

Viikkoraportti 2

Tällä viikolla opiskelin LZW-koodausta. Tällä hetkellä se toimii joten kuten: esimerkiksi 100000 satunnaista tavua menee hyvin pakkaus-purkuprosessin läpi, mutta jos otetaan 1000000 tavua, niin lopputuloksessa on vain 999999 tavua ja ainakin osa niistä väärää. Ongelma pahenee, jos koodisanojen kokoa kasvattaa yli 12 bittiin ja poistuu (ainakin siltä näyttää), kun koodisanan koko on pienempi. En tiedä vielä, missä vika on, mutta epäilen, että se liittyy LZW-purkamisessa olevaan kuuluisaan (?) poikkeustapaukseen.

Huffman-koodista poistin ArrayListit ja korvasin ne itse tehdyllä BitVectorilla. Lisäksi värkkäsin bittistreamit, joiden läpi pakattua koodia on helppo kuljettaa. Pääohjelmassa voi valita, käyttääkö Huffmania vai LZW:tä.

Seuraavaksi LZW-bugi pitää löytää ja korjata. LZW-koodia voisi muutenkin siistiä ja kommentoida paremmin.