Programski prevodioci

04 Semantička analiza

Fakultet tehničkih nauka, Novi Sad
22-23/Z
Dunja Vrbaški

Rezultat parsiranja

Šta je ulaz za narednu fazu, semantičku analizu?

- *(ništa)* Semantička analiza, pa i generisanje koda, se može izvršiti u okviru parsera

- *(struktura)* Realizuje se posebna međureprezentacija, apstraktno sintaksno stablo, čijim obilaskom se omogućava semantička analiza i generisanje koda

Semantička analiza

Semantika – značenje
Da li se poštuju pravila koja daju značenje programu?

- Da li je promenljiva definisana pre korišćenja?
- Da li postoji main()?
- Da li je identifikator vidljiv?
- Da li su identifikatori jednoznačni?
- Da li su tipovi u izrazima odgovarajući?
- Da li su tipovi argumenata i parametara odgovarajući?
- ...

Semantička pravila ćemo definisati neformalno, opisima i kroz primere.
Pravila više nisu kontekstno slobodna (kontekstno nezavisna).
Postoje formalizacije (operaciona, aksiomska semantika, denotaciona semantika)

Tokom parsiranja: preuzimaju se informacije, računaju se vrednosti i izvršavaju se određene akcije potrebne za naredne faze.

Često se koriste izmenjene, proširene gramatike koje, pored sintaksne strukture, sadrže informacije potrebne za semantičku analizu i kasnije faze.

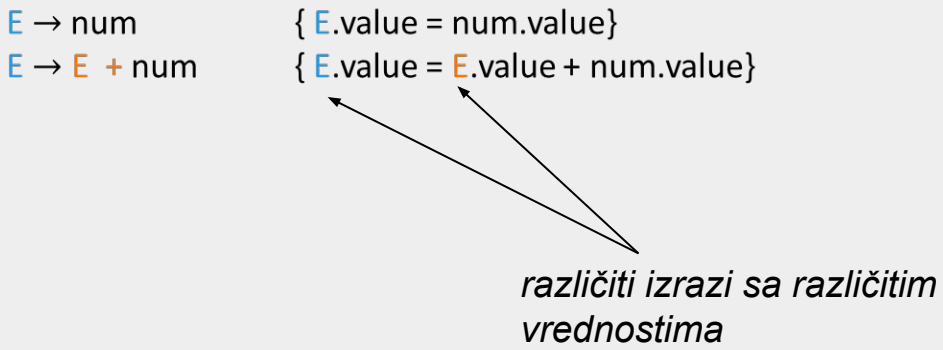
Produkcijama se dodeljuju semantička pravila.

Na primer:

- evidentiraju se ili izračunavaju neke vrednosti (x = 5, x = 2 + 3)
- dodaju se informacije o tipovima (int x)
- side effects (print, update global)
- ...

Token num
Skener postavlja vrednost prilikom formiranja tokena.
Leksema “123” postaje token num koji ima dodatnu (semantičku) vrednost 123

Pojam E → E + num
Parser postavlja vrednost pojma E (sa leve strane) na osnovu vrednosti tokena num i pojma E (sa desne strane; vrednost već postoji)



Neki tokeni i pojmovi ne moraju imati vrednosti.

- Atributivna (atributska) gramatika – dodeljuju se atributi elementima

- Čvor u stablu parsiranja ima atribut nastao na osnovu zadatih pravila (anotirano stablo)

- Atributi
  - sintetizovani – dobijaju se na osnovu atributa elemenata iz pravila (child nodes)
    - bottom-up parsiranje → vrednost se lako izračunava jer na steku možemo imati sve informacije
  - nasleđeni – zavise i od atributa parent čvora (npr. kontekst)
    - tek nakon formiranja stabla, prilagođavanje/izmene gramatike, prebacivanje u sintetizovane

