

Jatkoin tällä viikolla suorituskyykytestausta. Mukana on tällä viikolla uusi algoritmi LZ77 jonka toiminta on vielä todella heikkoa. Lisäksi huomasin vasta silloin kuin aloitin testauksen että hashmappini ei toimikaan hyvin LZW algoritmin kanssa joten en ehtinyt korjaamaan sitä viikkopalaukseen mennessä. Suuria ongelmia aiheutti myöskin excelin graafisten esitysten teko.

Testasin 5 eri tiedostoa, josta 3 ovat samoja kun viime viikolla, pieni teksti tiedosto README.TXT, valtava tekstitiedosto bible.txt ja html tiedosto web.html. Lisäksi otin mukaan 2 keskikokoista testitiedostoa medium.txt ja copypaste.txt, molempien koko on 269k tavua. Toinen sisältää random tekstiä mitä löysin netistä ja toinen on sisältää saman noin 10 rivin pätkän tekstiä copypastettuna.

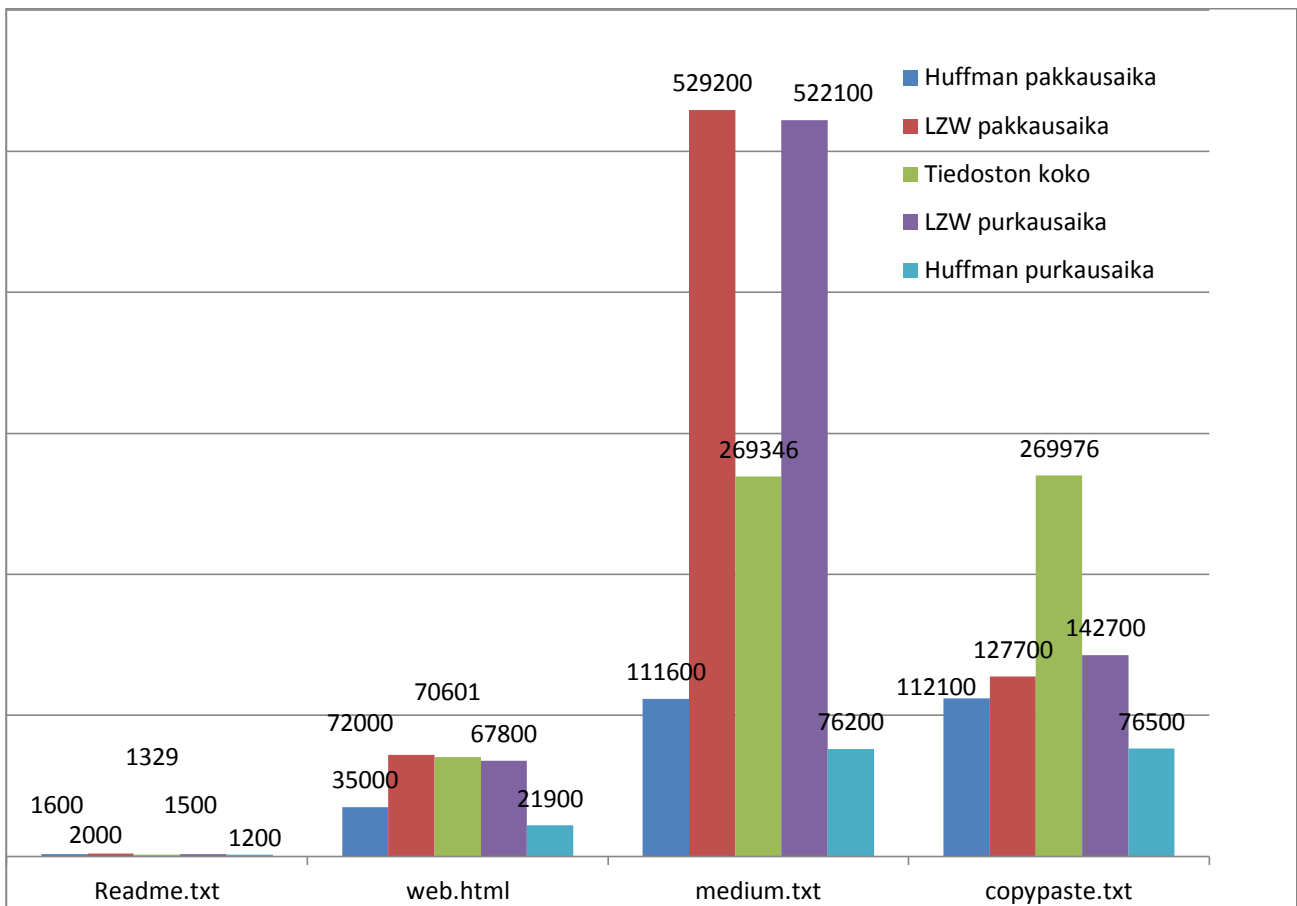
Laitoin kesken olevan LZ77 tulokset eri esityksiin kun LZW ja Huffmanin koodi, koska ne on varsinkin suoritusaajoiltaan niin eri luokkaa että muuten kuvaajista ei tulisi järkevän näköisiä.

