

Øving 08 - Algoritmer og datastruktur

Gruppe 17: John I. Eriksen og Emil Slettbakk

Huffman-problemer

Kompressjon av filene diverse.txt og diverse.lyx fungerte til slutt. Første fungerende variant av dekompressjonsdelen ville ikke skrive "spesialtegn" korrekt, som æ, ø, å, kulepunkter osv.

Løsningen gikk ut på å gå fra å lagre bytes som String til å benytte byteArray, hvilket hjelper med å skille de forskjellige tegnene. Innførte også en metode for å håndtere "padding" bedre, noe som tydeligvis er et poeng når det kommer til Huffman. I tillegg ble FileWriter byttet ut med DataOutputStream, siden denne (i følge det store Internett) er bedre egnet om man vil være sikker på at hver byte blir skrevet riktig.

Kompilering og kjøring

Programmet kan kjøres direkte i IDE-et, eller via kommandolinja. Argumentene for dette ligger i `main` metoden i både `Huffman` og `HuffmanDecompressor`. Begge variantene ble testet på både Linux og Mac.

"Direkte i koden" fra IDE-et, gjøres ved å skrive inn filnavn. Her har vi kompressjonen skrives til en fil med filendelse `.huff` for å indikere den Huffman-komprimerte fila. På samme måte i dekomprimeringen, hvor input-fila (altså den komprimerte) som skrives og gis filendelsen `.decomp` som ferdig dekomprimert fil. Dermed er det enkelt å se forskjell på dem, og originalfilene blir ikke overskrevet. Utklipp fra koden i kodeblokkene følger under.

```
// Compression
String inputFilePath;
String outputFilePath;
// (...)
inputFilePath = "diverse.txt"; // Type in the name of the file to compress.
outputFilePath = inputFilePath + ".huff"; // Add .huff suffix to file to not
overwrite original
```

```
// Decompression
String inputFilePath;
String outputFilePath;
// (...)
inputFilePath = "diverse.txt.huff"; // Type in the name of the file to be
decompressed.
outputFilePath = inputFilePath + ".decomp"; // Add .decomp suffix to file to not
overwrite original file.
```

Det kan også kjøres direkte i kommandolinja.

Først må programmene kompileres. Huffman og HuffmanDecompressor er separate, og begge må dermed kompileres:

```
javac Huffman.java HuffmanDecompressor.java
```

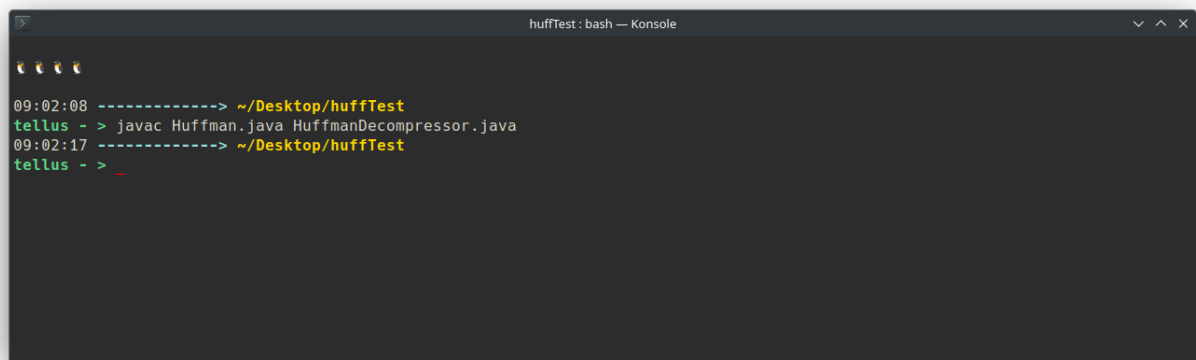
Kompressjon: Kjør Huffman-klassen med input.fil og output.fil:

```
java Huffman diverse.txt diverse.txt.huff
```

Dekompressjon: Kjør HuffmanDecompressor-klassen med input.fil og output.fil:

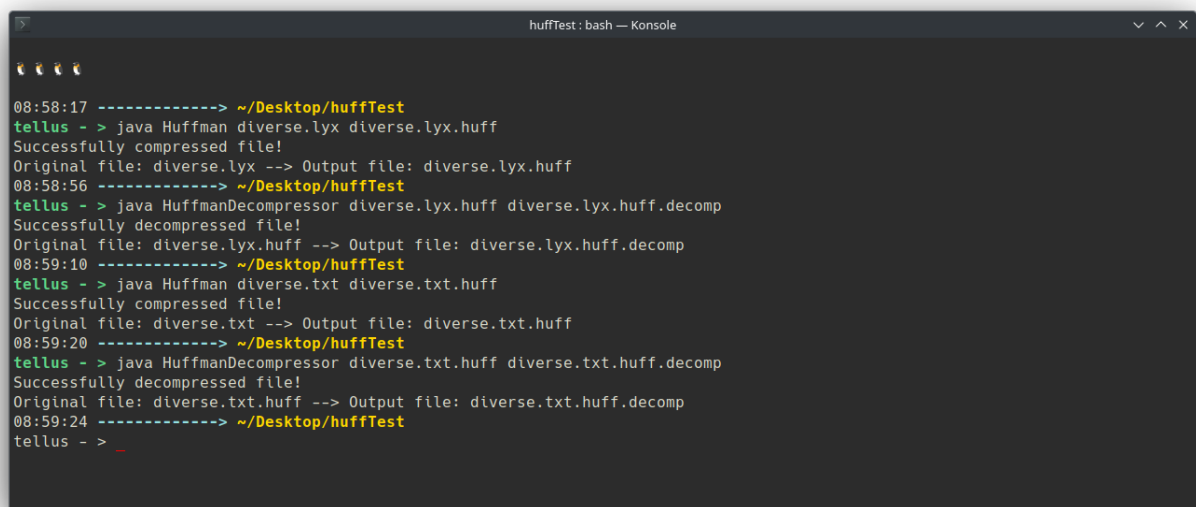
```
java HuffmanDecompressor diverse.txt.huff diverse.txt.huff.decomp
```

Kompilering i kommandolinja:



```
huffTest : bash — Konsole
09:02:08 -----> ~/Desktop/huffTest
tellus - > javac Huffman.java HuffmanDecompressor.java
09:02:17 -----> ~/Desktop/huffTest
tellus - > _
```

Komprimering og dekomprimering utført i kommandolinja

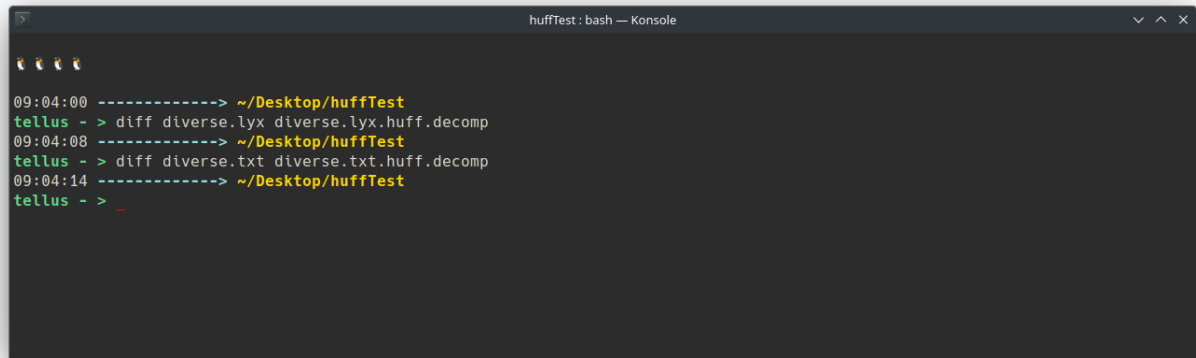


```
huffTest : bash — Konsole
08:58:17 -----> ~/Desktop/huffTest
tellus - > java Huffman diverse.lyx diverse.lyx.huff
Successfully compressed file!
Original file: diverse.lyx --> Output file: diverse.lyx.huff
08:58:56 -----> ~/Desktop/huffTest
tellus - > java HuffmanDecompressor diverse.lyx.huff diverse.lyx.huff.decomp
Successfully decompressed file!
Original file: diverse.lyx.huff --> Output file: diverse.lyx.huff.decomp
08:59:10 -----> ~/Desktop/huffTest
tellus - > java Huffman diverse.txt diverse.txt.huff
Successfully compressed file!
Original file: diverse.txt --> Output file: diverse.txt.huff
08:59:20 -----> ~/Desktop/huffTest
tellus - > java HuffmanDecompressor diverse.txt.huff diverse.txt.huff.decomp
Successfully decompressed file!
Original file: diverse.txt.huff --> Output file: diverse.txt.huff.decomp
08:59:24 -----> ~/Desktop/huffTest
tellus - > _
```

Sjekk med `diff`

Kompressjon og dekompressjon ble gjennomført uten feil på begge filene (diverse.txt og diverse.lyx).

Det ble utført manuelle punksjekker på de problemene som oppsto tidligere med spesialtegnene, i tillegg til å bli sjekket med `diff`.



```
huffTest : bash — Konsol  
09:04:00 -----> ~/Desktop/huffTest  
tellus - > diff diverse.lyx diverse.lyx.huff.decomp  
09:04:08 -----> ~/Desktop/huffTest  
tellus - > diff diverse.txt diverse.txt.huff.decomp  
09:04:14 -----> ~/Desktop/huffTest  
tellus - > _
```

Skjermbilde av `diff` med original vs dekomprimert. for begge filene i oppgaven.

Det ble gjort forsøk på å komprimere enwik8.txt, men etter "veldig lang tid" ble dette eksperimentet avbrutt.

Ingen pengepremie på oss. 🎁🎁🎁