

```
jan-erik@se20884419:~/Dokument$ cat hitchhikersguide.txt |  
tr -cs A-Za-z '\n' | tr A-Z a-z | sort | uniq -c |  
sort -rn | head -n 15  
2538 the  
1259 of  
1199 a  
1187 to  
1184 and  
892 it  
808 he  
681 said  
632 you  
630 in  
607 was  
586 that  
563 i  
396 s  
366 on  
jan-erik@se20884419:~/Dokument$
```

## ~ Laboration – Att räkna antalet förekomster av ord i en text ~

### Introduktion till programmering

— 2021-11-29 v1 —

Revisioner

2013-03-19 Daniel Bosk

2021-11-29 Jan-Erik Jonsson

***Läs igenom hela instruktionen innan du börjar!***

## Introduktion

År 1986 gav Jon Bentley en utmaning till Donald Knuth. Utmaningen var att skapa ett program som läser innehållet i en text, bestämmer de  $n$  mest förekommande orden, och skriver ut en sorterad lista tillsammans med antalet förekomster av respektive ord. Knuth löste naturligtvis problemet. Likaså gjorde Douglas McIlroy, fast med ett mycket kort skalskript.

McIlroys lösning:

```
n = 50
cat hitchhikersguide.txt | tr -cs A-Za-z '\n' | tr A-Z a-z | \
    sort | uniq -c | sort -rn | head -n $n
```

Denna kan köras direkt i terminalen på ett UNIX-lik system och varje kommando har en manualsida där du kan läsa vad kommandot gör med de givna argumenten.

## Syfte

Syftet med laborationen är att du ska få bekanta dig med ytterligare några C++-konstruktioner och tillämpa dessa för att lösa ett problem. Se läsanvisningarna om dessa.

## Läsanvisningar

Du bör ha läst om följande datastrukturer i standardbiblioteket för C++:

```
std::string
std::vector
std::map
```

Därefter bör du ha läst om följande funktioner från standardbibliotekets algoritmdel:

```
std::transform
std::sort
std::find
std::count
std::max_element
std::unique
```

Följande funktioner från C-biblioteket kan också vara till hjälp:

```
tolower
toupper
```

## Genomförande

- Du ska lösa Bentleys problem, men naturligtvis i C++.
- Som indata har du den bifogade textfilen *hitchhikersguide.txt*.
- Resultatet ska skrivas ut i terminalen.
- Utdata från ditt program ska vara:
  - de  $n$  mest förekommande orden,
  - sorterade i ordningen mest förekommande först.
  - Antalet,  $n$ , ska enkelt kunna ändras i programmet.

## Examination

Du redovisar din lösning muntligt för din labbhandledare. När du fått godkänt lämnar du in *källkoden* och *byggskriptet* i lärplattformens inlämningslåda för denna uppgift.