Elektrotehnički fakultet u Beogradu
Katedra za računarsku tehniku i informatiku

Predmet: Sistemski softver (13E113SS, 13S113SS)
Nastavnik: doc. dr Saša Stojanović
Školska godina: 2021/2022. (Zadatak važi počev od junskog roka 2022.)

Projekat za domaći rad
– Projektni zadatak –

Verzija dokumenta: 1.2

Važne napomene: Pre čitanja ovog teksta, **obavezno** pročitati opšta pravila predmeta i pravila vezana za izradu domaćih zadataka! Pročitati potom ovaj tekst **u celini i pažljivo**, pre započinjanja realizacije ili traženja pomoći. Ukoliko u zadatku nešto nije dovoljno precizno definisano ili su postavljeni kontradiktorni zahtevi, student treba da uvede razumne pretpostavke, da ih temeljno obrazloži i da nastavi da izgrađuje preostali deo svog rešenja na temeljima uvedenih pretpostavki. Zahtevi su namerno nedovoljno detaljni, jer se od studenata očekuje kreativnost i profesionalni pristup u rešavanju praktičnih problema!

Uvod

Cilj ovog projekta jeste realizacija alata u lancu prevođenja i emulatora za apstraktni računarski sistem. Opis apstraktnog računarskog sistema dat je u prilogu. Alati u lancu prevođenja koje treba realizovati obuhvataju asembler za navedeni apstraktni računarski sistem i linker nezavisan od ciljne arhitekture.

Rešenje projekta obuhvata izvorni kod kojim su implementirani asembler, linker i emulator. U daljem tekstu, rešenje projekta biće kratko nazivano samo *implementacija*. Izvorni asemblerski kod napisan za apstraktni računarski sistem, koji će asembler prevoditi, linker povezivati i emulator izvršavati, biće u nastavku referisan kao *korisnički program* kako bi se jasno napravila razlika u odnosu na *implementaciju*.

U cilju demonstracije i odbrane projekta, *implementacija* mora da se uspešno izvršava pod Linux operativnim sistemom kao konzolna aplikacija. Odbrana projekta pod Windows operativnim sistemom nije moguća.

Opšti zahtevi

Leksička analiza

Implementacija leksičke analize i parsiranja ulaznih tekstualnih datoteka ne predstavlja glavni zadatak projekta, ali jeste neophodna za uspešnu izradu rešenja, zbog čega je za te potrebe dozvoljeno je koristiti generatore leksera i parsera. Bez ulaska u detalje i namenu ovih alata, koji mogu biti korisni prilikom izrade rešenje, skreće se pažnja da su na virtuelnoj mašini instalirani alati *flex* i *bison* koji predstavljaju upravo generatore leksera i parsera respektivno. Pored generatora leksera i parsera dozvoljeno je i preporučeno koristiti standardne biblioteke kao što je STL (*Standard Template Library*). Takođe, dozvoljeno je koristiti i nestandardne biblioteke ukoliko one ne implementiraju stvari usko vezane sa srž projekta u koju spada generisanje mašinskog koda, pratećih metapodataka, formiranje relokacionih zapisa, relociranje i povezivanje predmetnih programa, emulacija itd.

Ocenjivanje projekta

Zahtevi u okviru zadataka navedenih u nastavku razvrstani su prema težini i obimu na tri nivoa: A (30 poena), B (35 poena) i C (40 poena). Podela zahteva po nivoima za svaki od zadataka definisana je neposredno nakon uvoda posmatranog zadatka. Prilikom odbrane projekta moguće je ostvariti:

- 30 poena, ako i samo ako su tačno urađeni svi zahtevi propisani za nivo A
- 35 poena, ako i samo ako su tačno urađeni svi zahtevi propisani za nivoe A i B
- 40 poena, ako i samo ako su tačno urađeni svi zahtevi propisani za nivoe A, B i C

Ukoliko student ne implementira ispravno sve zahteve propisane za nivo A za svaki od zadataka nije moguće uopšte pristupiti odbrani projekta. Angažovani na predmetu zadržavaju diskreciono pravo da odbranu projekta označe kao neuspešnu ukoliko ustanove da predato rešenje projekta sadrži grešku prilikom izvršavanja ili na bilo koji način odstupa od zahteva projektnog zadatka.

Zadatak 1: Asembler

Uvod

Cilj ovog zadatka jeste realizacija jednoprolaznog asemblera za procesor opisan u prilogu. Ulaz asemblera je tekstualna datoteka sa izvornim asemblerskim kodom napisanim u skladu sa sintaksom opisanom u nastavku. Izlaz asemblera je tekstualna datoteka koja predstavlja predmetni program (dozvoljeno je generisati kao izlaz, pored tekstualne datoteke, binarnu datoteku radi jednostavnijeg učitavanja izlaza asemblera u linker).

Format predmetnog programa bazirati na školskoj varijanti ELF formata čiji je referentni primer tekstualna datoteka kakva je korišćena na vežbama u delu gradiva koje se tiče konstrukcije asemblera. Dozvoljeno je praviti izmene u školskoj varijanti ELF formata (sekcije, tipovi zapisa o relokacijama, dodatna polja u postojećim tipovima zapisa, novi podaci o predmetnom programu i slično) ukoliko je to neophodno u skladu sa potrebama ciljne arhitekture. Prilikom razrešavanja svih nedefinisanih detalja za potrebe rešenja voditi se principima koje koristi GNU asembler.

Podela zahteva po nivoima

Podela zahteva po nivoima za ovaj zadatak tiče se asemblerskih direktiva i naredbi (ostale zahteve podrazumevano treba implementirati). Nivo u sklopu kojeg treba implementirati neku asemblersku direktivu naveden je u uglastim zagradama ispred tražene asemblerske direktive. Nivo u sklopu kojeg treba implementirati neku asemblersku naredbu naveden je unutar reda odgovarajuće tabele za posmatranu asemblersku naredbu.

Pokretanje iz terminala

Rezultat prevođenja *implementacije* ovog zadatka treba da ima `asembler` za naziv. Sve informacije potrebne za izvršavanje zadaju se kao argumenti komandne linije. Jednim pokretanjem `asembler` vrši asembliranje jedne ulazne datoteke. Naziv ulazne datoteke sa izvornim asemblerskim kodom zadaje se kao samostalni argument komandne linije. Način pokretanja `asembler` jeste sledeći:

```
asembler [opcije] <naziv_ulazne_datoteke>
```

Opcije komandne linije

Opis opcija komandne linije, zajedno sa opisom njihovih parametara, koje mogu biti zadate prilikom pokretanja `asembler` nalazi se u nastavku:

```
-o <naziv_izlazne_datoteke>
```

Opcija komandne linije `-o` postavlja svoj parametar `<naziv_izlazne_datoteke>` za naziv izlazne datoteke koja predstavlja rezultat asembliranja.

Primer pokretanja

Primer komande, kojom se inicira asembliranje izvornog koda u okviru datoteke *ulaz.s* sa ciljem dobijanja *izlaz.o* predmetnog programa, dat je u nastavku:

```
./asembler -o izlaz.o ulaz.s
```

Sintaksa izvornog asemblerskog koda

Sintaksa izvornog asemblerskog koda može se grubo podeliti na opšte detalje, asemblerske direktive i asemblerske naredbe. Opšti detalji definišu izgled jedne linije izvornog koda po pitanju zapisa labela, asemblerskih naredbi, asemblerskih direktivi, komentara itd. Asemblerska direktiva predstavlja operaciju koju asembler treba da izvrši u toku asembliranja. Asemblerska naredba predstavlja simbolički zapis mašinske instrukcije koji asembler treba da prevede u binarnu reprezentaciju.

Opšti detalji

Opšti detalji, predstavljeni u vidu (1) funkcionalnih zahteva koje asembler treba da ispuni i (2) sintaksnih pravila za koja asembler treba da proveri da li su ispoštovana, navedeni su redom po stavkama u nastavku:

- jedna linija izvornog koda sadrži najviše jednu asemblersku naredbu ili direktivu,
- zakomentarisan sadržaj se ignoriše u potpunosti prilikom asembliranja,
- jednolinijski komentar, koji se implicitno završava na kraju linije, započinje # karakterom,
- labela, čija se definicija završava dvotačkom, mora se naći na samom početku linije izvornog koda (opciono nakon proizvoljnog broja belih znakova) i
- labela može da stoji samostalno, bez prateće asemblerske naredbe ili asemblerske direktive u istoj liniji izvornog koda, što je ekvivalentno tome da stoji na samom početku prve naredne linije izvornog koda koja ima sadržaj.

Asemblerske direktive

`[nivo A] .global <lista_simbola>`

Izvozi simbole navedene u okviru liste parametara. Lista parametara može sadržati samo jedan simbol ili više njih razdvojenih zapetama.

`[nivo A] .extern <lista_simbola>`

Uvozi simbole navedene u okviru liste parametara. Lista parametara može sadržati samo jedan simbol ili više njih razdvojenih zapetama.

`[nivo A] .section <ime_sekcije>`

Započinje novu asemblersku sekciju, čime se prethodno započeta sekcija automatski završava, proizvoljnog imena navedenog kao parametar asemblerske direktive.

`[nivo A] .word <lista_simbola_ili_literal>`

Alocira prostor fiksne veličine po dva bajta za svaki inicijalizator (simbol ili literal) naveden u okviru liste parametara. Lista parametara može sadržati samo jedan inicijalizator ili više njih razdvojenih zapetama. Asemblerska direktiva alocirani prostor inicijalizuje vrednošću navedenih inicijalizatora.

`[nivo A] .skip <literal>`

Alocira prostor čija je veličina jednaka broju bajtova definisanom literalom navedenim kao parametar. Asemblerska direktiva alocirani prostor inicijalizuje nulama.

[nivo B] .ascii <string>

Alocira prostor fiksne veličine po jedan bajt za svaki karakter stringa (niska karaktera između znakova navoda). Asemblerska direktiva alocirani prostor inicijalizuje vrednostima koje odgovaraju navedenim karakterima prema ASCII tabeli.

[nivo C] .equ <novi_simbol>, <izraz>

Definiše novi simbol čija je vrednost jednaka navedenom izrazu.

[nivo A] .end

Završava proces asembliranja ulazne datoteke. Ostatak ulazne datoteke se odbacuje odnosno ne vrši se njegovo asembliranje.

Asemblerske naredbe

Mnemonik	Efekat	Flegovi	Nivo
halt	Zaustavlja izvršavanje instrukcija	-	A
int *regD*	push psw; pc <= mem16[(*regD* mod 8)*2];	-	A
iret	pop psw; pop pc;	psw	A
call *operand*	push pc; pc <= *operand*;	-	A
ret	pop pc;	-	A
jmp *operand*	pc <= *operand*;	-	A
jeq *operand*	if (equal) pc <= *operand*;	-	A
jne *operand*	if (not equal) pc <= *operand*;	-	A
jgt *operand*	if (signed greater) pc <= *operand*;	-	A
push *regD*	sp <= sp - 2; mem16[sp] <= *regD*;	-	A
pop *regD*	*regD* <= mem16[sp]; sp <= sp + 2;	-	A
xchg *regD*, *regS*	temp <= *regD*; *regD* <= *regS*; *regS* <= temp;	-	A
add *regD*, *regS*	*regD* <= *regD* + *regS*;	-	A
sub *regD*, *regS*	*regD* <= *regD* - *regS*;	-	A
mul *regD*, *regS*	*regD* <= *regD* * *regS*;	-	A
div *regD*, *regS*	*regD* <= *regD* / *regS*;	-	A
cmp *regD*, *regS*	temp <= *regD* - *regS*;	Z O C N	A
not *regD*	*regD* <= ~*regD*;	-	A
and *regD*, *regS*	*regD* <= *regD* & *regS*;	-	A
or *regD*, *regS*	*regD* <= *regD* | *regS*	-	A
xor *regD*, *regS*	*regD* <= *regD* ^ *regS*;	-	A
test *regD*, *regS*	temp <= *regD* & *regS*;	Z N	A
shl *regD*, *regS*	*regD* <= *regD* << *regS*;	Z C N	A
shr *regD*, *regS*	*regD* <= *regD* >> *regS*;	Z C N	A
ldr *regD*, *operand*	*regD* <= *operand*;	-	A
str *regD*, *operand*	*operand* <= *regD*;	-	A

Oznaka *regX* predstavlja oznaku nekog od programski dostupnih registara ciljne arhitekture. Programski dostupni registri su `r0`, `r1`, `r2`, `r3`, `r4`, `r5`, `r6`/`sp`, `r7`/`pc` i `psw`.

Oznaka *operand* obuhvata sve sintaksne notacije za navođenje operanada za različite načine adresiranja. Sintaksne notacije se razlikuju zavisno od toga da li se radi o asemblerskim naredbama za rad sa podacima ili asemblerskim naredbama skoka.

Asemblerske naredbe za rad sa podacima podržavaju različite sintaksne notacije za `operand`, opisane u nastavku, kojima se definiše vrednost podatka:

- `$<literal>` - vrednost `<literal>`
- `$<simbol>` - vrednost `<simbol>`
- `<literal>` - vrednost iz memorije na adresi `<literal>`
- `<simbol>` - vrednost iz memorije na adresi `<simbol>` apsolutnim adresiranjem
- `%<simbol>` - vrednost iz memorije na adresi `<simbol>` PC relativnim adresiranjem
- `<reg>` - vrednost u registru `<reg>`
- `[<reg>]` - vrednost iz memorije na adresi `<reg>`
- `[<reg> + <literal>]` - vrednost iz memorije na adresi `<reg> + <literal>`
- `[<reg> + <simbol>]` - vrednost iz memorije na adresi `<reg> + <simbol>`

Asemblerske naredbe skoka i poziva potprograma podržavaju različite sintaksne notacije za `operand`, opisane u nastavku, kojima se definiše vrednost odredišne adrese skoka:

- `<literal>` - vrednost `<literal>`
- `<simbol>` - vrednost `<simbol>` apsolutnim adresiranjem
- `%<simbol>` - vrednost `<simbol>` PC relativnim adresiranjem
- `*<literal>` - vrednost iz memorije na adresi `<literal>`
- `*<simbol>` - vrednost iz memorije na adresi `<simbol>`
- `*<reg>` - vrednost u registru `<reg>`
- `*[<reg>]` - vrednost iz memorije na adresi `<reg>`
- `*[<reg> + <literal>]` - vrednost iz memorije na adresi `<reg> + <literal>`
- `*[<reg> + <simbol>]` - vrednost iz memorije na adresi `<reg> + <simbol>`

Zadatak 2: Linker

Uvod

Cilj ovog zadatka jeste realizacija linkera nezavisnog od ciljne arhitekture koji na osnovu metapodataka (tabela simbola, relokacioni zapisi itd.) vrši povezivanje jednog ili više predmetnih programa generisanih od strane asemblera iz prvog zadatka.

Ulaz linkera je izlaz asemblera pri čemu je moguće zadati veći broj predmetnih programa koje je potrebno povezati. Linker podrazumevano smešta sekcije, počevši od nulte adrese, jednu odmah iza druge onim redom kako su definisane unutar predmetnog programa. Veći broj predmetnih programa na svom ulazu linker obrađuje u redosledu njihovog navođenja preko komandne linije. Prilikom smeštanja sekcije istog imena kao prethodno smeštena sekcija dolazi do njenog umetanja počev od mesta gde se istoimena sekcija ranije završila, stvarajući time agregaciju istoimenih sekcija, pri čemu nema preklapanja sa sekcijama sledbenicama što se postiže njihovim guranjem ka višim adresama.

Izlaz linkera je tekstualna datoteka sa sadržajem u skladu sa opisom u nastavku (dozvoljeno je generisati kao izlaz, pored tekstualne datoteke, binarnu datoteku radi jednostavnijeg učitavanja izlaza linkera u emulator).

Podela zahteva po nivoima

Podela zahteva po nivoima za ovaj zadatak tiče se opcija komandne linije (ostale zahteve podrazumevano treba implementirati). Nivo u sklopu kojeg treba implementirati neku opciju komandne linije naveden je u uglastim zagradama ispred tražene opcije komandne linije.

Pokretanje iz terminala

Rezultat prevođenja *implementacije* ovog zadatka treba da ima `linker` za naziv. Sve informacije potrebne za izvršavanje zadaju se kao argumenti komandne linije. Jednim pokretanjem `linker` vrši povezivanje jedne ili više ulaznih datoteka. Nazivi ulaznih datoteka, koje predstavljaju predmetne programe, zadaju se kao samostalni argumenti komandne linije. Način pokretanja `linker` jeste sledeći:

```
linker [opcije] <naziv_ulazne_datoteke>...
```

Opcije komandne linije

Opis opcija komandne linije, zajedno sa opisom njihovih parametara, koje mogu biti zadate prilikom pokretanja `linker` u proizvoljnom redosledu nalazi se u nastavku:

[nivo A] `-o <naziv_izlazne_datoteke>`

Opcija komandne linije `-o` postavlja svoj parametar `<naziv_izlazne_datoteke>` za naziv izlazne datoteke koja predstavlja rezultat povezivanja.

[nivo B] `-place=<ime_sekcije>@<adresa>`

Opcija komandne linije `-place` eksplicitno definiše adresu počev od koje se smešta sekcija zadatog imena pri čemu su adresa i ime sekcije određeni `<ime_sekcije>@<adresa>` parametrom. Ovu opciju moguće je navoditi više puta za različita imena sekcija kako bi se definisala adresa za veći broj sekcija iz ulaznih datoteka. Sve sekcije za koje ova opcija nije

navedena smeštaju se na podrazumevani način, opisan u sklopu uvoda ovog zadatka, počev odmah iza kraja sekcije koja je smeštena na najvišu adresu.

[nivo A] `-hex`

Opcija komandne linije `-hex` predstavlja smernicu linkeru da kao rezultat povezivanja generiše zapis, na osnovu kojeg se može izvršiti inicijalizacija memorije, u vidu skupa parova (*adresa, sadržaj*). Sadržaj predstavlja mašinski kod koji treba da se nađe na zadatoj adresi. Parovi se generišu samo za one adrese na koje treba smestiti sadržaj sa definisanom početnom vrednošću. Format zapisa, na primeru u kojem je početna vrednost sadržaja jednaka njegovoj adresi, prikazan je u nastavku:

```
0000: 00 01 02 03 04 05 06 07
0008: 08 09 0A 0B 0C 0D 0E 0F
0010: 10 11 12 13 14 15 16 17
```

Povezivanje je moguće samo u slučaju da ne postoje (1) višestruke definicije simbola, (2) nerazrešeni simboli i (3) preklapanja između sekcija iz ulaznih predmetnih programa kada se uzmu u obzir `-place` opcije komandne linije. Ukoliko za zadate ulazne datoteke linkera nije ispunjen neki od prethodnih uslova linker mora da prijavi grešku uz odgovarajuću poruku. Nazivi simbola ili sekcija koji su uzrok greške treba da budu deo poruke o grešci.

Prilikom pokretanja linkera navođenje tačno jedne od `-relocateable` i `-hex` opcija komandne linije je obavezno. Linker ne treba da generiše nikakav izlaz ako nije navedena tačno jedna od dve prethodno navedene opcije komandne linije.

[nivo C] `-relocateable`

Opcija komandne linije `-relocateable` predstavlja smernicu linkeru da kao rezultat povezivanja generiše predmetni program, istog formata kao i izlaz asemblera, u kojem se sve sekcije smeštaju takođe od nulte adrese (potpuno se ignorišu potencijalno navedene `-place` opcije komandne linije). Predmetni program dobijen na ovakav način može kasnije biti naveden kao ulaz linkera.

Povezivanje je moguće samo u slučaju da nema višestrukih definicija simbola. Ukoliko za zadate ulazne datoteke linkera postoji višestruka definicija simbola linker mora da prijavi grešku uz odgovarajuću poruku. Naziv višestruko definisanog simbola treba da bude deo poruke o grešci.

Prilikom pokretanja linkera navođenje tačno jedne od `-relocateable` i `-hex` opcija komandne linije je obavezno. Linker ne treba da generiše nikakav izlaz ako nije navedena tačno jedna od dve prethodno navedene opcije komandne linije.

Primer pokretanja

Primer komande, kojom se pokreće povezivanje predmetnih programa *ulaz1.o* i *ulaz2.o* pri čemu se (1) definišu adrese na koje se smeštaju odgovarajuće sekcije i (2) zahteva generisanje zapisa za inicijalizaciju memorije, dat je u nastavku:

```
./linker -hex
    -place=iv_table@0x0000 -place=text@0x4000
    -o mem_content.hex
    ulaz1.o ulaz2.o
```

Zadatak 3: Emulator

Uvod

Cilj ovog zadatka jeste realizacija interpretativnog emulatora za računarski sistem opisan u prilogu. Ulaz emulatora jeste datoteka za inicijalizaciju memorije dobijena kao izlaz linkera uz navedenu `-hex` opciju komandne linije. Emulacija je moguća samo ukoliko je ulaznu datoteku moguće uspešno učitati u memorijski adresni prostor emuliranog računarskog sistema. Nakon pokretanja emulatora jedini ispis u konzoli jeste ispis direktno iz *korisničkog programa*, dok *implementacija* ne ispisuje ništa samostalno do završetka emulacije. Emulacija se završava u onom trenutku kada emulirani procesor izvrši `halt` instrukciju *korisničkog programa*. Nakon završetka emulacije *implementacija* ispisuje stanje emuliranog procesora u sledećem formatu:

```
-----------------------------------------------
Emulated processor executed halt instruction
Emulated processor state: psw=0b0000000000000000
r0=0x0000    r1=0x0000    r2=0x0000    r3=0x0000
r4=0x0000    r5=0x0000    r6=0x0000    r7=0x0000
```

Podela zahteva po nivoima

Podela zahteva po nivoima za ovaj zadatak definisana je u nastavku. Za nivo A potrebno je emulirati celokupan posmatrani računarski sistem osim periferija terminal i tajmer. Za nivo B, pored svega definisanog nivoom A, potrebno je emulirati i periferiju terminal. Za nivo C, pored svega definisanog nivoom B, potrebno je emulirati i periferiju tajmer.

Pokretanje iz terminala

Rezultat prevođenja *implementacije* ovog zadatka treba da ima `emulator` za naziv. Sve informacije potrebne za izvršavanje zadaju se kao argumenti komandne linije. Jednim pokretanjem `emulator` emulira jedno izvršavanje programa iz ulazne datoteke. Naziv ulazne datoteke, koja predstavlja izlaz linkera uz navedenu `-hex` opciju komandne linije, zadaje se kao samostalni argument komandne linije. Način pokretanja `emulator` jeste sledeći:

```
emulator <naziv_ulazne_datoteke>
```

Primer pokretanja

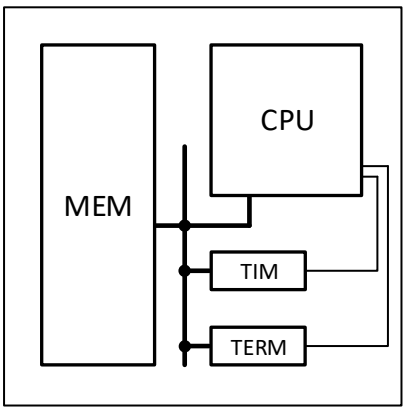
Primer komande, kojom se pokreće emulacija na osnovu zapisa za inicijalizaciju memorije u ulaznoj datoteci *mem_content.hex*, dat je u nastavku:

```
./emulator mem_content.hex
```

Prilog: Opis računarskog sistema

Uvod

Apstraktni računarski sistem se sastoji od procesora, operativne memorije, tajmera i terminala. Sve komponente računarskog sistema su međusobno povezane preko sistemske magistrale. Tajmer i terminal su povezani pored sistemske magistrale i direktno sa procesorom preko linija za slanje zahteva za prekid. Uprošćen šematski prikaz posmatranog apstraktnog računarskog sistema prikazan je na sledećoj slici:



Opis procesora

U nastavku je opisan deo 16-bitnog dvoadresnog procesora sa Von-Neuman arhitekturom. Adresibilna jedinica je jedan bajt, a raspored bajtova u reči je little-endian. Veličina memorijskog adresnog prostora je 2^{16}B.

Memorijski mapirani registri

Memorijski mapirani registri jesu registri kojima se pristupa instrukcijama za pristup memorijskom adresnom prostoru. Počev od adrese `0xFF00` memorijskog adresnog prostora nalazi se prostor veličine 256 bajtova rezervisan za memorijski mapirane registre. Memorijski mapirani registri koriste se za rad sa periferijama u računarskom sistemu.

Mehanizam prekida

Tabela prekidnih rutina odnosno IVT (interrupt vector table) nalazi se počev od adrese `0x0000` memorijskog adresnog prostora i ima osam ulaza. Svaki ulaz zauzima dva bajta i sadrži adresu odgovarajuće prekidne rutine. Ulazi u IVT odgovaraju sledećim prekidnim rutinama:

- ulaz 0 sadrži adresu prekidne rutine koja se izvršava prilikom pokretanja odnosno resetovanja čitavog procesora (ne izvodi se kompletna sekvenca obrade prekida već se samo vrši skok na adresu koja se nalazi u okviru datog ulaza),
- ulaz 1 sadrži adresu prekidne rutine koja se izvršava ukoliko se pokuša izvršavanje nekorektne instrukcije (nepostojeći operacioni kod, neispravan način adresiranja itd.),
- ulaz 2 sadrži adresu prekidne rutine koja se izvršava kada stigne zahtev za prekid od tajmera (opis principa rada tajmera i način njegove konfiguracije dat je u zasebnom poglavlju),

- ulaz 3 sadrži adresu prekidne rutine koja se izvršava kada stigne zahtev za prekid od terminala (opis principa rada terminala dat je u zasebnom poglavlju) i
- ostali ulazi su slobodni za korišćenje od strane programera.

Svaka mašinska instrukcija procesora izvršava se atomično. Zahtevi za prekid se opslužuju tek nakon što se trenutna mašinska instrukcija atomično izvrši do kraja. Procesor prilikom prihvatanja zahteva za prekid i ulaska u prekidnu rutinu samo postavlja programsku statusnu reč i povratnu adresu na stek.

Procesorski registri

Procesor poseduje osam opštenamenskih 16-bitnih registara označenih sa `r<num>` gde `<num>` može imati vrednosti od nula do sedam. Registar `r7` se koristi kao `pc` registar. Vrednost registra `r7` sadrži adresu instrukcije koja naredna treba da se izvrši. Registar `r6` se koristi kao `sp` registar. Vrednost registra `r6` sadrži adresu zauzete lokacije na vrhu steka (stek raste ka nižim adresama).

Pored opštenamenskih registara postoji `psw` registar (statusna reč procesora). Statusna reč procesora se sastoji od flegova koji pružaju mogućnost konfiguracije mehanizma prekida i predstavljaju status izvršavanog programa. Izgled statusne reči procesora jeste:

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
I	Tl	Tr										N	C	O	Z

Značenje flegova u `psw` registru:

- `Z` (Zero) - rezultat prethodne operacije je nula,
- `O` (Overflow) - prekoračenje,
- `C` (Carry) - prenos,
- `N` (Negative) - rezultat prethodne operacije je negativan,
- `Tr` (Timer) - maskiranje prekida od tajmera (0 - omogućen, 1 - maskiran),
- `Tl` (Terminal) - maskiranje prekida od terminala (0 - omogućen, 1 - maskiran) i
- `I` (Interrupt) - globalno maskiranje spoljašnjih prekida (0 - omogućeni, 1 - maskirani).

Format procesorskih instrukcija

Instrukcije mogu biti veličine od jedan do pet bajtova. Instrukcija u najopštijem slučaju ima sledeći format:

I	II	III	IV	V
InstrDescr	RegsDescr	AddrMode	DataHigh	DataLow

Prvi bajt `InstrDescr` predstavlja opis instrukcije u kom se nalaze operacioni kod i modifikator instrukcije. Operacioni kod definiše o kojoj instrukciji je reč. Modifikator dodatno govori šta tačno instrukcija treba da uradi. Format prvog bajta instrukcije je sledeći:

7	6	5	4	3	2	1	0
OC_3	OC_2	OC_1	OC_0	MOD_3	MOD_2	MOD_1	MOD_0

Značenje bitova `InstrDescr` bajta instrukcije:

- $OC_3OC_2OC_1OC_0$ - operacioni kod instrukcije i
- $MOD_3MOD_2MOD_1MOD_0$ - modifikator instrukcije.

Drugi bajt `RegsDescr` definiše registre koje instrukcija koristi preko njihovih indeksa. Opštenamenskim registrima dodeljeni su indeksi koji odgovaraju njihovim imenima. Indeks registra `r0` je nula, indeks registra `r1` je jedan itd. Indeks `psw` registra je osam. Format drugog bajta instrukcije je sledeći:

7	6	5	4	3	2	1	0
RD_3	RD_2	RD_1	RD_0	RS_3	RS_2	RS_1	RS_0

Značenje bitova `RegsDescr` bajta instrukcije:

- $RD_3RD_2RD_1RD_0$ - indeks prvog registra (uglavnom *destination*) koji instrukcija koristi i
- $RS_3RS_2RS_1RS_0$ - indeks drugog registra (uglavnom *source*) koji instrukcija koristi.

Treći bajt `AddrMode` definiše način adresiranja operanda za instrukcije koje ne barataju samo registrima. Format trećeg bajta instrukcije je sledeći:

7	6	5	4	3	2	1	0
Up_3	Up_2	Up_1	Up_0	AM_3	AM_2	AM_1	AM_0

Značenje bitova `AddrMode` bajta instrukcije:

- $Up_3Up_2Up_1Up_0$ - način ažuriranja kod registarsko indirektnih adresiranja i
- $AM_3AM_2AM_1AM_0$ - način adresiranja.

