About Postfix Expressions

Lector dr. Florin Fortiş
Universitatea de Vest din Timişoara
Asist. Alexandra Fortiş
Universitatea "Tibiscus" din Timişoara

ABSTRACT. In this paper we study the problem of equivalence between the languages of infix and postfix expressions over a universal algebra (A,Ω) of type τ, having a finite domain of operators. Based on these results one can establish a result of characterization for the closure operator associated with a universal algebra (A,Ω) in terms of formal languages.

1 Introduction

Let (A,Ω) be a universal algebra of type τ, with a finite the domain of operators, and $\tau(\omega)=2,\forall\omega\in\Omega$ (that is, the arity of any operator $\omega\in\Omega$ is 2).

Definition 1.1 *The infix (postfix) expressions over the universal algebra* (A,Ω) *of type* τ*, can be defined as follows:*

1. *For every* $a\in A$*,* a *is an infix (postfix) expression.*
2. *If* e_1 *and* e_2 *are infix (postfix) expressions, then* $(e_1\omega e_2)$ *(*$e_1e_2\omega$*, respectively) is an infix (postfix) expression.*

Definition 1.2 *Let* e_i^1 *and* e_i^2 *be infix expressions, and* e_p^1 *and* e_p^2 *postfix expressions. The postfix version of a given infix expression can be defined as follows:*

1. *For every* $a\in A$*,* a *is a postfix version of the infix expression a.*
2. *Let* e_p^1 *be a postfix version of the infix expression* e_i^1*,* e_p^2 *be a*

postfix version of the infix expression e_i^2*, and the binary operator* $\omega \in \Omega$*. Then* $e_p^1 e_p^2 \omega$ *is a postfix version for the infix expression* $(e_i^1 \omega e_i^2)$*.*

Next, we will prove that every infix expression has a unique postfix version, and that also the reversely holds.

2 The grammars of infix-like and postfix-like expressions

Let $V_N = \{X_0\}$, $V_T = \{e\}$, $V_{()} = \{(,)\}$. Define now the grammars

$$G_i = (V_N, V_T \cup V_{()} \cup \Omega, X_0, F_i)$$

and

$$G_p = (V_N, V_T \cup \Omega, X_0, F_p)$$

where

$$F_i = \{X_0 \to X_0 \omega X_0, \forall \omega \in \Omega\} \cup \{X_0 \to e\}$$

and

$$F_p = \{X_0 \to X_0 X_0 \omega, \forall \omega \in \Omega\} \cup \{X_0 \to e\}.$$

Lemma 2.1 *Let* $\sigma : V_T \cup V_{()} \cup \Omega \to \wp(A) \cup V_{()} \cup \Omega$ *be defined by* $\sigma(e) = A$*,* $\sigma(\omega) = \omega$*,* $\forall \omega \in \Omega$ *and* $\sigma(p) = p$*, for* $p \in V_{()}$*. Then the sets* $\sigma(L(G_i))$ *and* $\sigma(L(G_p))$ *represent the sets of all infix (postfix, respectively) expressions over the universal algebra* (A, Ω).

Proof: It follows immediately from the definition of infix (postfix) expressions and the definitions for the grammars G_i and G_p, as given above.

3 The unambiguosity of the studied grammars

Recall that a grammar is unambiguous if every generated word possesses exactly one leftmost derivation according to the rules of the grammar (here we use only generative devices).

Proposition 3.1 *The grammar* G_i *is unambiguous. Consequently, the language* $L(G_i)$ *is unambiguous.*

Proof: Let us define the application $Op: L(G_i) \to \mathbf{N}$, where $Op(e_i)$ represents the number of operators in the expression $e_i \in L(G_i)$.

The only expression in $L(G_i)$ with $Op(e_i) = 0$ is $e_i = e$, and there is a unique derivation of the expression e_i according to the rules in the grammar G_i, that is $S \Rightarrow e_i = e$.

Moreover, every expression $e_i \in L(G_i)$ with $Op(e_i) = 1$ possesses a unique leftmost derivation. Indeed, according to the second rule for infix expressions generation, every infix expression with $Op(e_i) = 1$ has the form $e_i = (e\omega_1 e)$, with $\omega_1 \in \Omega$, and the only leftmost derivation for the expression e_i is

$$X_0 \Rightarrow (X_0\omega_1 X_0) \Rightarrow (e\omega_1 X_0) \Rightarrow (e\omega_1 e)$$

Let us now presume that for every infix expression $e_i^{'} \in L(G_i)$ with $Op(e_i^{'}) \le k$ there is a unique leftmost derivation, and let $e_i \in L(G_i)$, with $Op(e_i) = k+1$. The first step in the derivation for e_i must be $S \Rightarrow (S\omega_1 S)$.

