Jednoprolazni asembler

Student: Mentor:

Tijana Simić 200/17 Saša Stojanović

# Uvod

Ovaj asembler je domaci zadatak iz predmeta Sistemski Softver. U direktorijumu src se nalazi izvorni kod resenja, kao i izvrsna verzija programa. Takodje, postoje i test fajlovi koji sluze za proveravanje ispravnosti programa.

Program je pisan u jeziku C++ na operativnom sistemu Linux.

Za prevodjenje ovog programa se koristi prevodilac g++ i izvrsni fajl se pravi komandom:

g++ asembler.cpp instructionParser.cpp lexer.cpp main.cpp symbolData.cpp -o main

Nakon toga pokretanje je moguce komandom:

./main inputFile.txt outputFile.cpp

Potrebno je poslati ulazni i izlazni fajl kao argumente komandne linije.

# Parsiranje teksta

Kod se cita liniju po liniju iz ulaznog fajla i salje se lexeru na parsiranje kako bi se linija podelila na reci, tj. tokene, obrisali komentari, suvisne beline, itd. To je uradjeno tako sto se ide slovo po slovo i formiraju se reci. Nakon toga linija koda je spremna za dalju obradu.

# Asembler

## Simboli

Glavni deo programa se nalazi u fajlu asembler.cpp gde se obradjuju konkretne reci. U zavisnosti od toga da li je rec direktiva ili instrukcija, sledi odgovarajuca obrada. Kada je rec simbol, on se odmah dodaje u tabelu simbola. Pored simbola, u tabeli simbola se nalaze i sekcije i njihova vrednost je nula, a oznacene su kao lokalne. Pri nailasku na neki simbol, ako nije u tabeli simbola, a naisli smo na koriscenje, dodajemo ga u tabelu simbola, dodajemo novi ulaz u listi obracanja unapred za taj simbol, i dodajemo ulaz u tabeli relokacija za tekucu sekciju. To radimo sve dok ne dodjemo do definicije simbola, kada je simbol definisan ne dodajemo ulaze u listi obracanja unapred. Za jedan simbol se u tabeli simbola pamti njegovo ime, redni broj simbola, redni broj sekcije u kojoj je definisan, vrednost, flag koji govori da li je lokalni ili globalni simbol, i ako je simbol sekcija, pamti se i velicina sekcije.

Ako je simbol lokalni, u izlaznu tabelu se upisuje njegova vrednost, u tabelu relokacija broj sekcije u kojoj je definisan, a ako je globalni ili eksterni, u izlaznu tabelu se upisuje vrednost 0, a u tabelu relokacija redni broj simbola.

Ako je simbol definisan .equ direktivom, razresavanje equ direktive ce se raditi na kraju, kada su svi simboli definisani.

Razresavanje simbola se radi na kraju, kada dodje direktiva .end, tada svi simboli imaju vrednost i onda se simboli definisani .equ direktivnom mogu razresiti i izracunati njihovi indeksi klasifikacije.

## Literali

Literali se odmah upisuju u izlazni fajl, pri cemu se vodi racuna o tome da li literal treba da bude velicine 1B ili 2B. U programu se radi sa oznacenim celim brojevima. Za te potrebe koriscen je tip \_\_int8\_t, ako je literal potrebno predstaviti na 1B, i \_\_int16\_t ako je literal potrebno predstaviti na 2B.

## Instrukcije

Instrukcije mogu imati 1 ili 2 operanda, a svaka instrukcija, osim sto poseduje instrukcijski descriptor, poseduje i deskriptore operanada. Asembler vidi 5 tipova adresiranja, neposredno, registarsko direktno, registarsko indirektno, registarsko indirektno sa pomerajem i memorijsko.

Kod PC relativnog adresiranja treba voditi racuna o pomeraju, ako nam je prvi operand tog adresiranja, a drugi operand ima samo descriptor, onda je pomeraj do sledece instrukcije 3B, a ako drugi operand ima I dodatna 2B za svoju vrednost, onda je pomeraj do sledece instrukcije 5B. Ako je drugi operand adresiran PC relativno, ili je rec o instrukciji skoka, pomeraj je 2B.

