_n \in S$$

It follows that a loose semigroup S cannot contain an idempotent. Thus S cannot have a unit element or a zero element and cannot contain a non-empty finite subsemigroup.

If we denote by $|s|$ the largest integer n for which a factorization (4.1) exists, we obtain a function $S \to N$ with the following properties

(4.2) $|s| > 0$ for all $s \in S$

(4.3) $|s| = 1$ iff s is indecomposable, i.e. iff $s \in S - S^2$

(4.4) $|st| \geqslant |s| + |t|$.

Let Σ be a set equipped with a bijection $\varphi : \Sigma \to S - S^2$. This bijection is extended to a morphism $\varphi : \Sigma^+ \longrightarrow S$. Since S is generated by its indecomposable elements it follows that the morphism φ is surjective. Let $\sim$ be the congruence relation in Σ^+ defined by $w \sim w'$ iff $w\varphi = w'\varphi$. The following two properties of this congruence are easily verified

(4.5) Each $\sigma \in \Sigma$ is a congruence class in itself

(4.6) Each congruence class is finite.

Conversely given any congruence $\sim$ in Σ^+ satisfying (4.5) and (4.6), the quotient semigroup $S = \Sigma^+/\sim$ is loose. This shows the extent by which loose semigroups differ from free ones.

Example 4.1. Each free semigroup Σ^+ is loose. Each free commutative semigroup Σ_c^+ is loose.

Given any loose semigroup S, we denote by S° the monoid obtained from S by adjoining a unit element. Let K be any commutative semiring. An (S, K) grammar is defined to be an (S°, K)-grammar $G = (\Xi, \xi_0, R)$ which has no rules $\xi \to 1$. If further G has no rules $\xi \to \xi'$ then we say that G is a positive (S, K)-grammar.

We define the

$$S[\Xi] = S^\circ[\Xi] - 1$$

We note that $S[\Xi]$ is a loose semigroup and that for any word

$$w = m_0 \xi_{i_1} m_1 \cdots m_{p-1} \xi_{i_p} m_p$$

we have

$$|w| = p + |m_0| + \cdots + |m_p|$$

$ww'\delta = w\delta + w'\delta$

$\|ww'\| \geq \|w\| + \|w'\|$?

$\|w\tau\| = \|x \underline{\tau} y\|$

$\geq \|x\| + \|\underline{\tau}\| + \|y\| \geq 1 + \|x\| + \|\bar{\tau}\| + \|y\| = 1 + \|x\bar{\tau}y\|$

where $|m_i| = 0$ whenever $m_i = 1 \in S^\circ$. The integer p is called the degree of w and is denoted

$1 + \|\bar{\tau}\| \leq \|\underline{\tau}\|$

$\|w\| = 2|w| - w\delta$ $\quad 1 + \|w\| \leq \|w\tau\|$ $\quad w = x\xi y$ $\quad w\tau = x\underline{\tau}y$

by $w\delta$.

The fact that the grammar is positive may be expressed by the inequality

$$(4.7) \qquad 2 \leqslant 2|\underline{r}| - \underline{r}\delta$$

Indeed, if $\underline{r} = 1$ then $|\underline{r}| = 0 = \underline{r}\delta$ and if $\underline{r} = \xi$ then $|\underline{r}| = \underline{r}\delta = 1$. In all other cases the inequality (4.7) holds.

We claim that in a positive grammar the inequality

$$(4.8) \qquad |d| + \overset{\|u\|}{2|u| - u\delta} \leqslant \overset{\|ud\|}{2|w| - w\delta}$$

holds for any derivation

$$(4.9) \qquad d : u \to \cancel{\sigma}\, w$$

The inequality clearly holds if $d = 1$ and $u = w$. Assume that $d = d'r$ and that the inequality holds for d'. Thus d is the composition

$$u \xrightarrow{d} x\xi y \xrightarrow{\underline{r}} x\underline{r}y = w$$

H. $|d| + \|w\| \leq \|wd\|$

$|d| + \|u\| = 1 + |d'| + \|u\cancel{d}\|$ $\cancel{+\|d\|}$

$\leq 1 + \|ud'\| \leq \|ud'r\|$

Then

$$
\begin{aligned}
|d| + 2|u| - u\delta &= 1 + d' + 2|u| - u\delta \\
&\leqslant 1 + 2|x\xi y| - (x\xi y)\delta \\
&= 2 + 2|x| + 2|y| - x\delta - y\delta \\
&\leq 2 + 2|x\underline{\tau}y| - 2|\underline{\tau}| + (x\underline{\tau}y)\delta - \underline{\tau}\delta \\
&= 2|w| - w\delta - (2|\underline{\tau}| - \underline{\tau}\delta - 2) \\
&\leqslant 2|w| - w\delta
\end{aligned}
$$

as required.