Thus, we can decompose the expression e_i as follows $e_i = (e_l \omega_1 e_r)$, where $e_l, e_r \in L(G_i)$, with $Op(e_l) \le k$ and $Op(e_r) \le k$. That is, the expressions e_l and e_r posses unique leftmost derivations, so there exists at least one leftmost derivation for the expression $e_i = (e_l \omega_1 e_r)$,

$$X_0 \Rightarrow (X_0\omega_1 X_0) \Rightarrow^* (e_l\omega_1 X_0) \Rightarrow^* (e_l\omega_1 e_r) = e_i$$

Presume now that there exists another leftmost derivation for the expression e_i, for example

$$X_0 \Rightarrow (X_0\omega_1 X_0) \Rightarrow^* (e_l^{'}\omega_1 X_0) \Rightarrow^* (e_l^{'}\omega_1 e_r^{'}) = e_i \qquad (1)$$

with $Op(e_l) < Op(e_l^{'})$. Observe first that, if there exists another derivation of the form

$$X_0 \Rightarrow (X_0\omega_1^{'} X_0) \Rightarrow^* (e_l^{'}\omega_1^{'} X_0) \Rightarrow^* (e_l^{'}\omega_1^{'} e_r^{'}) = e_i \qquad (2)$$

then $\omega_1^{'} = \omega_1$, $e_l^{'} = e_l$ and $e_r^{'} = e_r$. Indeed, if there exists the derivations (1) and (2) for the infix expression e_i, then there exist the infix expressions e_l, $e_l^{'}$, e_r and $e_r^{'}$. Now, if $\omega_1 \ne \omega_1^{'}$, we have either $e_l = e_l^{'}\omega_1^{'} e_r^{''}$, or $e_l^{'} = e_l \omega_1 e_r^{'}$, with $e_r^{''} \in L(G_i)$ and $e_r^{'} = e_r^{''} e_r^{1}$ in the first case, and $e_r = e_r^{''} e_r^{1}$ in the second case. Obviously, because $e_r^{'} \in L(G_i)$, e_r^{1} must begin with an operator $\omega \in \Omega$, and end with a right unmatched parenthesis. However, because $e_r^{"} \in L(G_i)$, such a construction cannot exist.

Now, if there exist two different leftmost derivations of the type (1) for the infix expression e_i, then either e_l possesses two different leftmost derivations, or e_r possesses two different leftmost derivations.

Observation 3.2 *If in the definition for infix expressions we use rules of the form* $X_0 \to X_0 \omega X_0$ *instead of* $X_0 \to (X_0 \omega X_0)$, *the grammar can easily lose its property of being unambiguous.*

Indeed, let G_a be the grammar

$$G_a = (\{X_0\}, \{a,+\}, X_0, \{X_0 \to a, X_0 \to X_0 + X_0\}).$$

Then the word $a + a + a$ posses at least two leftmost derivations:

$$X_0 \Rightarrow X_0 + X_0 \Rightarrow a + X_0 \Rightarrow a + X_0 + X_0 \Rightarrow^* a + a + a$$

$$X_0 \Rightarrow X_0 + X_0 \Rightarrow X_0 + X_0 + X_0 \Rightarrow a + X_0 + X_0 \Rightarrow^* a + a + a$$

4 The result of equivalence

Proposition 4.1 *For every infix expression* $e_i \in L(G_i)$, *there exists at least one postfix expression* $e_p \in L(G_p)$, *representing a postfix version of the original infix expression.*

Proof: Let us consider the matricial grammar [, pp.143], []

$$G_{i,p}^m = (\{X_0, X_0^i, X_0^p\}, \{e,c\} \cup \{(,)\} \cup \Omega, X_0, F_{i,p}^m)$$

where

$$F_{i,p}^m = \{[X_0 \to X_0^i c X_0^p], [X_0^i \to e, X_0^p \to e]\} \cup \{[X_0^i \to (X_0^i \omega X_0^j), X_0^p \to X_0^p X_0^p \omega], \forall \omega \in \Omega\}$$

and let $L_{left}(G_{i,p}^m)$ be the language generated by the matricial grammar $G_{i,p}^m$ under leftmost restriction on derivations (see [, pp.146]).

