Časopis pro pěstování matematiky

Domonkos Asztalos

A note on van Wijngaarden syntax

Časopis pro pěstování matematiky, Vol. 99 (1974), No. 4, 347--351

Persistent URL: http://dml.cz/dmlcz/117855

Časopis pro pěstování matematiky, roč. 99 (1974), Praha

A NOTE ON VAN WIJNGAARDEN SYNTAX

D. ASZTALOS, Budapest

(Received January 13, 1972)

In [1] a concept of a generalized grammar with an enumerable set of productions was introduced.

Some of the generalized grammars, called "ω-regular grammars", were considered and recursivness of languages generated by ω-regular grammars was proved.

The aim of this paper is to give a formal definition of van Wijngaarden's syntax and to establish a family of grammars as given by van Wijngaarden's syntax that would be ω-regular.

Definition 1. The *generalized grammar* is the 4-tuple $G = (V, T, R, \sigma)$ where V is an enumerable set of letters v_k, $k = 1, \ldots$; T is a recursive subset of V; R is a recursive set of ordered pairs $x \to y$ where either $x \in V \setminus T$, $y \in V^* - \{\varepsilon\}$, or $x = \sigma$, $y = \varepsilon$, ε is a null string; σ is an element of $V \setminus T$.

Definition 2. A generalized grammar G is said to be *ω-regular* if

(a) There exists a recursive function $b_G(x) \geqq 0$, $x \in V^*$ such that if $v_k \to x \in R$ then $k \leqq b_G(x)$.

(b) There exists an integer $s_G \geqq 0$ such that if $v_k \to v_m \in R$ and $k \geqq s_G$ then $m \geqq k$.

In [1] "Theorem 1." has been proved that is, if G is ω-regular then the language $L(G)$ is recursive.

The *van Wijngaarden's syntax* is a grammar where the enumerable set of productions (rules) is represented by means of finite number of production schemes containing some "metanotions" expandable by suitable rules into enumerable set of "notions". Below we are giving a formal definition of van Wijngaarden's syntax.

Definition 3. The *context-free grammar* without an initial symbol is a 3-tuple $H = (N, T, R)$ where N is a finite set of nonterminals; T is a finite set of terminals $(N \cap T) = \emptyset$; R is a finite set of ordered pairs $x \to y$ where $x \in N$, $y \in (N \cup T)^*$. The language generated by H is

$$L(H) = \{x : n \Rightarrow_H x\,, \quad n \in N\,, \quad x \in T^*\}\,.$$

Definition 4. The *rule-generator* is a 3-tuple $RG = (H, PR, T')$ where H is a CF grammar without an initial symbol; PR is a finite set of strings $\alpha \to \pi$ where $\alpha \in \in (N \cup T')^*$, $\pi \in (N \cup T' \cup \{,\})^*$, $T' = T \cup T_1$ where T_1 is a finite set of letters.

We can get rules from any $\alpha \to \pi \in PR$ by replacing each $n \in N$ in $\alpha \to \pi$ with their terminal productions. There is however a restriction: if a given $n \in N$ has two or more entries in a prerule $\alpha \to \pi \in PR$, then in a given derivation of a rule each entry of n must be replaced by the same terminal production.

Example 1.

$$\begin{aligned} H &= (N = \{D, E\}, T = \{s\}, R = \{D \to Ds; D \to s; D \to E; \\ &\quad E \to \varepsilon\}), \\ RG &= (H, PR = \{Dsa \to Da, a; Dsb \to Db, b; Dsc \to Dc, c; \\ &\quad s \to Da, Db, Dc; a \to \bar{a}; b \to \bar{b}; c \to \bar{c}\}, \\ T' &= T \cup \{a, b, c, \bar{a}, \bar{b}, \bar{c}\}. \end{aligned}$$

The enumerable set of rules wich can be generated by RG is

$$\begin{aligned} s &\to s^k a, s^k b, s^k c, \\ s^k a &\to s^{k-1} a, a, \\ s^k b &\to s^{k-1} b, b, \quad k = 0, 1, \dots, \\ s^k c &\to s^{k-1} c, c, \\ a &\to \bar{a}, \\ b &\to \bar{b}, \\ c &\to \bar{c}, \end{aligned}$$

The set of rules

$$L(RG) = \{a \to p : \alpha \to \pi \in PR, \alpha \to \pi \overset{S}{\underset{H}{\Rightarrow}} a \to p, a \in T'^*, \\ p \in (T' \cup \{,\})^*\}.$$

may be called a *rule-language generated by a rule generator RG*. The letter "s" indicates that the derivation is synchronous (See the restriction mentioned under Definition 4.)