Taking $w = \xi$ and $w = s \in S$ in (4.9) we obtain

Proposition 4.1. If

$$d : \xi \to s \,, \quad s \in S$$

is a derivation in positive (S, K)-grammar, then

$$|d| < 2|s| \quad \blacksquare$$

Corollary 4.2. For each $s \in S$, there is only a finite number of derivations $d : \xi \to s$ $\blacksquare$

Exercise 4.1. Let $\Sigma = \{\sigma\}$ be a single letter alphabet and let $A = 2\sigma$ be the N-subset of Σ^+ consisting of the letter σ with multiplicity 2. Show that A is positive algebraic, but that every positive grammar $G = (\Xi, \xi_0, R)$ defining A must contain a rule $\xi_0 \to \sigma$ with ~~multiplicity~~ weight 2. Compare this with Proposition 3.1

Exercise 4.2. For each of the sets A in Examples 2.2 – 2.3 construct a positive grammar for the set A with unit element removed.

Exercise 4.3. Show that if $\varphi : S \to S'$ is a morphism of loose semigroups then $|s| \leq |s\varphi|$ for all $s \in S$.

5. Linguistic interpretation; parsing

Temporary gap

pp. 29-34

6. Introduction to the fixed point approach.

In Chapters II and III we shall develop a completely different approach to algebraic sets, ~~completely~~ sidestepping derivations. To prepare the reader, we shall give here a heuristic discussion of a special case ~~that will nevertheless~~ which however is sufficient to give the reader a taste of things to come.

We return to the grammar

$$\xi \to \sigma\xi\tau \quad , \quad \xi \to 1$$

considered in Example I, 1.2 and defining the subset

$$A = \{\sigma^n\tau^n \mid n \geqslant 0\}$$

of $M = \{\sigma, \tau\}^*$.

We shall interpret ξ as representing an unknown subset X of M. The rule $\xi \to \sigma\xi\tau$ will thus be interpreted as the inclusion $\sigma X \tau \subset X$ and similarly the rule $\xi \to 1$ will be interpreted as $1 \in X$. Using $+$ instead of $\cup$ we thus obtain the inclusion

(1.1) $$\sigma X \tau + 1 \subset X$$

The set A does satisfy this inclusion but much more is true. Indeed let B be any subset of M satisfying (1.1)

Then $1 \in B$ and since $\sigma B \tau \subset B$ it follows inductively that $\sigma^n \tau^n \in B$ for all $n \geq 0$. Thus $A \subset B$, and A is the <u>minimal</u> solution of the inclusion. Next we note that for A we actually have

$$A = \sigma A \tau + 1$$

so that A also is the minimal solution of the equation

$$X = \sigma X \tau + 1$$

The process of passing from grammars to equations is not limited to the example above. The grammar

$$\xi \to \sigma\eta \, , \quad \xi \to \xi^2\tau \, , \quad \xi \to \sigma$$

$$\eta \to \xi\eta \, , \quad \eta \to \tau^2 \, , \quad \eta \to 1$$

in which ξ, η are variables and $M = \{\sigma, \tau\}^*$ leads to the system of equations

$$X = \sigma Y + X^2 \tau + \sigma$$

$$Y = XY + \tau^2 + 1$$

On the right hand side of each equation we have a polynomial (in the non-commutative sense) with X and Y as variables and with elements of M as coefficients. Jointly the two equations may be written as

(6.2) $\quad (X, Y) = (X, Y)\,\mathbb{P}$

where

$$\mathbb{P} : 2^M \times 2^M \longrightarrow 2^M \times 2^M$$

is a polynomial transformation in which the various monomial correspond to the rules of the grammar.

In Chapter II we shall treat positive (S,K)-grammers where S is loose semigroup and K is any commutative semiring. In this case it will be shown that the equation (6.2) has a unique solution, and that the first coordinate of this solution is the language defined by the grammar. This is the positive case

In Chapter III we shall define continuous semirings and study (M,K)-grammars where M is any monoid and K is a continuous semiring. The equation (6.2) may then have many solutions, but one of them is the smallest (in the sense of a partial order inherent in K). The language defined by the grammar is then the first coordinate of the minimal solution of (6.2). This is the <u>continuous</u> <u>case</u>.