- S-atributska gramatika je ona koja ima samo sintentizovane atribute

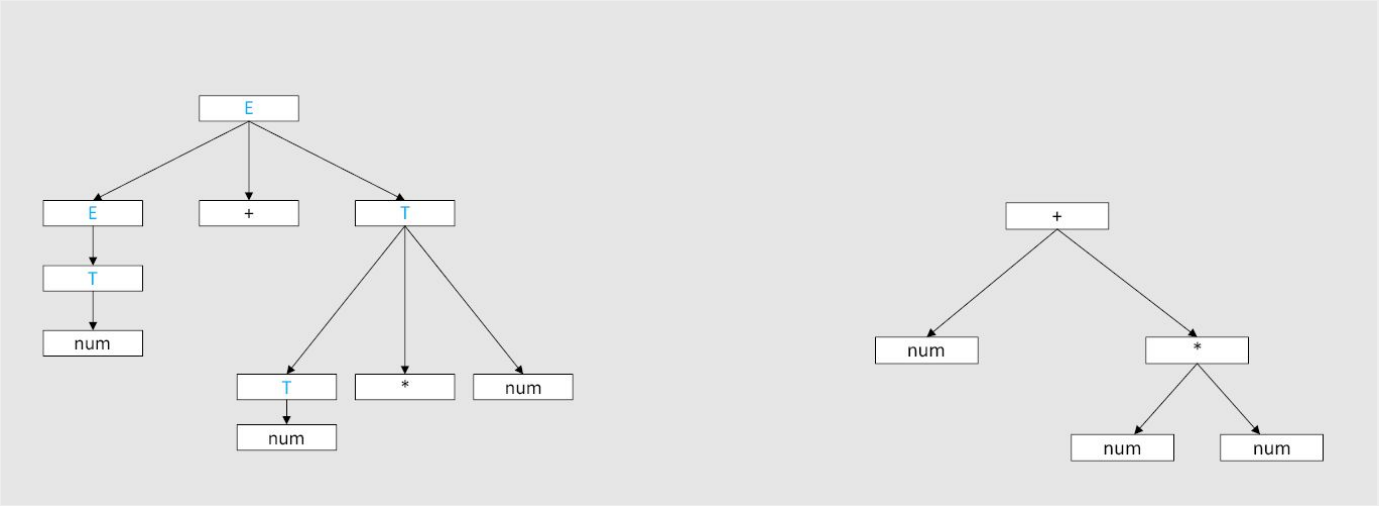
Apstraktno sintaksno stablo

Informacije se koriste za izgradnju apstraktnog sintaksnog stabla koji se dalje koristi u narednim fazama.

Međureprezentacija.

Stablo parsiranja – stablo koji sadrži sve pojmove i tokene i odgovara izvođenju

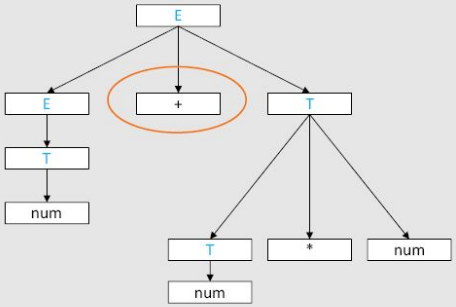
Apstraktno sintaksno stablo (AST) - stablo koje sadrži samo značajne delove sintaksnog stabla koji su nam potrebni za nastavak, za dalje faze.
(šta je, zapravo, programer hteo)

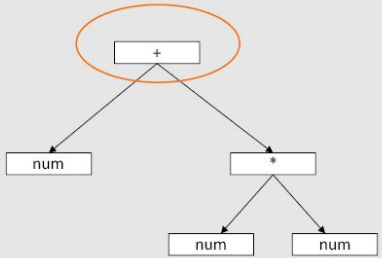


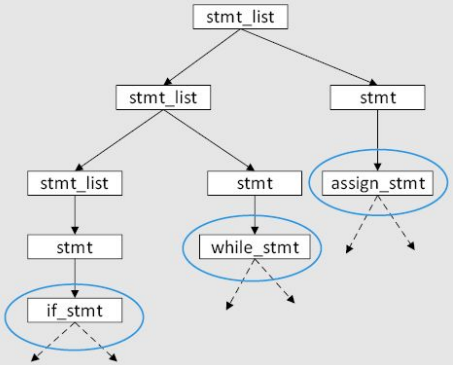
konkretno sintaksno stablo
parse tree

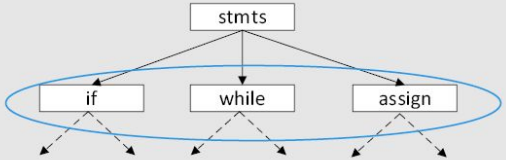
apstraktno sintaksno stablo
abstract syntax tree

Jednom kad znamo da je parsirani niz tokena validan većina informacija je nepotrebna.
Stablo parsiranja (parse tree) se tranformiše u AST (syntax tree).



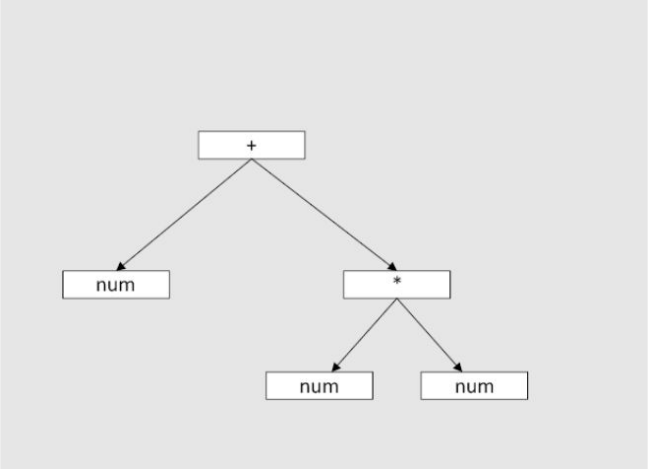






Operatori prelaze u unutrašnje čvorove (umesto listova)

Liste se “uravnavaju” (flatten)



Zašto AST?

Obilascima možemo, na primer, proveravati semantička pravila.

Na primer, usklađenost tipova.

Čvor +
    Čvor num (int)
    Čvor *
        Čvor num (int)
        Čvor num (int)

OK