Načini adresiranja

Najveći deo instrukcijskog seta procesora koristi samo podatke iz registara odnosno isključivo i podrazumevano koristi registarsko direktno adresiranje bez podrške za neki drugi načina adresiranja. Mali podskup instrukcijskog seta ima podršku za druge načine adresiranja opisane u nastavku.

Opis načina adresiranja koristi nekoliko pojmova kako bi definisao svaki od načina adresiranja. Pojam *operand* kod grupe instrukcija za rad sa podacima predstavlja sam podatak dok kod grupe instrukcija skoka predstavlja odredišnu adresu skoka. Pojam *odabrani registar* podrazumeva registar određen bitovima $RS_3RS_2RS_1RS_0$ unutar `RegsDescr` bajta instrukcije. Pojam *payload instrukcije* odnosi se na `DataHigh` i `DataLow` bajtove instrukcije. Zavisno od vrednosti bitova $AM_3AM_2AM_1AM_0$ procesor podržava sledeće načine adresiranja:

- `0b0000` - neposredno; *operand* je jednak *payload instrukcije*,
- `0b0001` - registarsko direktno; *operand* je jednak vrednosti u *odabranom registru*,
- `0b0101` - registarsko direktno sa 16-bitnim označenim sabirkom; *operand* je jednak zbiru vrednosti u *odabranom registru* i *payload instrukcije* koji predstavlja označeni sabirak.
- `0b0010` - registarsko indirektno; *operand* je u memoriji na adresi ukazanoj vrednošću u *odabranom registru* (vrednost *odabranog registra* se ažurira prilikom izvršavanja instrukcije u skladu sa vrednošću $Up_3Up_2Up_1Up_0$ bitova),
- `0b0011` - registarsko indirektno sa 16-bitnim označenim pomerajem; *operand* je u memoriji na adresi ukazanoj zbirom vrednosti u *odabranom registru* i *payload instrukcije* koji predstavlja označeni pomeraj (vrednost *odabranog registra* se ažurira prilikom izvršavanja instrukcije u skladu sa vrednošću $Up_3Up_2Up_1Up_0$ bitova),
- `0b0100` - memorijsko; *operand* je u memoriji na adresi ukazanoj *payload instrukcije*.

Vrednost $Up_3Up_2Up_1Up_0$ bitova definiše na koji način će biti ažurirana vrednost *odabranog registra* prilikom izvršavanja instrukcije koja koristi neki od registarsko indirektnih načina adresiranja. U zavisnosti od vrednosti $Up_3Up_2Up_1Up_0$ bitova vrednost *odabranog registra* se ažurira na sledeći način:

- `0b0000` - nema ažuriranja,
- `0b0001` - umanjuje se za dva pre formiranje adrese *operanda*,
- `0b0010` - uvećava se za dva pre formiranja adrese *operanda*
- `0b0011` - umanjuje se za dva nakon formiranja adrese *operanda* i
- `0b0100` - uvećava se za dva nakon formiranja adrese *operanda*.

Pregled procesorskih instrukcija

Instrukcija za zaustavljanje procesora

InstrDescr	
7654	3210
0000	0000

Zaustavlja procesor kao i dalje izvršavanje narednih instrukcija. Veličina instrukcije je jedan bajt.

Instrukcija softverskog prekida

InstrDescr		RegsDescr	
7654	3210	7654	3210
0001	0000	DDDD	1111

Generiše softverski zahtev za prekid. Broj ulaza u IVT za koji se generiše zahtev za prekid nalazi se u odredišnom registru. Veličina instrukcije je dva bajta.

```
push psw; pc<=mem16[(reg[DDDD] mod 8)*2];
```

Instrukcija povratka iz prekidne rutine

InstrDescr	
7654	3210
0010	0000

Izlazi iz prekidne rutine povratkom na adresu sa steka i restaurira `psw` registar vrednošću sa steka. Veličina instrukcije je jedan bajt.

```
pop psw; pop pc;
```

Instrukcija poziva potprograma

InstrDescr		RegsDescr		AddrMode	
7654	3210	7654	3210	7654	3210
0011	0000	1111	SSSS	UUUU	AAAA

Poziva potprogram skokom na adresu definisanu *operandom* pre čega sačuva povratnu adresu na steku. Veličina instrukcije je tri ili pet bajtova (zavisno od načina adresiranja).

```
push pc; pc<=operand;
```

Instrukcija povratka iz potprograma

InstrDescr	
7654	3210
0100	0000

Izlazi iz potprograma povratkom na adresu sa steka. Veličina instrukcije je jedan bajt.

```
pop pc;
```

Instrukcija skoka

InstrDescr		RegsDescr		AddrMode	
7654	3210	7654	3210	7654	3210
0101	MMMM	1111	SSSS	UUUU	AAAA

Skače bezuslovno ili uslovno, u skladu sa modifikatorom instrukcije, na adresu definisanu *operandom*. Veličina instrukcije je tri ili pet bajtova (zavisno od načina adresiranja). Zavisno od modifikatora instrukcija vrši sledeći tip skoka:

```
MMMM==0b0000: /* jmp */ pc<=operand;
MMMM==0b0001: /* jeq */ if (equal) pc<=operand;
MMMM==0b0010: /* jne */ if (not equal) pc<=operand;
MMMM==0b0011: /* jgt */ if (signed greater) pc<=operand;
```

Instrukcija atomične zamene vrednosti

InstrDescr		RegsDescr	
7654	3210	7654	3210
0110	0000	DDDD	SSSS

Zamenjuje vrednost odredišnog i izvorišnog registra atomično bez mogućnosti da zamena bude prekinuta usled asinhronog zahteva za prekid. Veličina instrukcije je dva bajta.

```
temp<=reg[DDDD]; reg[DDDD]<=reg[SSSS]; reg[SSSS]<=temp;
```