En myöskään ottanut kuviin mukaan bible.txt pakkausta ja purkua koska Huffman ainostaan pystyi tekemään tämän järkevässä ajassa. LZW pääsi hurjaan 3026386 ms ja LZ77 ei ole vielä valmista, tosin pistin molemmat pakkamaan samaan aikaan joten se ehkä hieman hidastaa menoa.

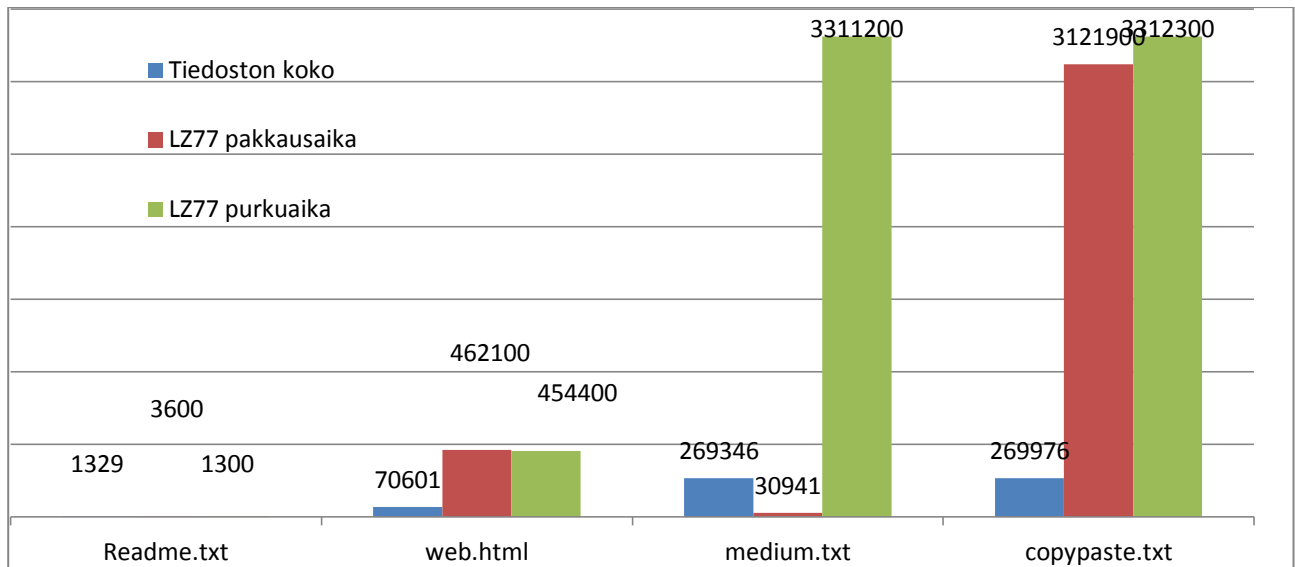
Bible.txt pakkauksen koot ovat kuitenkin huffmanilla : 3068618 tavua ja LZWllä 2156266 tavua, alkuperäinen koko oli 5498995. Lisäksi sen pakkaus vei huffmanille 20731 ms ja purku 14873 ms.

