LetterCounter - Tehnička dokumentacija

# Opis zadatka:

Klijent je zatražio Java framework koji broji samoglasnike i suglasnike u .txt i .xml datotekama. Specifičan uvjet je način na koji se oni broje u .xml datotekama – suglasnici se broje kroz cijelu datoteku, dok se samoglasnici broje samo u vrijednostima elemenata i atributa XML node-ova.

Zatim, korištenjem tog framework-a, klijent je zatražio izradu aplikacije koja iz komandne linije prima naziv datoteke te ispisuje broj samoglasnika u zadanoj datoteci.

U slučaju da je framework-u predana datoteka druge ekstenzije, ili u slučaju za datoteka ne postoji na disku korisnika, framework o tome treba obavijestiti korisnika. U slučaju neke druge greške, potrebno je ispisati detalje greške zbog lakšeg debugging-a.

# Upute za korištenje:

Rješenje zadatka izrađeno je kao Maven projekt. Kako bi ga se moglo koristiti, potrebno je (iz root direktorija projekta) pozvati sljedeću naredbu:

$ mvn clean package

Ovom naredbom izradit će se .jar datoteka u /target direktoriju, koju možemo pozvati iz terminala koristeći program java. Npr.:

$ cd target  
$ java -jar vowel-counter-0.0.1.jar ~/dummy.xml

# Opis rješenja:

Za potrebe rješavanja ovog problema, izrađen je LetterCounter framework, koji ispunjava navedene uvjete.

## LetterCounter

LetterCounter klasa u sebi sadrži sve metode korištene u izradi ovog rješenja. Također sadrži privatni konstruktor bez argumenata, koji služi tome da se onemogući instanciranje ove klase.

Osim toga, u sebi sadrži kao članove String-ove vowels i consonants, koji služe kao popis samoglasnika i suglasnika. Oni se koriste u metodama za provjeru slova unutar datoteke.

## LetterCount

LetterCount klasa služi kao spremnik za broj izbrojanih samoglasnika i suglasnika. Ti brojevi su pohranjeni u članovima vowels i consonants. Klasa ima tri getter metode, koje služe dobivanju vrijednosti izbrojenih samoglasnika: vowels(), suglasnika: consonants(), te ukupnog broja izbrojanih slova: total().

## count(String filename)

Glavna te ujedno jedina javna (statična) metoda ovog framework-a je count(String filename). Ova metoda prima jedan argument, filename. On predstavlja ime, odnosno putanju do datoteke za koju korisnik želi izbrojati samoglasnike i suglasnike.

Metoda kao rezultat vraća instancu LetterCount klase.

Ova metoda može baciti FileNotFoundException u slučaju da navedena datoteka ne postoji na disku korisnika. Također može baciti IllegalArgumentException u slučaju da datoteka ima neku drugu ekstenziju osim .txt ili .xml.

Unutar metode iz naziva datoteke se pravi instanca File klase, te se zatim jednostavnim grananjem vrši provjera tipa datoteke, te odabir metode za brojanje slova. U slučaju da je predana .txt datoteka, poziva se countInTxt(File file) metoda, a u slučaju .xml datoteke, poziva se countInXml(File file) metoda.

## countInTxt(File file)

Ova metoda služi za brojanje samoglasnika i suglasnika u .txt datoteci. To se postiže korištenjem Java klase Scanner. While petlja iterira po dostupnim linijama datoteke, te u for petlji prolazi po svakom znaku, te ispituje nalazi li se on u ranije navedenim vowels i consonants String-ovima.

Na kraju vraća brojeve u obliku instance LetterCount objekta.

## countInXml(File file)

Zbog specifičnog uvjeta zadatka, navedenog na početku ove dokumentacije, ova metoda delegira brojanje slova. Suglasnici se broje korištenjem prethodne metode – countInTxt(File file). Ovdje se gleda isključivo izbrojeni suglasnici, dok se samoglasnici odbacuju.

Samoglasnici se broje u countVowelsInXml(File file) metodi.

Na kraju se brojevi vraćaju u obliku nove instance LetterCount objekta.

## countVowelsInXml(File file)

Ova metoda služi za brojanje samoglasnika u vrijednostima svih atributa i elemenata u .xml datoteci.

U tu svrhu, koristi se DOM parser. Prvo se .xml datoteka uz pomoć DocumentBuilder-a parsira u objekt tipa Document. Zatim iz njega dobavljamo popis svih čvorova (NodeList).

Kako bi prebrojali samoglasnike u svim elementima, jednostavno uzimamo nulti čvor (root) iz NodeList-a, te brojimo samoglasnike u njegovom tekstualnom sadržaju, kojeg dobivamo pomoću .getTextContent() metode. Ovo radimo koristeći countVowels(String source) metodu, kojoj predajemo dobiveni tekstualni sadržaj kao argument.

Brojanje samoglasnika u svim atributima radimo iteracijom kroz sve čvorove u NodeList-u. Korištenjem metode .getAttributes(), dobivamo mapu atributa u čvoru. Zatim predajemo vrijednost atributa metodi countVowels(String source).

Metoda na kraju vraća zbrojeni rezultat prethodne dvije radnje.

## countVowels(String source)

Ova metoda broji samoglasnike u zadanom String-u. To radi iteriranjem kroz sve znakove tog String-a, te traži nalaze li se oni negdje u ranije navedenom vowels String-u.