

F4270 UNIX, počítačové sítě (jaro 2016), Cvičení 9

Příkazy: awk. Stáhněte si <http://physics.muni.cz/~zbona/unix/awktut.pdf>. Prostudujte si.

- A1** Napište awk příkaz, který očísluje řádky na vstupu a vypíše je i s čísly (jako n1).
[guy@nix ~]\$ awk '{print NR " " \$0}' /etc/passwd
- A2** Napište awk příkaz, který vypíše druhý a pátý sloupec za souboru.
[guy@nix ~]\$ awk -F ":" '{print \$2 " " \$3}' /etc/passwd
- A3** Napište awk příkaz, který vypíše každou desátou řádku vstupu.
[guy@nix ~]\$ awk '(NR % 10) == 0' /etc/passwd
- A4** Napište awk skript, který v souboru /etc/passwd převede písmena v loginech na velká.
[guy@nix ~]\$ awk 'BEGIN { FS=":"; OFS=":" } { \$1=toupper(\$1); print}' /etc/passwd
- A5** Napište awk skript, který otočí pořadí slov na každé řádce vstupu (napíše je v opačném pořadí)

```
#soubor: skript_01.awk
{
  for (pole = NF; pole > 1; pole --)
  {
    printf ("%s" OFS, $pole);
  }
  printf ("%s" ORS, $1);
}
```

```
[guy@nix ~]$ awk -F ":" -f skript_01.awk /etc/passwd
```

- A6** Stáhněte si soubor <http://physics.muni.cz/~zbona/unix/reakce>
[guy@nix ~]\$ wget <http://physics.muni.cz/~zbona/unix/reakce>
- A7** V souboru reakce jsou uvedeny koeficienty pro výpočet chemických reakčních konstant. Napište skript v awk, který vyčíslí reakce s '*' podle vztahu

$$k_* = A \times 10^{-9} + B \times 10^{-11}\xi + C \times 10^{-13}\xi^2,$$

pro $\xi = 55.5$. Ostatní podle vztahu

$$\log(k) = A + B/\xi + C/\xi^2,$$

pro $\xi = 37.6$.

```
#soubor skript_eval.awk
{
  if(NR>=2)
  if( match($1,"*"))
  { x=37.6; print $1("x"): "10**($2+$3/x+$4/x**2);}
  else
  { x=55.5; print $1("x"): "$2*1e-9+$3*x*1e-11+$4*1e-13*x**2; }
}
```

```
[guy@nix ~]$ awk -f skript_eval.awk reakce
```

- A8** Napište awk skript, který vypíše tolik náhodných čísel mezi 0 a 1, kolik dostane zadáno parametrem.

```
#!/bin/sh
# Toto je skript, který obaluje awk skript
# Sám dostane počet potřebných parametrů ve svém parametru.
awk -v pocet="$1" 'BEGIN {
  srand();
  for (i=0; i<pocet; ++i)
  {
    print rand ();
  }
}'
```

A9 Napište awk skript, který z výstupu `getent passwd` vypíše plná jména uživatelů, jejichž login má tvar čtyři písmena.

```
#!/bin/awk -f

BEGIN {
    FS = ":";
}

$1 ~ /^[[:alpha:]]{4}$/ {
    print $5;
}
```

A10 Napište awk skript, který vypíše počet linků, tedy tagů `<a` ve vstupním html souboru.

```
#!/bin/awk -f
# Jako RS zvolíme "<", protože podle toho se dělí HTML soubor
# nejvhodněji. Specifikace ani nepřipouští víceznakové RS (v tom
# případě je výsledek nedefinovaný), regexp může být jen ve FS, tak to
# ani jinak nejde.
#
BEGIN {
    RS = "<";
    pocetAcek = 0;
}

($1 ~ /^a$/) {
    ++ pocetAck;
}

END {
    print pocetAcek;
}
```

A11 Napište awk skript, který dostane na vstupu html soubor a vypíše adresy, na které se tento soubor odkazuje (to jest za značku `` vypíše `adresa`). Pozor na to, že ta značka může být rozložena přes víc řádků a můžou tam být kolem `a`, `href` mezery. Volbou vhodného RS a FS si výrazně usnadníte práci.

```
#!/bin/awk -f
# Je to trochu komplikovanější, aby to fungovalo i v případě, že není
# href hned za a, ale je mezi tím ještě něco. Jinak by stačilo rovnou brát i=1.
BEGIN {
    RS = "<";
    FS = "=";
}

$1 ~ /^[[:blank:]]*a[[:blank:]]+/ {
    i=1
    while (i <= NF)
    {
        if ($i ~ /^[[:blank:]]+href[[:blank:]]*$/)
        {
            split ($(i+1), pole, "[[:blank:]]*" "[[:blank:]]*");
            print pole[2]
        }
        else if ($i ~ />/)
        {
            break;
        }
        ++i;
    }
}
```