Instrukcija aritmetičkih operacija

InstrDescr		RegsDescr	
7654	3210	7654	3210
0111	MMMM	DDDD	SSSS

Vrši odgovarajuću aritmetičku operaciju, u skladu sa modifikatorom instrukcije, nad vrednostima u odredišnom i izvorišnom registru. Veličina instrukcije je dva bajta. Zavisno od modifikatora instrukcija obavlja sledeću operaciju:

```
MMMM==0b0000: /* add */ reg[DDDD]<=(reg[DDDD] + reg[SSSS]);
MMMM==0b0001: /* sub */ reg[DDDD]<=(reg[DDDD] - reg[SSSS]);
MMMM==0b0010: /* mul */ reg[DDDD]<=(reg[DDDD] * reg[SSSS]);
MMMM==0b0011: /* div */ reg[DDDD]<=(reg[DDDD] / reg[SSSS]);
MMMM==0b0100: /* cmp */ temp<=(reg[DDDD] - reg[SSSS]); updt psw;
```

Instrukcija logičkih operacija

InstrDescr		RegsDescr	
7654	3210	7654	3210
1000	MMMM	DDDD	SSSS

Vrši odgovarajuću logičku operaciju, u skladu sa modifikatorom instrukcije, nad vrednostima u odredišnom i izvorišnom registru. Veličina instrukcije je dva bajta. Zavisno od modifikatora instrukcija obavlja sledeću operaciju:

```
MMMM==0b0000: /*  not */ reg[DDDD]<=~reg[DDDD];
MMMM==0b0001: /*  and */ reg[DDDD]<=(reg[DDDD] & reg[SSSS]);
MMMM==0b0010: /*   or */ reg[DDDD]<=(reg[DDDD] | reg[SSSS]);
MMMM==0b0011: /*  xor */ reg[DDDD]<=(reg[DDDD] ^ reg[SSSS]);
MMMM==0b0100: /* test */ temp<=(reg[DDDD] & reg[SSSS]); updt psw;
```

Instrukcija pomeračkih operacija

InstrDescr		RegsDescr	
7654	3210	7654	3210
1001	MMMM	DDDD	SSSS

Vrši odgovarajuću pomeračku operaciju, u skladu sa modifikatorom instrukcije, nad vrednostima u odredišnom i izvorišnom registru. Veličina instrukcije je dva bajta. Zavisno od modifikatora instrukcija obavlja sledeću operaciju:

```
MMMM==0b0000: /* shl */ reg[DDDD]<=(reg[DDDD] << reg[SSSS]); updt psw;
MMMM==0b0001: /* shr */ reg[DDDD]<=(reg[DDDD] >> reg[SSSS]); updt psw;
```

Instrukcija učitavanja podatka

InstrDescr		RegsDescr		AddrMode	
7654	3210	7654	3210	7654	3210
1010	0000	DDDD	SSSS	UUUU	AAAA

Učitava podatak definisan *operandom* u odredišni registar. Veličina instrukcije je tri ili pet bajtova (zavisno od načina adresiranja).

```
reg[DDDD]<=operand;
```

Instrukcija smeštanja podatka

InstrDescr		RegsDescr		AddrMode	
7654	3210	7654	3210	7654	3210
1011	0000	DDDD	SSSS	UUUU	AAAA

Smešta vrednost iz odredišnog registra u podatak definisan *operandom*. Veličina instrukcije je tri ili pet bajtova (zavisno od načina adresiranja).

```
operand<=reg[DDDD];
```

Sve kombinacije instrukcija i operanada, za koje ne postoji razumno tumačenje, proglasiti greškom.

Opis terminala

Terminal predstavlja ulazno/izlaznu periferiju koja se sastoji od displeja (izlaz) i tastature (ulaz). Terminal poseduje dva programski dostupna registra `term_out` i `term_in` kojima se pristupa kroz memorijski adresni prostor (memorijski mapirani registri). Navedeni programski dostupni registri mapirani su u memorijski adresni prostor na sledeći način:

Registar	Opseg adresa
`term_out`	`[0xFF00-0xFF01]`
`term_in`	`[0xFF02-0xFF03]`

Registar `term_out` predstavlja registar izlaznih podataka. Terminal prati upise u ovaj registar i prilikom svakog upisa vrednosti u `term_out` registar ispisuje znak na displej. Ispisani znak određen je sadržajem ASCII tabele za vrednost koja je upisana u `term_out` registar.

Registar `term_in` predstavlja registar ulaznih podataka. Terminal vrši sledeće dve operacije svaki put kada se pritisne bilo koje dugme na tastaturi: (1) upisuje ASCII kod pritisnutog tastera u `term_in` registar kako bi čitanjem njegove vrednosti bilo moguće ustanoviti koje dugme je pritisnuto i (2) generiše zahtev za prekid. Dve prethodno opisane operacije terminal vrši čim se pritisne bilo koje dugme; pri čemu treba naglasiti da terminal nipošto ne čeka terminator unosa odnosno *enter* dugme. Terminal ne vrši *echo* pritisnutog dugmeta odnosno ne prikazuje ga na displeju. Za potrebe implementacije ove funkcionalnosti emulatora moguće je, na primer, koristiti `<termios.h>` zaglavlje. Ukoliko *korisnički program* ne pročita vrednost `term_in` registra na vreme, odnosno pre nego što korisnik pritisne naredno dugme, ASCII kod prethodno pritisnutog dugmeta biće bespovratno izgubljen. Ne treba vršiti baferisanje ASCII kodova u okviru *implementacije*. Smatrati da je procesor posmatranog računarskog sistema dovoljno brz da garantovano, ukoliko je *korisnički program* ispravno napisan, može da stigne da pročita vrednost `term_in` registra u prekidnoj rutini pre nego što ona bude pregažena.

Opis tajmera

Tajmer predstavlja periferiju koja periodično generiše zahtev za prekid. Tajmer poseduje jedan programski dostupan registar `tim_cfg` kojem se pristupa kroz memorijski adresni prostor (memorijski mapiran registar). Navedeni programski dostupan registar mapiran je u memorijski adresni prostor na sledeći način:

Registar	Opseg adresa
`tim_cfg`	`[0xFF10-0xFF11]`

Registar `tim_cfg` predstavlja konfiguracioni registar tajmera. Perioda generisanja zahteva za prekid od strane tajmera definisana je vrednošću `tim_cfg` registra na sledeći način: `0x0` -> 500ms, `0x1` -> 1000ms, `0x2` -> 1500ms, `0x3` -> 2000ms, `0x4` -> 5000ms, `0x5` -> 10s, `0x6` -> 30s i `0x7` -> 60s. Inicijalna vrednost `tim_cfg` registra, nakon pokretanja odnosno resetovanja emuliranog računarskog sistema, jeste `0x0000`.

