

Viikkoraportti 5

Viimeisellä viikolla olen hionut projektia arviointia ja esittelyä varten. Nyt on esittelyt suoritettu, enää olisi pienet viimeistelyt lopullista arviointia varten. Tälle viimeiselle viikolle jäi vielä yksi aika massiivinen bugi, jota en ollut huomannut. Viime viikolla kehuskelin, miten aakkoston ja datatiedoston kirjoittaminen yhteen sujui helposti. Todellisuudessa tämä ei ollutkaan ihan niin helposti toteutettavissa. Yhdellä testitiedostolla merkit menivät suorastaan täysin sekaisin. Virhe oli hyvin hämmentävä, koska muut tiedostot toimivat oikein.

Epäilen virheen olleen käyttämässäni OutputStream -luokassa. Kirjoittelin sillä datatavut ja puun koodaukseen tarvittavat characterit täysin samalla tavalla, koska kirjoittajalle charin antaminen byten sijasta näytti toimivan aivan hyvin. Kun aloin lukea juttua internetin koodausfoorumeilta, varoiteltiin siellä, ettei charien kirjoittaminen surutta tavuina toimi aina ihan oletetusti. Vaihdoinkin sitten tiedostonkirjoittajan DataOutputStream- olioksi, joka on ilmeisesti jonkinlainen kääritty versio OutputStreamista. DataOutputStreamilla (ja DataInputStreamilla) onkin metodit erikseen kaikkien eri tietotyyppien lukemiselle.

Tämän muutoksen jälkeen pakkausohjelma alkoi taas toimimaan moitteettomasti, mutta aakkostotiedoston koko kasvoi. Tämän johtunee siitä, että OutputStream kirjoitti merkit tavuina, kun taas DataOutputStreamissa merkki vie 16 bittiä, eli kaksi tavua. Periaatteessa nuo merkit onnistuisi hyvin ilmaisemaan 8:lla bitillä (UTF-8) mutta en uskalla lähteä kikkailemaan 8-bittisten merkkien kanssa. Olen tyytyväinen siihen, että ohjelma toimii, vaikka aakkoston koko onkin vähän suurempi!

Kirjoitin myös testausdokumenttia ohjelmalle. Varsinaisella ohjelmalla ei ole automatisoituja testejä. En ole keksinyt järkevää tapaa testata tiedostoa kirjoittavia metodeja. Automaattiset testit olisi kyllä ollut poikaa. Esim tuon ylläolevan bugin olisin testien avulla varmasti huomannut ennen sitä viimeistä iltaa ennen demoa! Huffman-koodaajalle voisin kyllä yrittää kirjoittaa vielä jokusen testin. Sitä ehkä pystyisikin melko helposti testaamaan.