

1 Specifikacija sintakse programskih jezika, meta jezici

Za opis programskih jezika često se koriste kontekstno-slobodne gramatike. Bekus-Naurova forma (BNF) je konvencija za zapisivanje pravila kontekstno-slobodnih jezika. Proširena Bekus-Naurova forma (EBNF) dodaje određene sintaksičke izraze BNF notaciji i omogućava jednostavniji zapis pravila gramatike.

1.1 BNF (Backus-Naur form)

Meta jezik je jezik koji služi da se pomoću njega opiše neki drugi jezik ili isti taj jezik. Tako se na primer služimo srpskim jezikom da bismo opisali gramatiku srpskog jezika. Bitno je razlikovati term meta jezika od terma jezika koji se opisuje.

BNF (Bekus-Naurova forma) je formalni meta jezik za predstavljanje kontekstno-slobodnih gramatika odnosno gramatika programskih jezika.

U BNF se koriste sledeće konvencije za zapisivanje pravila gramatike:

- Umesto simbola \rightarrow koristi se simbol $::=$
- Pomoćni simboli se navode među zagradama $\langle i \rangle$
- Vertikalna crta $|$ (izbor) razdvaja desne strane pravila koja odgovaraju istom simbolu sa leve strane pravila
- Prazna niska ϵ se zapisuje kao $\langle \text{empty} \rangle$.

Sledi nekoliko primera BNF-a :

- BNF za cifru:

```
<cifra> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0
```

- BNF za neoznačen ceo broj:

```
<neoznaceniCeoBroj> ::= <cifra> | <cifra> <neoznaceniCeoBroj>
<cifra> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0
```

Drugo pravilo može da se napiše i kao:

```
<neoznaceniCeoBroj> ::= <cifra> | <neoznaceniCeoBroj> <cifra>
```

- BNF za ceo broj:

```
<ceoBroj> ::= <neoznaceniCeoBroj>
           | + <neoznaceniCeoBroj>
           | - <neoznaceniCeoBroj>
<neoznaceniCeoBroj> ::= <cifra> | <cifra> <neoznaceniCeoBroj>
<cifra> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0
```

- BNF za realne brojeve:

```
<realanBroj> ::= - <neoznaceniRealanBroj>
              | + <neoznaceniRealanBroj>
              | <neoznaceniRealanBroj>
<neoznaceniRealanBroj> ::= <neoznaceniCeoBroj>
                          | <neoznaceniCeoBroj> . <neoznaceniCeoBroj>
<neoznaceniCeoBroj> ::= <cifra> | <cifra> <neoznaceniCeoBroj>
<cifra> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0
```

- BNF za identifikator (niska cifara ili slova koja ne sme počinjati cifrom):

```
<identifikator> ::= <slovo>
                  | <identifikator> <slovo>
                  | <identifikator> <cifra>
<cifra> ::= 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0
```

Primetimo da ovde ne smemo da zamenimo mesta pomoćnim simbolima (proverite!).

1.2 EBNF (Extended Backus-Naur form)

Proširena Bekus-Naurova forma ima istu izražajnu moć kao BNF s tim što su u njoj izvršene izmene koje doprinose čitljivosti zapisu pravila. Zapravo, BNF koristi rekurziju (definisanje pomoćnog simbola preko njega samog) da bi se izrazile relacije a EBNF koristi iteraciju.

Konvencije za EBNF su¹:

- Pomoćni simboli se zapisuju velikim početnim slovom
- Završni simboli se zapisuju pod jednostrukim navodnicima ako se sastoje od jednog karaktera, a crnim slogom (engl. *bold*) ako su višeslovni. Npr. '+' ali div ili integer
- Oble zagrade () se koriste za grupisanje
- Vitičaste zagrade { i } ograđuju deo koji se ponavlja 0 ili više puta
- Uglaste zagrade [i] opisuju opcionu konstrukciju
- Značenje ostalih oznaka je isto kao kod BNF.

¹Po nekim konvencijama se koriste sledeći sufixi: ? umesto [], * umesto {}, + sa značenjem jedno ili više pojavljivanja simbola.

Sve ove konstrukcije mogu biti izražene i u BNF-u što je pokazatelj toga da BNF i EBNF imaju istu izražajnu moć.

Sledi nekoliko primera EBNF-a :

- EBNF za neoznačen ceo broj.

```
NeoznaceniCeoBroj ::= Cifra{Cifra}
Cifra ::= '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9' | '0'
```

- EBNF za ceo broj.

```
CeoBroj ::= [ '+' | '-' ] NeoznaceniCeoBroj
NeoznaceniCeoBroj ::= Cifra{Cifra}
Cifra ::= '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9' | '0'
```

Pri čemu se umesto prva dva pravila može pojaviti pravilo:

```
CeoBroj ::= [ '+' | '-' ] Cifra{Cifra}
```

- EBNF za realne brojeve.

```
RealanBroj ::= [ '+' | '-' ] Cifra{Cifra} [ '.' Cifra{Cifra} ]
Cifra ::= '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9' | '0'
```

- EBNF za identifikator.

```
Identifikator ::= Slovo{Slovo|Cifra}
```

Zadatak 1

Napisati BNF/EBNF za aritmetički izraz.

