

Tällä viikolla minulla on mennyt paljon enemmän aikaa lukemiseen kuin koodaamiseen.

Sain selville, että JPEG käyttää valmiiksi laskettuja Huffman -tauluja, joiden käyttäminen ei opettaisi Trie muodostusta. Aloin siis toteuttaa omaa Huffman -koodausta.

Tein ensin perinteisen Huffman -koodauksen, eli laskin jokaiselle arvolle oman esiintyvyytensä. Huomasin kuitenkin, että testikuvassani oli nollia 92%. Olisi siis täysin tehotonta, jos jokainen nolla koodattaisiin tiedostoon omana merkinään. Päätin jättää nollat pois esiintyvyytlaskennasta ja muuttaa koodiani niin, että Trie-rakenteeseen tulevat kaikki muut arvot ja nollaa ei koodata ollenkaan, vain kuinka monta nollia on missäkin kohdassa.

Tällä hetkellä Huffman -koodaukseeni kuuluu koodattuna osat (jotka olen testannut omina yksikköinä):

- frekvenssitaulun muodostus
- Node -luokka
- puun muodostus Node -minimikeon avulla. En tiedä ehdinkö implementoimaan omaa minimikekoa tässä projektissa.

Huffman -koodauksesta puuttuu vielä tiedostoonkirjoitus. Aluksi ajattelin jonkunlaista koodiarvot -taulukkoa jonka sisältö kirjoitettaisiin tiedostoon niin, että jokaisessa taulukon kohdassa olisi joko kyseisen noden arvo tai viittaus taulukon indekseihin joissa tämän noden lapset ovat. Sitten tutustuin tähän esimerkkiin: <http://algs4.cs.princeton.edu/55compression/Huffman.java.html>. Aion seuraavalla viikolla tehdä tiedostoonkirjoituksen sekä -luvun tuolla tyylillä.

Luulen että voisin tarvita ohjausta, tai ainakin kommentin, miten trie-puun tiedostoonkirjoitus kannattaisi tehdä ja onko linkin tapa paras.