

Perl – operacje na plikach

W dużym stopniu będziemy mieli do czynienia z plikami tekstowymi. Dostęp do plików realizowany jest przez tzw. uchwyt – zmienne, których sama wartość niewiele nas interesuje, lecz które wskazują na plik.

Otwieranie pliku:

```
open (my $FH, '<', "/etc/passwd") or die("Cannot open file!");
```

Zmienna \$FH jest naszym uchwytem. Odczytajmy plik linia po linii:

```
while (my $linia = <$FH> ) {  
    print $linia;  
}
```

Oczywiście możemy tutaj użyć instrukcji foreach (foreach my \$linia (<\$FH>)), ale to utworzy nam tymczasową tablicę zawierającą wszystkie linie z pliku, co nie jest korzystne pamięciowo. W pętli while za każdym razem z uchwytu odczytywana jest jedna linia.

Zamiast '<' możemy otworzyć plik do zapisu ">" lub dopisać do pliku ">>". Po zakończeniu korzystania z pliku należy go zamknąć (close \$FH;) aby wyczyścić bufor i zapisać zaległe dane.

Bardziej rozbudowany przykład:

```
open(my $FH, '<', "/etc/passwd") or die("cannot access file!");  
open(my $OPH, ">", "test.txt") or die("cannot access savefile!");
```

```
while (my $linia = <$FH>) {  
    if (index($linia,"is2014") == -1) {  
        next;  
    }  
    my @Q=split(':', $linia);  
    print "Znalazłem: ".$Q[2]." ".$Q[4]."\n";  
    print $OPH $Q[2]." ".$Q[4]."\n";  
}
```

```
close $FH;  
close $OPH;
```

Użyte funkcje:

index(\$string, \$substring) – zwraca położenie \$substring w \$string. -1 gdy go nie znajdzie. Użyteczne w wyszukiwaniu.

split(\$delimiter, \$lancuch) – zwraca tablicę ciągów po podzieleniu łańcucha wejściowego.

Uwaga: Podejście print split(':', \$linia)[1] nie uruchomi się ze względu na nieznany typ podczas parsowania.

print \$HANDLE string... – drukuje ciąg tekstowy do pliku wskazanego przez \$HANDLE.

Gdy już przeczytaliśmy...

Jeżeli przeczytaliśmy odpowiednie linijki i chcemy się cofnąć, możemy przenieść się na początek poleceniem seek(\$FH, 0, 0).

Jeszcze jedno użyteczna instrukcja do operacji na tekście: substr – fragment ciągu znaków.

```
my $str="Ala ma kota";  
print substr($str,4,2);
```

Wydrukuję:

ma

Można również używać tej instrukcji do zamiany fragmentów:

```
my $str="Ala ma kota";  
substr($str,4,2,"miała");  
print $str;
```

"Ala miała kota".

Dodatkowo używając "index" można wykonywać zamiany znanych części zmiennej. Długość łańcucha tekstowego uzyskujemy używając funkcji `length($string)`.

Wprowadzenie do wyrażeń regularnych – operator `=~`

Podstawowym narzędziem do modyfikowania łańcuchów tekstowych w Perlu są jednak rozbudowane mechanizmy wyrażeń regularnych. Powyższy przykład można przepisać:

```
my $str="Ala ma kota";  
$str =~ s/ma/miała/;  
print $str;
```

z takim samym wynikiem.

Zawsze wygląda do tak:

funkcja/regex/zamiennik/modyfikator;

Najczęściej stosowane są funkcje:

s – zamień (substitute)

m – dopasuj (match) – np. `if ( $str =~ m/ma/ ) { ...`

Często stosowane modyfikatory:

g – global – zamienia wszystkie wystąpienia, nie tylko pierwsze

i – ignore case – ignoruje małe/wielkie litery.

x – eXclude whitespace – wyklucza białe znaki chyba, że są escape'owane.

Modyfikatory można łączyć np. `s/ma/miała/ig`

Zadanie

Napisz skrypt, który wyświetli ilość użytkowników nieużywających powłoki `tcsh`. Do sprawdzenia użyj wyrażeń regularnych.