Definition 5. The *van Wijngaarden's syntax* is a 4-tuple $W = (H, PR, T', \sigma)$ where $H = (N, T, R)$ is a CF grammar without an initial symbol. Any element of N is called metanotion (See [2].). PR is a finite set of strings $\alpha \to \pi$ (See Definition 4.). There may be some strings $\alpha \to \pi$ such that for them only a null-derivation in RG is valid, i.e. $\alpha \in T'^*$, $\pi \in (T' \cup \{,\})^*$.

T' is a finite set of letters as above; σ is a sentence-notion of W. $\sigma \in A$ where

$$A = \{a : \alpha \to \pi \in PR, \alpha \to \pi \Rightarrow^{S}_{H} a \to p, a \in T'^*\} .$$

A denotes the set of notions generated by RG. According to Definition 4., $W = (RG, \sigma)$.

Let us denote $T' \cup N \cup \{,\}$ by F. We write $\varphi \to_W \psi$ if $\varphi_1, \varphi_2 \in F^*$ and there exists a finite set of rules L_{RG} such that $\varphi = \varphi_1, a, \varphi_2, \psi = \varphi_1, p, \varphi_2$ where $a \to p \in L_{RG}$. We write $\varphi \Rightarrow_W \psi$ if there exists a sequence $(L_{RG}(0), \ldots, L_{RG}(n-1))$ such that $L_{RG}(k) \subseteq L_{RG}(k+1)$, $L_{RG}(i) \subset L(RG)$ and they are finite sets of rules for all $0 \leqq k \leqq n-1$; $(\varphi_0, \varphi_1, \ldots, \varphi_n)$ is a sequence where $\varphi_0 = \varphi$, $\varphi_n = \psi$ and

$$\varphi = \varphi_0 \xrightarrow[L_{RG}(0)]{} \varphi_1 \to \ldots \to \varphi_{n-1} \xrightarrow[L_{RG}(n-1)]{} \varphi_n = \psi .$$

The expression $\varphi_i \xrightarrow[L_{RG}(i)]{} \varphi_{i+1}$ means that there is a rule $a \to p \in L_{RG}(i)$ such that $\varphi_i = \varphi_{i1}, a, \varphi_{i2}$; $\varphi_{i+1} = \varphi_{i_1}, p, \varphi_{i_2}$; $\varphi_i, \varphi_{i_1}, \varphi_{i_2} \in F^*$, $0 \leqq i \leqq n-1$.
The language generated by W is

$$L(W) = \{w : \sigma \Rightarrow_W w,\ w \in (\tilde{T} \cup \{,\})^* ,$$
$$\text{where } \tilde{T} \subset (T'^* \times \{s\})^*,\ s \in T' \text{ is a fixed letter}\} .$$

For each W the set of metanotions can be divided into two parts: $N = N^f \cup N^\infty$ where N^f is the set of metanotions, each element of which has a finite set of terminal productions, and N^∞ stands for the set of metanotions each element of which has an infinite set of terminal productions.

Lemma 1. *For each van Wijngaarden's syntax there exists an equivalent van Wijngaarden's syntax such that* $N^f = \emptyset$. $L_W(RG)$ is finite when $N^\infty = \emptyset$ and W can be considered as finite provided that $L_W(RG)$ is finite, too.

Definition. 6 We will say that W is *ω-regular*, when for the set of notions $A = \{a^i_{j,k}\}$ of W there exists such a numeration $\mathfrak{N}$, that for each rule

$$a^i_{0,k} \to a^i_{1,k}, a^i_{2,k}, \ldots, a^i_{l,k} \in L_W(RG)$$

the condition to be ω-regular is satisfied.

Lemma 2. *For W to be ω-regular it is necessary that if $\alpha \to \pi \in PR$, $\alpha = \alpha_1\, n\, \alpha_2$, $n \in N^\infty$ then $\pi = \pi_1\, n\, \pi_2$.*

Proof. Let the condition of Lemma 2. be not satisfied. Then exists such a rule $\alpha \to \pi \in PR$, $\alpha = \alpha_1\, n\, \alpha_2$, $n \in N^\infty$ that n has no entry in π.

In this case the left and right sides of the pre-rule are independent one from the other. We can thus obtain a fixed right hand side and a varying left hand side, i.e. the rules

$$a_{n_i} \to p \in L(RG), \quad i = 1, 2, \ldots$$

If $b_W(p) = K$ then for the first $K + 1$ elements of sequence $\{n_i\}$ we can find a number such that $N = \max_{1 \leq i \leq K} n_i$. In this case is true that $N > K$.

