

TESTAUSDOKUMENTTI

Aihe: Sanaindeksoija

Sovellus lukee käyttäjän antamista tekstitiedostoista sanat hakupuurakenteeseen, jonka jälkeen käyttäjän on mahdollista kysellä sovellukselta sanojen tai niiden yhdistelmien (haluamiensa merkkijonojen) esiintymisestä kyseisissä tekstitiedostoissa.

Ohjelman testaus

Ohjelman testaus suoritettiin pääosin käsin koodauksen edistyessä käyttämällä hyväksi apukomennoilla standardisyöttövirtaan tulostettavaa informaatiota. Manuaalitestattavia asioita olivat:

- käyttäjän syötteet ohjelmalle: tiedostojen lukumäärä, tiedostojen nimet, haun merkkijonot
 - todennettiin ohjelman toiminnalla
- annetun tiedoston löytyminen ja virheellisyys
 - todennettiin standardisyöttövirtaan tuotettavalla informaatiolla tiedoston nimen antamisen jälkeen
- rivien käsittelyn ja samalla hakupuun solmujen tallennuksen toteutus
 - todennettiin hakupuun tulostuksen avulla vertaamalla tulostusta syöttötiedoston materiaaliin
 - testaus aloitettiin pienestä, yksinkertaisesta syötteestä, jota kasvatettiin ja monipuolistettiin vähitellen
- rivien tallennuksen onnistuminen
 - todennettiin tallennustaulukon tulostamisen avulla
- merkkisolmujen tallennuspaikkojen selvittäminen
 - todennettiin tuottamalla standardisyöttövirtaan käsiteltävien merkkien int-arvot ja vertaamalla niitä toteutettuun ratkaisuun `selvitaInd()`
- hakutuloksen paikkansapitävyys
 - todennettiin vertaamalla ohjelman tuottamaa tulostusta syöttötiedoston materiaaliin
 - testaus aloitettiin pienestä syötteestä ja sitä kasvatettiin ja haettavaa merkkijonoa hankaloitettiin vähitellen huomioiden myös tallennukseen kelpaamattomien merkkien käsittelyn
- kahden hakutuloksen leikkauksen toteutus
 - todennettiin vertaamalla ohjelman tuottamaa tulostusta syöttötiedoston materiaaliin
 - testaus aloitettiin pienestä syötteestä ja alahakuja monimutkaistettiin vähitellen

Manuaalitestauksen lisäksi toteutettiin JUnitia hyväksikäyttäen yksikkötestausta. Yksikkötestaus kohdistui public-metodeihin, jotka palauttivat arvon. Näin testattiin lisäksi seuraavat toiminnot:

- sisäänluvun onnistuminen: palauttaa luettujen rivien lukumäärän, joka täsmää tiedoston rivimäärään
- tiedoston virheellisyys: palauttaa totuusarvon false, jos tiedostoa ei löydy
- merkkijonon löytyminen: palauttaa ensimmäisen jonon esiintymän
- merkkijonon löytymättömyys: palauttaa nullin
- rivin löytyminen tallennustaulukosta: palauttaa pyydetyn rivin sisällön
- hakupuun juuren löytyminen: palauttaa Node-solmun, jonka merkkiarvo on ''
- lapsen löytyminen: palauttaa Node-solmun arvolla, joka löytyy sisäänlukutiedostosta
- lapsen puuttuminen: palauttaa nullin, koska merkillä alkavaa sanaa ei materiaalissa ole
- lapsen tallentuminen: tekee tallennuksen äsken todettuun puuttuvaan soluun ja lukee sen takaisin
- lapsen listan löytyminen: palauttaa listan aiemmin löytyväksi tarkastetun merkin osalta
- lapsen listan puuttuminen: palauttaa nullin äsken lisätyn lapsen osalta
- lapsen listan ensimmäisen solmun tallentaminen: tekee tallennuksen äsken lisätylle lapsisolmulle ja lukee sen
- lisäys lapsen listaan onnistuu: tekee tallennuksen äsken lisättyyn listaan ja lukee sen

Ohjelman toteutus oli varsin virheetöntä alusta asti, jonka vuoksi suurimittaiseen toiminnan oikeellisuuden testaukseen ei ollut syytä.

Ohjelman nopeutta testattiin ohjaajan kehittämällä JUnit-testillä, jossa varioitiin hakupuuhun tallennettavan rivin pituutta ja rivien lukumäärää ja mitattiin tallennuksen ja haun kestoa.

Testisyötteen

Testisyötteinä käytettiin seuraavia tekstitiedostoja, jotka löytyvät materiaalista:

- testi1.txt: 2 riviä, blankkoja rivin alussa ja lopussa
- testi2.txt: 3 riviä, samalla kirjaimella/-milla alkavia sanoja
- testi3.txt: 5 riviä, lyhyitä rivejä ja toistuvia alkuja, hakupuun tulostusta varten
- rautatiekpl1.txt: Juhani Ahon kirja Rautatie, kappale 1, 393 riviä
- rautatie.txt: Juhani Ahon kirja Rautatie, 2992 riviä
- holmes.txt ja dracula.txt: kirjat The Adventures of Sherlock Holmes ja Dracula, yhteensä 23539 riviä, käytettiin yhdessä

testaamaan toiminnallisuutta isomman massan yhteydessä

Metodien oikeellisen toiminnan testaamiseen tehtyjä JUnit-testejä käytettiin rautatiekpl1.txt-tiedoston yhteydessä.

Nopeustestissä tulokset olivat seuraavat:

- 1000 rivin lisäys 110 ms, 4-merkkisen sanan haku 0 ms
- 20000 rivin lisäys 3656 ms, 6-merkkisen sanan haku 0 ms
- 60000 rivin lisäys 9954 ms, 6-merkkisen sanan haku 0 ms

Testauksen toistettavuus

JUnit-testit ovat löydettävissä materiaalista hakemistoista
/Tiraharj/Sanaindeksoija/test/Kayttoliittyma ja
/Tiraharj/Sanaindeksoija/test/Logiikka.

Syötetiedostot sijaitsevat hakemistossa /Tiraharj.