

Programski jezik PINS'26

1 Leksikalna pravila

Programski jezik PINS'26 je pisan v abecedi ASCII in vsebuje naslednje leksikalne enote:

- *Konstante:*

- *število:*

- Neprazno zaporedje desetiških števk, pred katerimi lahko stoji predznak (+ or -).

- *znak:*

- Znak zapisan v enojnih navednicah ('). Znak je lahko (a) katerikoli ASCII znak s kodo v obsegu {32...126}, pri čemer morata biti enojna navednica in obratna poševnica (\) uvedeni z obratno poševnico, (b) znak za konec vrstice, ki je zapisan kot \n, ali (c) katerikoli ASCII znak zapisan s kodo v obliki \xx, pri čemer je X šestnajstiška števka (0...9 in a...f).

- *niz znakov:*

- Niz znakov, lahko prazen, zapisan v dvojnih navednicah ("). Znak je lahko (a) katerikoli ASCII znak s kodo v obsegu {32...126}, pri čemer morata biti dvojna navednica in obratna poševnica (\) uvedeni z obratno poševnico, (b) znak za konec vrstice, ki je zapisan kot \n, ali (c) katerikoli ASCII znak zapisan s kodo v obliki \xx, pri čemer je X šestnajstiška števka (0...9 in a...f).

- *Simboli:*

- = ; , && || ! == != > < >= <= + - * / % ^ ()

- *Imena:*

- Neprazno zaporedje črk (A...Z in a...z), desetiških števk (0...9) in podčrtajev (_), ki (a) se začne s črko ali podčrtajem in (b) ni ključna beseda.

- *Ključne besede:*

- fun var if then else while do let in end

- *Komentarji:*

- Niz znakov, ki se začne z // in se konča na koncu vrstice.

- *Belo besedilo:*

- Presledek in znaki HT, LF in CR. Konec vrstice označuje znak LF, HT je širok 8 znakov.

Leksikalni elementi morajo biti razpoznani od leve proti desni po pravilu najdaljšega ujemanja na skrajno levem mestu.

2 Sintaksna pravila

Sintakso programskega jezika PINS'26 določa kontekstno neodvisna gramatika z začetnim simbolom *program* in naslednjimi produkcijami:

program
→ (*definition*)*

definition

- `fun IDENTIFIER (parameters)`
- `fun IDENTIFIER (parameters) = statements`
- `var IDENTIFIER = initializers`

parameters

- `(IDENTIFIER (, IDENTIFIER) *) ?`

statements

- `statement ; (statement ;) *`

statement

- `expression`
- `expression = expression`
- `if expression then statements (else statements) ? end`
- `while expression do statements end`
- `let (definition) + in statements end`

expression

- `INTCONST | CHARCONST | STRINGCONST`
- `IDENTIFIER ((arguments)) ?`
- `prefix-operator expression`
- `expression postfix-operator`
- `expression binary-operator expression`
- `(expression)`

arguments

- `(expression (, expression) *) ?`

initializers

- `(initializer (, initializer) *) ?`

initializer

- `(INTCONST *) ? const`

const

- `INTCONST | CHARCONST | STRINGCONST`

Simboli INTCONST, CHARCONST in STRINGCONST označujejo celoštevilске konstante, znakovne konstante in nize, zaporedoma.

Prioriteta operatorjev je sledeča

<i>postfiksni operatorji</i>	<code>^</code>	NAJVIŠJA PRIORITETA
<i>prefiksni operatorji</i>	<code>! + - ^</code>	
<i>multiplikativni operatorji</i>	<code>* / %</code>	
<i>aditivni operatorji</i>	<code>+ -</code>	
<i>primerjalni operatorji</i>	<code>== != < > <= >=</code>	
<i>konjunkcija</i>	<code>&&</code>	
<i>disjunkcija</i>	<code> </code>	NAJNIŽJA PRIORITETA

Multiplikativni in aditivni operatorji so levo asociativni, primerjalni operatorji niso asociativni.

3 Semantična pravila

Imena. Vsa imena so v istem imenskem prostoru.

Vsako ime je vidno v celotnem območju dosega, v katerem je definirano. Območje dosega je ustvarjeno na tri načine:

1. Vse definicije imen zunaj funkcij so v globalnem območju dosega.
2. Funkcija ustvari novo območje dosega (ime funkcije je zunaj ustvarjenega dosega).
3. Stavek `let-in` ustvari novo območje dosega (vse lokalne definicije so znotraj območja obsega).

Tipi. Obstaja en sam podatkovni tip: 32-bitno predznačeno število v dvojiškem komplementu, ki služi kot vrednost in kot naslov. Znakovne konstante predstavljajo cela števila glede na ASCII kodiranje. Nizi znakov predstavljajo naslov, na katerem so shranjena števila, ki so dobljena iz znakov niza glede na ASCII kodiranje.

Zaporedje stavkov, ki tvori telo funkcije, se mora končati

1. s stavkom, ki vsebuje zgolj izraz, ali
2. z `let-in` stavkom, katerega zaporedje stavkov se konča s stavkom pod točko 1 ali 2.

Leve vrednosti. Leva vrednost je

1. spremenljivka ali
2. izraz, katerega najbolj zunanji operator je postfiksni `^`.

Operacijska semantika.

1. V prireditvenem stavku se najprej izračuna vrednost na desni strani, potem naslov na levi strani.
2. Pogoj v pogojnem stavku in zanki je izpolnjen, če je njegova vrednost različna od 0.
3. Pri izračunu izraza z dvomestnim operatorjem se najprej izračuna levi, nato desni argument.
4. Pri klicu funkcije se argumenti izračunajo od desne proti levi.