It means there exists such a rule in $L(RG)$

$$v(n_i) \to p \quad \text{that} \quad n_i > b_W(p).$$

The condition of Lemma 2. isn't sufficient for W to be ω-regular (see the following example).

Example. In W there are two pre-rules:

$$Ds \to Dss,\ Dss \to Ds,\ D \in N^\infty,\ s \in T.$$

Then for all $s_G \geqq 0$ we can find a number k where $k > s_G$, $v_k \to v_m \in L(RG)$ and $k > m$.

Definition 7. Let $\alpha_0^i \to \alpha_1^i, \alpha_2^i, \ldots, \alpha_l^i \in PR$, then $D(\alpha_j^i)$ is the set of terminal productions derived from α_j^i by RG.

Let us see pre-rules where $l = 1$ and I denotes the set of numbers of such rules.

Theorem. *Let $D(\alpha_0^i) \cap D(\alpha_1^j)$ be a finite set for each pair $i, j \in I$. If $\alpha \to \pi \in PR$ and $\alpha = \alpha_1\, n\, \alpha_2$ then $\pi = \pi_1\, n\, \pi_2$ where $n \in N^\infty$; $\alpha_1, \alpha_2 \in T'^*$; $\pi_1, \pi_2 \in (T' \cup \{,\})^*$. If this two conditions are true for a W then W is ω-regular.*

Proof. A. The set of notions $A = \{a_{j,k}^i\}$ is recursively ennumerable. If the second condition of the theorem is true, there exists such a numeration $\mathfrak{N}$ of A that if for a finite sequence of positive integers $n_{i_1}, n_{i_2}, \ldots, n_{i_l}$ is true that $v(n_{i_k}) \to p_{n_i} \in L(RG)$ for $k = 1, 2, \ldots, l$, then

$$n_{i_{k+1}} - n_{i_k} = 1\,, \quad k = 1, 2, \ldots, l - 1\,.$$

Let us fix such a numeration $\mathfrak{N}$ of A. Then for each pair $i, j \in I$ exists

$$N_{i,j} = \max\ \{n_k : v(n_k) \in D(\alpha_0^i) \cap D(\alpha_1^j)\}$$

because $D(\alpha_0^i) \cap D(\alpha_1^j)$ is finite.

Let assume that $N = \max_{i,j \in I} N_{i,j}$, then for $s_G = N$ the second condition of ω-regularity is satisfied.

B. Let $b_W(x) \geqq 0$ be a recursive function defined on A^*. If for $\mathfrak{N}$ and $b_W(x)$ the first condition of ω-regularity is not complied with, then there are two cases, which we must consider.

Let $S = \{\langle n_i, b_W(p_{n_i})\rangle\}$ denote a sequen cesuch pairs of numbers, that $n_i > b_W(p_{n_i})$.

1st case. S is finite. If $N = \max_{n_i \in S} n_i$ then $\bar{b}_W(x) = b_W(x) + N \geqq 0$ and is recursive. For $\mathfrak{N}$ and $\bar{b}_W(x)$ the first condition of ω-regularity is satisfied.

2nd case. Let S be infinite. Since the second condition of theorem is true, for each $b_W(p_{n_i})$ there exists only a finite set of pairs P of S such that for each pair of P the $b_W(p_{n_i})$ is the same.

Let $|P|$ denote the number of elements in P. We can effectively define $|P|$ because H is context-free. It means, we can define a number $b'(p)$ for each p:

$$b'(p_{n_i}) = n_{i_1} + l - 1 .$$

Let assume

$$\tilde{b}_W(x) = \begin{cases} b_W(x) & \text{if} \quad n_i \leqq b(p_{n_i}) \\ b'(p_{n_i}) & \text{if} \quad n_i > b(p_{n_i}) \end{cases}$$

then for $\mathfrak{N}$ and $\tilde{b}_W(x)$ the first condition of ω-regularity is satisfied.

The part A and part B together prove the theorem.

References

[1] *Mazurkiewicz, A. W.*, A Note on Ennumerable Grammars. Information and Control *14*, 555—558 (1969).

[2] *van Wijngaarden, A.* (ed.) (1968). Draft Report on the Algorithmic Language ALGOL 68. Matematisch Centrum, MR 93, Amsterdam.

Author's address: Computer Centre of Hungarian Planning Office, Budapest, Angol u. 27, Hungary.