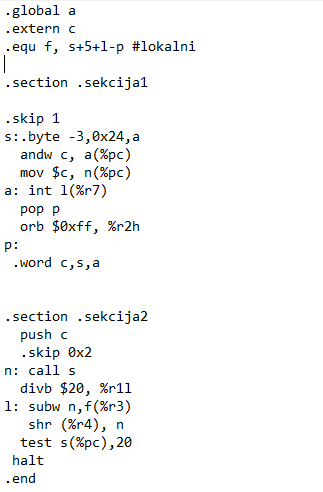
# Testovi

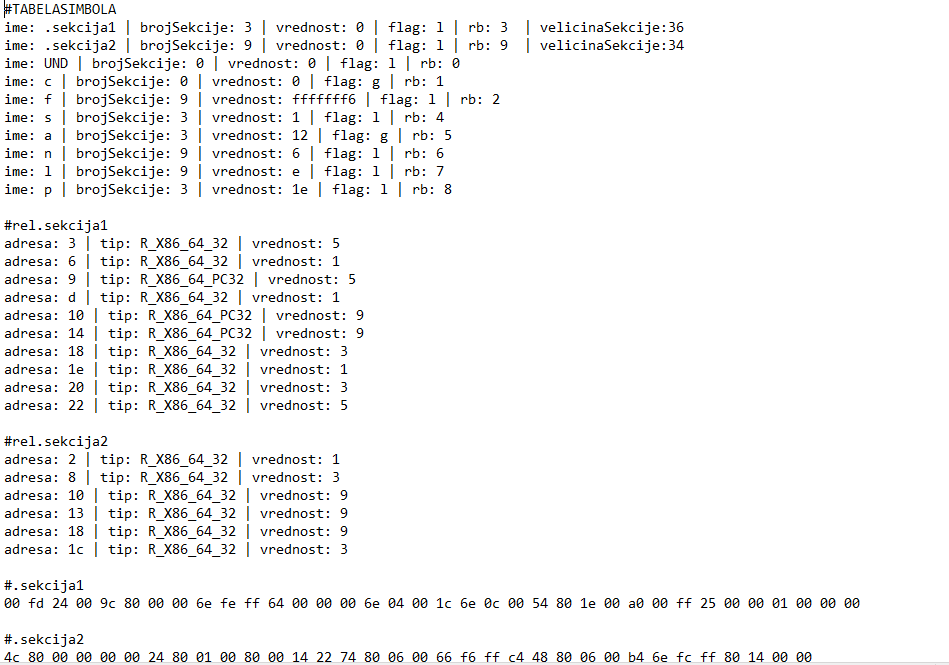
Za test 1 potrebno je proslediti kao argument komandne linije ulazni fajl inputFile.txt a za test 2 in1.txt. Prvi test je orijentisan na instrukcije i na poziv svih postojecih adresiranja, dok je drugi test orijentisan na skokove.

Definisano je nekoliko simbola, kako lokalnih, tako i globalnih i eksternih, kao i simboli dobijeni .equ direktivom. Posto su neki simboli definisani u jednoj, a neki u drugoj sekciji, doc ice do relokacija sto se moze videte u tabelama relokacija. Vrednosti simbola se mogu pratiti u tabeli simbola.

## Test 1

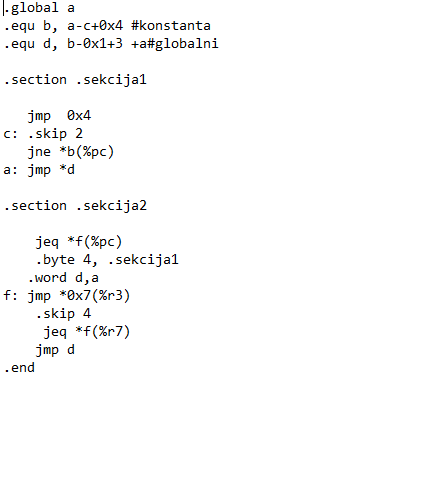
Ulaz:



Izlaz:

## Test 2

Ulaz:



Izlaz: