

# Scriptingtalen

Daan Pape  
Universiteit Gent

24 juni 2012

# 1 Sed

Sed biedt uitgebreide ondersteuning voor lijnadressering. Volgende zijn de belangrijkste:

- `x~N`: dit behandelt alle lijnen vanaf `x` en elke volgende lijn welke een veelvoud is van `N`.
- `/regex/`: matcht alle lijnen waarin `regex` voorkomt.

Inplaats van na een *range* aanduiding meteen een commando te plaatsen kan men ook een commandogroep plaatsen:

- `range{}`: binnen de commandogroep kan men alles plaatsen, ook ranges, en dit zal uitgevoerd worden als het binnen de range valt.
- `range!{}`: binnen de commandogroep kan men alles plaatsen, ook ranges, en dit zal uitgevoerd worden als het buiten de range valt.

Zoals veel programma's heeft sed enkele regex specialiteiten:

1. `\{x\}`: duidt op het `x` keer herhalen van de voorige karakter match.
2. `s/./&/`: de `&` is een backreference naar de volledige match.

Specifieke sed commando's:

- `h`: plaatst de `pattern space` in de `hold space`, deze wordt dus overschreven.
- `H`: voegt de `pattern space` toe aan de `hold space`, deze wordt dus groter.
- `g`: plaatst de `hold space` in de `pattern space`, deze wordt dus overschreven.
- `G`: voegt de `hold space` toe aan de `pattern space`, deze wordt dus groter.
- `n`: leest de volgende lijn input in, in de `pattern space`.
- `N`: voegt de volgende lijn input toe aan de `pattern space`.
- `x`: verwisselt de inhoud van de `pattern space` en de `hold space`.

Sed truucjes:

1. `:a;N;$!ba`: lees een bestand in de `pattern space` in 1 lijn `n`.

# 2 Awk

Sed is een uitgebreidere scriptingtaal en bevat enkele handige dingen. Men moet allereerst alles in `{}` blokken schrijven. volgende zijn handige weetjes:

- `$x`: duidt op kolom `x`.
- `getline var < "-"`: lees de input van `stdin` in in de variabele `var`.
- `index($0, type){}`: doet het zelfde als `/type/{}`:

Awk bezit alles in lijnen en kolommen. De *field separators* zijn daarbij zeer belangrijk. Er zijn vele variabelen

1. IFS: stelt de input field separator in.
2. OFS: stelt de output field separator in.
3. FS: stelt zowel de input als de output field separator in.
4. ARGV: een array die alle command line arguments bevat.
5. ARGC: het aantal command line argumenten.

### 3 Bash

Bash is een uitgebreide scriptingtaal die gebruik maakt van alle mogelijke commando's. Zo zijn uitgebreide conditionele opdrachten mogelijk:

- [ `expr1 -a expr2` ]: beide uitdrukkingen moeten waar zijn
- [ `expr2 -o expr2` ]: een van de uitdrukkingen moet waar zijn
- `$(echo ${#string})`: geeft de lengte van de string terug
- `${#}`: geeft terug hoeveel commandolijn argumenten er werden opgegeven
- `$(mktemp)`: geeft een pad naar een tijdelijk bestand terug. Dit moet wel verwijderd worden na gebruik.
- `{a..b}`: geeft alle getallen tussen `a` en `b` terug met `a` en `b` inclusief.

Met het `getopts` commando kunnen commandolijnargumenten worden opgevraagd, stel `:f::`

- `:`: een `:` in het begin onderdrukt errors van `getopts` zelf.
- `:`: de letters in het midden duiden commando argumenten aan.
- `:`: de `:` op het einde geeft aan dat er een verplicht argument moet zijn.

### 4 Python

Regex in python kan door middel van de `re` bibliotheek:

1. `re.sub("match", "substitutie", string)`
2. `bool(re.search("match", string))`: geeft `True` terug bij een match

Handige standaard python functies zijn:

1. `ord(char)`: geeft de ASCII plaats terug van het karakter
2. `sorted(list)`: geeft een gesorteerde lijst terug
3. `len(...)`: geeft de lengte terug
4. `str.rstrip(chars)`: verwijderd de whitespace rechts van de string indien geen karakterlijst werd opgegeven.

5. `str.lstrip(chars)`: verwijderd de whitespace links van de string indien geen karakterlijst werd opgegeven.
6. `str.strip(chars)`: verwijderd de whitespace links en rechts van de string indien geen karakterlijst werd opgegeven.
7. `str.center(int)`: geeft een string terug met whitespace links en rechts zodat de string in het midden staat.

Een dictionary kan je aanmaken door `var=dict()` of `var={}`. Om een `assertionError` op te gooien kan men het `assert statement, errorstring` (make shure that) gebruiken.