

Yleisrakenne

Merkkilaskuri -luokka laskee kunkin merkin esiintymismäärät pakattavassa tiedostossa ja palauttaa taulukon jossa määrät on tallennettu merkkiä vastaavan indeksin kohdalle. Pakkaaja -luokka hoitaa tiedostojen pakkaamisen. Aluksi merkkilaskurilla lasketaan merkkien esiintymismäärät. Jokaista pakattavassa tiedostossa esiintyvää merkkiä varten luodaan solmu. Solmu -luokka on tietorakenne joka sisältää merkin jota solmu vastaa, merkin määrän sekä solmun vasemman ja oikean lapsen. Solmuista muodostetaan puu siten että solmut jotka vastaavat usein esiintyviä merkkejä ovat lyhyen matkan päässä puun juuresta. Puun rakentamisessa hyödynnetään minimikeon toimintoja toteuttavaa luokkaa Keko. Merkkien lyhennemerkit muodostetaan kulkemalla puuta juuresta alaspäin. Vasemmalle kulkeminen vastaa nollaa ja oikealle kulkeminen ykköstä. Lopuksi pakkausmetodi kirjoittaa merkkipuun ja pakatun datan tiedostoon. Bittikirjoitin -luokka mahdollistaa pakatun tiedoston kirjoittamisen tiedostoon bitteinä.

Purkaja -luokka pystyy purkamaan tällä ohjelmalla pakattuja tiedostoja. Aluksi merkkipuu luetaan pakatun tiedoston alusta. Sitten pakattu tiedosto käydään läpi bitti kerrallaan. Kun tiedostosta luetaan 0, puussa kuljetaan vasemmalle ja kun 1 oikealle. Kun saavutaan solmuun joka on lehti, kirjoitetaan solmua vastaava merkki. Bittilukija -luokka mahdollistaa pakatun tiedoston lukemisen bitti kerrallaan.

Puutteet/parannusehdotukset

Ohjelma ei pysty käsittelemään tiedostoja jotka sisältävät merkkejä joiden numeroarvo on suurempi kuin 255 (8 bittiä, laajennettu ASCII).

Lähteet

http://en.wikipedia.org/wiki/Huffman_coding

https://www.siggraph.org/education/materials/HyperGraph/video/mpeg/mpegfaq/huffman_tutorial.html