Skladovni stroj. Skladovni stroj je 32-biten: naslovi, celoštevilске vrednosti in registri so 32-bitni.

Skladovni stroj ima tri registre:

1. programski števec (IP),
2. skladovni kazalec (SP),
3. klicni kazalec (FP).

Ukazi skladovnega stroja so naslednji:

1. **LOAD:** z vrha sklada vzame naslov in na vrh sklada naloži vrednost, ki je na tem naslovu.

```
addr ← MEM[SP], SP++  
value ← MEM[addr]  
SP--, MEM[SP] ← value
```

2. **SAVE:** z vrha sklada vzame naslov in vrednost ter shrani vrednost na ta naslov.

```
addr ← MEM[SP], SP++  
value ← MEM[SP], SP++  
MEM[addr] ← value
```

3. POPN: z vrha sklada vzame vrednost n in
z vrha sklada umakne $|n|$ vrednosti (če $n \geq 0$)
ali na vrh sklada naloži nanj $|n|$ vrednosti 0 (če $n < 0$).

$$n \geq 0 \implies \forall i = 1 \dots \lfloor n/4 \rfloor: SP++$$

$$n < 0 \implies \forall i = 1 \dots \lfloor n/4 \rfloor: SP--, MEM[SP] \leftarrow 0$$
4. PUSH value: na vrh sklada naloži vrednost value.

$$SP--, MEM[SP] \leftarrow value$$
5. NAME label: na vrh sklada naloži naslov addr, ki ga določa oznaka label.

$$SP--, MEM[SP] \leftarrow addr$$
6. REGN label: na vrh sklada naloži vrednost registra r.

$$SP--, MEM[SP] \leftarrow r$$
7. OPER op: z vrha sklada vzame en ali dva operanda in na vrh sklada naloži rezultat operacije o.
 enomestna operacija:

$$oper_1 \leftarrow MEM[SP], SP++$$

$$result \leftarrow op\ oper_1$$

$$SP--, MEM[SP] \leftarrow result$$
 dvomestna operacija:

$$oper_2 \leftarrow MEM[SP], SP++$$

$$oper_1 \leftarrow MEM[SP], SP++$$

$$result \leftarrow oper_1\ op\ oper_2$$

$$SP--, MEM[SP] \leftarrow result$$
8. UJUMP: z vrha sklada vzame novo vrednost programskega števca.

$$addr \leftarrow MEM[SP], SP++$$

$$PC \leftarrow addr$$
9. CJUMP: z vrha sklada vzame vrednost in dva naslova ter v programski števec vpiše enega od naslovov glede na vrednost: če je ta ničelna, v programski števec vpiše prvega, sicer drugega.

$$addr_1 \leftarrow MEM[SP], SP++$$

$$addr_2 \leftarrow MEM[SP], SP++$$

$$value \leftarrow MEM[SP], SP++$$

$$value = 0 \implies PC \leftarrow addr_1$$

$$value \neq 0 \implies PC \leftarrow addr_2$$
10. CALL: z vrha sklada vzame naslov, shrani klicni kazalec in programski števec, nastavi klicni kazalec in v programski števec vpiše naslov.

$$addr \leftarrow MEM[SP], SP++$$

$$SP--, MEM[SP] \leftarrow FP$$

$$SP--, MEM[SP] \leftarrow IP + 4$$

$$FP \leftarrow SP + 8$$

$$IP \leftarrow addr$$
11. RETN: z vrha sklada vzame velikost parametrov in rezultat funkcije, restavrira klicni in skladovni kazalec, nastavi programski števec, umakne parametre z vrha sklada in na vrh sklada potisne rezultat.

$$size \leftarrow MEM[SP], SP++$$

$$result \leftarrow MEM[SP], SP++$$

$$SP \leftarrow FP$$

$$FP \leftarrow MEM[SP - 4]$$

$$PC \leftarrow MEM[SP - 8]$$

$SP \leftarrow SP + \text{size} + 4$
 $SP--, \text{MEM}[SP] \leftarrow \text{result}$

12. INIT: z vrha sklada vzame naslov opisa začetne vrednosti in naslov spremenljivke ter inicializira spremenljivko z opisano začetno vrednostjo.

$\text{init} \leftarrow \text{MEM}[SP], SP++$
 $\text{addr} \leftarrow \text{MEM}[SP], SP++$
 $\text{MEM}[\text{addr} \dots] \leftarrow \text{začetna vrednost}$

Začetna vrednost na naslovu je init opisana kot

$$n \ n_1 \ l_1 \ d_{1,1} d_{1,2} \dots d_{1,l_1} \ n_2 \ l_2 \ d_{2,1} d_{2,2} \dots d_{2,l_2} \dots n_n \ l_n \ d_{n,1} d_{n,2} \dots d_{n,l_n},$$

na naslovu addr pa se razvije v

$$(d_{1,1} d_{1,2} \dots d_{1,l_1})^{n_1} (d_{2,1} d_{2,2} \dots d_{2,l_2})^{n_2} \dots (d_{n,1} d_{n,2} \dots d_{n,l_n})^{n_n}.$$

Psevdoukazi skladovnega stroja so naslednji:

1. LABEL label: oznaka label.
2. SIZE size: odmeri size bajtov pomnilnika.
3. DATA value: 32-bitna konstanta value v pomnilniku.

(V primeru dvomov pogledajte v `Machine.java`.)