Primer korisničkog programa

Program prikazan u nastavku predstavlja primer programa napisanog na asembleru opisanom u prvom zadatku. Izvorni asemblerski kod programa je razdvojen na dve tekstualne datoteke *interrupts.s* i *main.s*. Program obavlja sledeće radnje:

- definiše sadržaj IVT odnosno postavlja adrese prekidnih rutina,
- prekidna rutina terminala ispisuje na displeju znak koji odgovara pritisnutom dugmetu,
- prekidna rutina tajmera ispisuje na displeju znak 'T',
- konfiguriše tajmer tako da generiše zahtev za prekid na svaku sekundu i
- čeka da korisnik pritisne pet puta dugme pre nego što zaustavi procesor.

```
# file: interrupts.s
.section ivt
   .word isr_reset
   .skip 2 # isr_error
   .word isr_timer
   .word isr_terminal
   .skip 8
.extern my_start, my_counter
.section isr
.equ term_out, 0xFF00
.equ term_in, 0xFF02
.equ ascii_code, 84 # ascii('T')
# prekidna rutina za reset
isr_reset:
   jmp my_start
# prekidna rutina za tajmer
isr_timer:
   push r0
   ldr r0, $ascii_code
   str r0, term_out
   pop r0
   iret
# prekidna rutina za terminal
isr_terminal:
   push r0
   push r1
   ldr r0, term_in
   str r0, term_out
   ldr r0, %my_counter # pcrel
   ldr r1, $1
   add r0, r1
   str r0, my_counter # abs
   pop r1
   pop r0
   iret
.end
```

```
# file: main.s
.global my_start
.global my_counter
.section my_code
.equ tim_cfg, 0xFF10
.equ init_sp, 0xFEFE
my_start:
   ldr r6, $init_sp
   ldr r0, $0x1
   str r0, tim_cfg
wait:
   ldr r0, my_counter
   ldr r1, $5
   cmp r0, r1
   jne wait
   halt
.section my_data
my_counter:
   .word 0
.end
```

Komande za prevođenje, povezivanje i emuliranje ovog primera jesu:

```
./asembler -o interrupts.o interrupts.s
./asembler -o main.o main.s
./linker -hex -place=ivt@0x0000 -o program.hex interrupts.o main.o
./emulator program.hex
```

Predaja projekta

Potrebno je realizovati prethodno opisane alate prema datim zahtevima pod Linux operativnim sistemom na amd64 arhitekturi koristeći C/C++ programski jezik. Potrebni alati su opisani u okviru materijala za predavanja i vežbe. Preporuka je rešenje izrađivati u okviru VMware virtuelne mašine koju je moguće preuzeti sa eLearning platforme. Virtuelna mašina sadrži već instalirane sve neophodne alate a kao integrisano okruženje za razvoj programa moguće je koristiti VS Code koje treba povezati sa GNU toolchain setom.

Podešavanje okruženja potrebnog za uspešno prevođenje izvornog koda i pokretanje programa takođe spada u deo postavke projektnog zadatka. Odbrana rešenja projektnog zadatka vrši se isključivo pod prethodno opisanim okruženjem u okviru virtuelne mašine. Pristup internetu prilikom odbrane rešenja projektnog zadatka biće onemogućen što znači da treba koristiti samo alate dostupne na virtuelnoj mašini.

Pravila za predaju projekta

Projekat se predaje isključivo kao jedna .zip arhiva unutar koje se nalazi samo jedan direktorijum imena *resenje* čiji je sadržaj opisan u nastavku. Unutar direktorijuma *resenje* može se naći (1) opciono *makefile* ili neka druga skripta za prevođenje rešenja zadatka, (2) opciono poddirektorijum *misc* predviđen za dodatne stvari kao što su flex i bison ulazne datoteke (izlazne datoteke generisane ovim alatima ne treba nikako predavati jer će one biti generisane) i (3) obavezno sledeća tri poddirektorijuma: *tests*, *src* i *inc*. Sadržaj ovih poddirektorijuma treba da bude sledeći:

- *tests*: ulazne datoteke za prikaz funkcionalnosti rešenja zadatka,
- *src*: sve .c i .cpp datoteke sa izvornim kodom i
- *inc*: sve .h i .hpp datoteke sa izvornim kodom.

Opisani sadržaj .zip arhive ujedno treba da bude i jedini njen sadržaj. Nije dozvoljeno predavati izvršne datoteke, datoteke generisane od strane alata flex i bison niti bilo šta drugo što iznad nije eksplicitno navedeno. Nepoštovanje pravila za predaju projekta po pitanju sadržaja, strukture i naziva direktorijuma, povlači negativne poene ili zabranu izlaska na odbranu u datom ispitnom roku.

Prikaz primitivnog primera (rešenje svakako treba da sadrži mnogo veći broj datoteka sa izvornim kodom u odnosu na prikazano) sadržaja .zip arhive nalazi se u nastavku:

```
.
└── resenje
    ├── makefile
    ├── misc
    │   ├── flex
    │   └── bison
    ├── inc
    │   ├── assembler.hpp
    │   ├── linker.hpp
    │   └── emulator.hpp
    ├── src
    │   ├── assembler.cpp
    │   ├── linker.cpp
    │   └── emulator.cpp
    └── tests
        ├── test1.s
        ├── test2.s
        └── test3.s
```

Zapisnik revizija

Ovaj zapisnik sadrži spisak izmena i dopuna ovog dokumenta po verzijama.

Verzija 1.1

Strana	Izmena
19	Dodata inicijalizacija pokazivača na vrh steka.

Verzija 1.2

Strana	Izmena
4	Izmenjena pravila za ocenjivanje projekta
20	Neznatno izmenjen opis predaje projekta

Verzija 1.3

Strana	Izmena