LZ-dynaastia otti muutenkin turpiin pahasti huffmanilta kaikessa muussa paitsi copypaste tiedoston pakkauskooassa. Alla on graafiset esitykset kaikesta. Olen kaikki 3 algoritmia laittanut samaan pakkauskokoa demonstroivassa esityksessä. **Niin ja aika on sitten 10^5 sec (koska en excel noobina osaa laittaa sitä näyttämään hyvältä muuten), ja koko tavuina kaikissa esityksissä.**

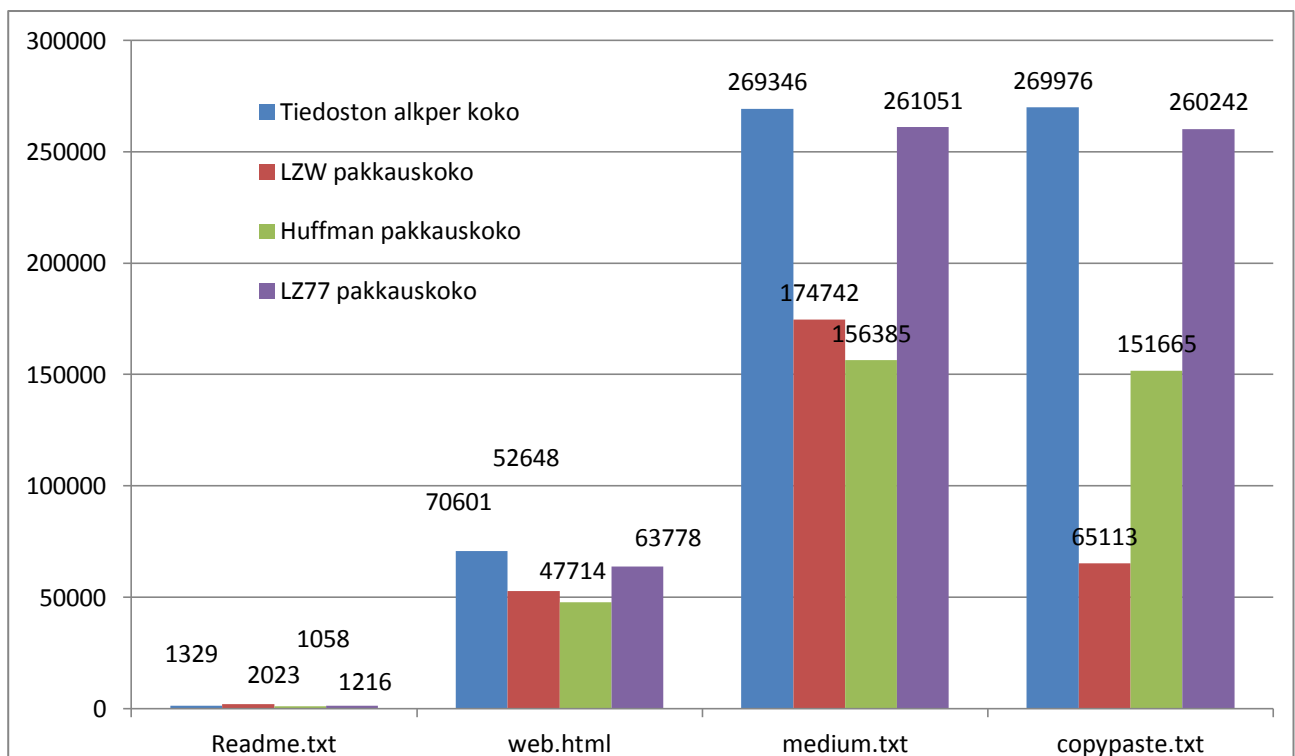
Pakkaus ja purkuajat suhteessa tiedostonkokoan LZW ja huffmanin koodi algoritmeilla:



Pakkaus ja purkuajat suhteessa tiedostonkokoan LZ77 algoritmilla.



Pakkauskoko vertailu LZW, Huffmanin koodi ja LZ77 algoritmeilla.



Niin ja asia mikä suoritussykyyn liittyen ihmetyttää on: miksi LZ77 toimii noin hitaasti purkaessa, sen pitäisi olla teoriassa tosi nopea silloin..