

Viikkoraportti 4

Tällä viikolla sain ratkaistua edellisellä viikolla purkualgoritmissa huomaamani ongelman, jossa purettavaksi ilmestyi toisinaan, vaikkakin harvoin dekoodaajalle tuntemattomia koodeja. Näin tapahtuu, kun tekstissä on toisteisuutta, esim. "AAA". Tätä yhdellä videolla kutsuttiin "tricky caseksi". Ratkaisuna ongelmaan oli arvata koodi algoritmin edellisestä tulostamasta merkkijonosta, johon liitetään edellisellä kerralla sanakirjaan tallennetun merkkijonon ensimmäinen merkki.

Seuraavaksi ryhdyin rakentamaan luokkaa bitti- ja tavumuunnoksia varten. Tästä vaiheesta tuli työteliäs, sillä olin pakkausalgoritmin toimivuudesta varmistuakseni tallentanut 1. versiossa pakatut merkit kokonaislukuina. LZW:n parantaa tehoa pakkaamalla merkkejä tiettyyn bittimittaan, esim. 9 bitin mittaiseksi, kun sanakirjan suurin koodi on alle 2 potenssiin 9. Jouduin mylläämään koodia runsaasti ja tiedostonkäsittelyäkin muutin jonkun verran. Viikkotyöajan loputtua pakkauksen purkualgoritmi ei taaskaan toimi tai ainakaan purettu ja alkuperäiseen muotoon tallennettu tiedosto ei vastaa alkuperäistä jostain syystä.

Yllätyin kurssin tiukasta aikataulusta katsottuani (kovin vähien) tulevien viikkojen ohjelmaa. Olen miettinyt lista- ja puurakenteita, joilla korvaisi Javan valmiit tietorakenteet, mutta mitään konkreettisia hahmotelmia en ole koneelle tai paperille vielä ehtinyt tehdä. Myöskin vertaisarvioinnin jouduin jättämään väliin, vaikka arvioitavaksi annettu Huffman-algoritmi olisi ollut kiintoisa kohde arviolle verrattuna omaan työhöni.