Rešenje:

BNF:

```
<izraz> ::= <izraz> + <term> | <izraz> - <term> | <term>
<term> ::= <term> * <faktor> | <term> / <faktor> | <faktor>
<faktor> ::= (<izraz>) | <neoznaceniCeoBroj>
```

EBNF:

```
Izraz ::= Term{('+'|'-')Term}
Term ::= Faktor{('*'|'/')Faktor}
Faktor ::= '('Izraz')'|NeoznaceniCeoBroj
```

1.3 Sintaksni dijagrami

Sintaksni dijagrami predstavljaju pravila gramatike posredstvom grafova. Imaju istu izražajnu moć kao BNF i EBNF. Konstruišu se na sledeći način:

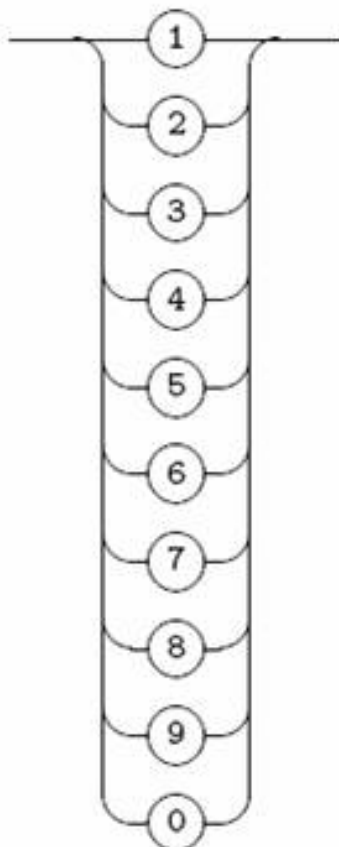
- Svakom pravilu gramatike dodeljuje se po jedan graf koji nosi ime pomoćnog simbola sa leve strane pravila
- Čvorovi grafa su obeleženi simbolima desne strane pravila i to pravougaonikom ako odgovaraju pomoćnom simbolu gramatike, a elipsom ako su pridruženi završnom simbolu gramatike. Lukovi u grafu povezuju čvorove na način opisan samim gramatičkim pravilom.
- Jezik predstavljen grafom se dobija obilaskom puteva u grafu.

Pogledajmo kako izgledaju sintaksni dijagrami nekoliko već pomenutih gramatika:

- Sintaksni dijagram za *Cifra*:

Cifra ::= '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9' | '0'

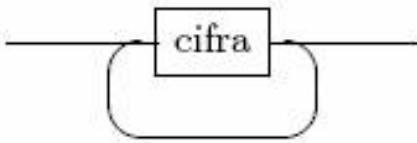
Cifra



- Sintaksni dijagram za NeoznacenoBroj:

$$\text{NeoznacenoBroj} ::= \text{Cifra}\{\text{Cifra}\}$$

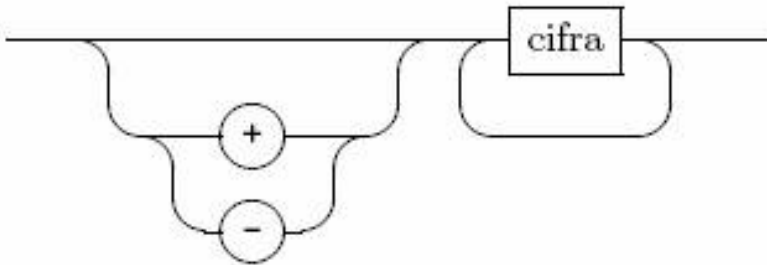
NeoznacenoBroj



- Sintaksni dijagram za CeoBroj:

$$\text{CeoBroj} ::= ['+' | '-'] \text{Cifra}\{\text{Cifra}\}$$

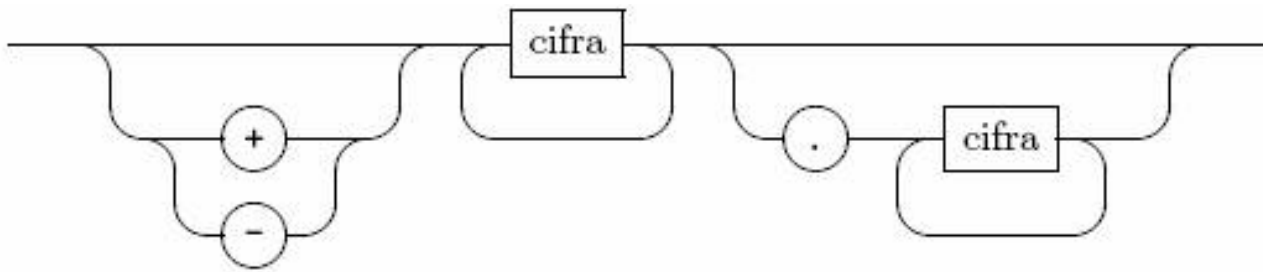
CeoBroj



- Sintaksni dijagram za RealanBroj: (DOMAĆI)

$$\text{RealanBroj} ::= ['+' | '-'] \text{Cifra}\{\text{Cifra}\} ['.' \text{Cifra}\{\text{Cifra}\}]$$

RealanBroj



- Sintaksni dijagram za Identifikator:

Identifikator ::= Slovo { Slovo | Cifra }

Identifikator