Observe now that if $e_i \in L(G_i)$, there exists some words of the form $e_i c e_p$ in $L_{left}(G_{i,p}^m)$. By considering only leftmost derivations in the matricial grammar, we obtain that for every infix expression in $L(G_i)$ there exist at least one postfix expression $e_p^i \in L(G_p)$ such that $e_i c e_p^i$ belongs to

$L_{left}(G_{i,p}^m)$. Also, observe that such a leftmost derivation in $G_{i,p}^m$ corresponds with the definition for infix versions of a postfix expression.

Proposition 4.2 *The grammar G_p is unambiguous. Consequently, the language $L(G_p)$ is unambiguous.*

Proof: Define the application $Op : L(G_p) \to \mathbf{N}$, where $Op(e_p)$ represents the number of operators in the postfix expression e_p.

The only expression with $Op(e_p) = 0$ is $e_p = e$, and it has a unique derivation, namely $X_0 \Rightarrow e$.

If e_p is a postfix expression with $Op(e_p) = 1$, then there exists a unique leftmost derivation for e_p,

$$X_0 \Rightarrow X_0 X_0 \omega_1 \Rightarrow e X_0 \omega_1 \Rightarrow e e \omega_1 \text{ X\_\{0\}}$$

Let now e_p be an expression with $Op(e_p) = k+1$, and let us presume that for every expression $e_p^{'}$ with $Op(e_p^{'}) \le k$, there exists a unique leftmost derivation. The first step in the derivation for e is $X_0 \Rightarrow X_0 X_0 \omega_1$, so we can decompose the postfix expression e_p as follows: $e_p = e_f e_l \omega_1$, where $e_f, e_l \in L(G_p)$, and $Op(e_f) \le k$ and $Op(e_l) \le k$, so there exist, by the hypothesis we have made, unique leftmost derivations for e_f and e_l. Then the derivation

$$X_0 \Rightarrow X_0 X_0 \omega_1 \Rightarrow^* e_f X_0 \omega_1 \Rightarrow^* e_f e_l \omega_1$$

is a leftmost derivation for $e_p \in L(G_p)$.

Let us presume now that there exists another leftmost derivation for e_p. Then there can exist two different leftmost derivations of the form

$$X_0 \Rightarrow X_0 X_0 \omega_1 \Rightarrow^* e_f X_0 \omega_1 \Rightarrow^* e_f e_l \omega_1$$

That is, either e_f possesses two different leftmost derivations, or e_l possesses two different leftmost derivations.

Corrolary 4.3 *For every word $e_i \in L(G_i)$ there exists exactly one word $e_p \in L(G_p)$, representing the unique postfix version of the infix expression e_i.*

Proof: Consider again the matricial grammar $G_{i,p}^m$, and consider only leftmost derivations. From the Propositions 3.1 and 4.2 we easily deduce that there exists a unique leftmost derivation for a word of the form $e_i \omega e_p$. More, it follows from Proposition 4.1 that e_p is a postfix version for e_i, and due to the fact that the grammars G_i and G_p are unambiguous we easily deduce that it is the unique postfix version for e_i.

Now we can easily deduce the following result:

Theorem 4.4 *The sets* $L(G_i)$ *and* $L(G_p)$ *are isomorphic, and the isomorphism is emphasized by* $\sigma(e_i) = e_p$, *if* $e_i c e_p \in L(G^m_{i,p})$.

Consequence 4.5 *The set of all the postfix expressions over the universal algebra* (A,Ω), *of type* τ, *with a finite domain of binary operators,* Ω, *is isomorphic with the set of all the infix expressions over the universal algebra* (A,Ω), *of type* τ, *with a finite domain of binary operators.*

References

[Jeb84] Jebeleanu, T., A theory on postfix notation, Seminarul de Informatică şi Analiză Numerică (SIAN), 18 (1984), Universitatea din Timişoara

[Pau81] Păun, Gh., Gramatici matriciale, Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1981

[PP77] Purdea, I.; Pic, Gh., Tratat de algebră modernă (Vol. 1), Editura Academiei, Bucureşti, 1977

[Sal73] Salomaa, A., Formal Languages, Academic Press, New York, 1973