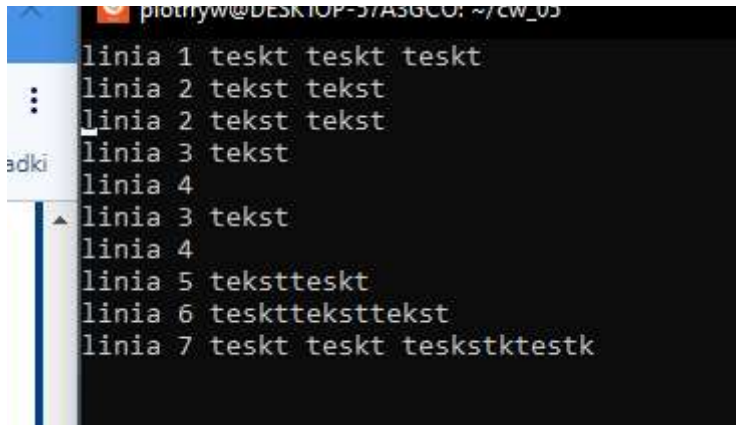


## Sprawozdanie ćwiczenia 5

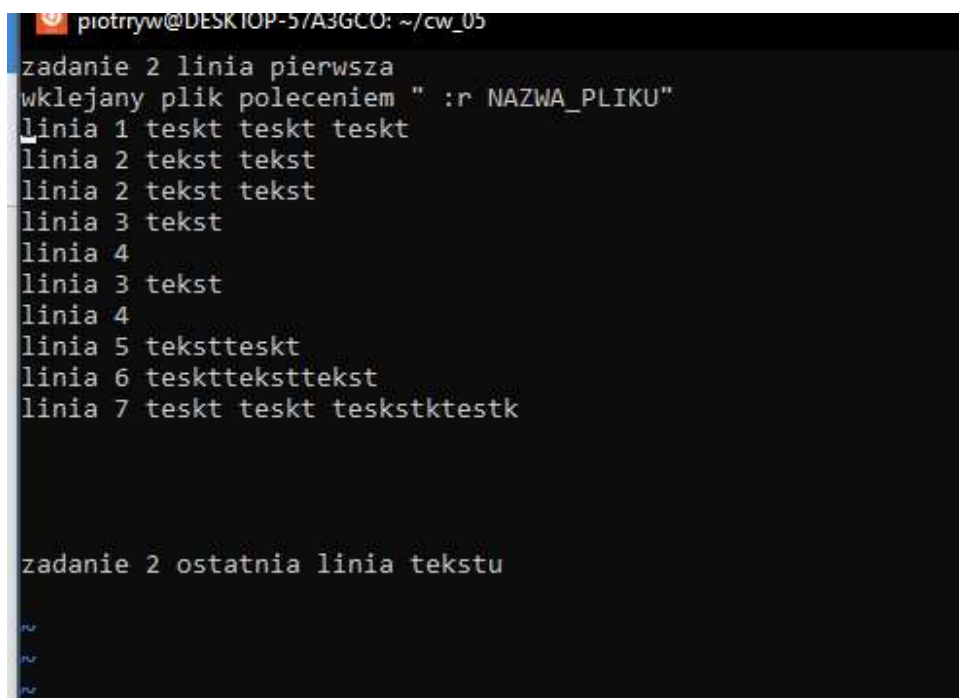
Piotr Rywczak

1. Przetestować działanie komend kopiujących i wycinających całe linie tekstu bądź pojedyncze słowa. Zademonstrować prowadzącemu wycięcie i wklejenie w innym miejscu 3 linii tekstu poniżej kursora na raz.



```
piotryw@DESKTOP-57A3GCO: ~/cw_05
linia 1 tekst tekst tekst
linia 2 tekst tekst
linia 2 tekst tekst
linia 3 tekst
linia 4
linia 3 tekst
linia 4
linia 5 teksttekst
linia 6 tekstteksttekst
linia 7 tekst tekst tekststktestk
```

2. Stworzyć plik tekstowy (dowolny, wpisać przykładowe parę słów, zapamiętać nazwę). Następnie otworzyć nowy plik tekstowy i przetestować działanie komendy :r.



```
piotryw@DESKTOP-57A3GCO: ~/cw_05
zadanie 2 linia pierwsza
wklejany plik poleceniem " :r NAZWA_PLIKU"
linia 1 tekst tekst tekst
linia 2 tekst tekst
linia 2 tekst tekst
linia 3 tekst
linia 4
linia 3 tekst
linia 4
linia 5 teksttekst
linia 6 tekstteksttekst
linia 7 tekst tekst tekststktestk

zadanie 2 ostatnia linia tekstu
```

3. Proszę przeanalizować powyższą komendę (za co odpowiada który fragment).

`:%s/stary/nowy/g`

%s – szukanie we wszystkich liniach dokumentu; g – zamiana wszystkich znalezionych wystąpień; stary – miejsce gdzie się wpisuje szukane słowo; nowy – miejsce gdzie się wpisuje to na co zamieniamy znalezione słowo.

4. Dokonać zamiany całej kolumny tekstu, z ':' na '--' w pliku:

```

piotryw@DESKTOP-57A3GCO: ~/
0:fwefwef
1:grggrerhsh
2:f23f2gagasg
3:fdqf qwfqwf
4:ff wegwgwg wgw
5: fqfqef qwegqwegwg wg

```

```

piotryw@DESKTOP-57A3GCO: ~/cw
0--fwefwef
1--grggrerhsh
2--f23f2gagasg
3--fdqf qwfqwf
4--ff wegwgwg wgw
5-- fqfqef qwegqwegwg wg

```

```
:%s/./--/g
```

5. Zaproponować fragment kodu dla pliku konfiguracyjnego (~/.vimrc), którym uruchomi się podświetlanie składni różnych języków programowania (syntax), wyłączy automatyczne wcięcia (autoindent) i ustawi listę poziomów cofania operacji na 50.

podświetlanie składni różnych języków programowania (syntax)

```
:set syntax = python
```

wyłączy automatyczne wcięcia (autoindent)

```
:set noautoindent
```

ustawi listę poziomów cofania operacji na 50

```
:set undolevels=50
```

6. Które z poniższych ciągów tekstowych nie są reprezentowane przez powyższe wyrażenie regularne?

- `[a-zA-Z0-9._- ]+@[a-zA-Z0-9._- ]\.[a-zA-Z]{2,4}`

1. [@nowhere.com](mailto:@nowhere.com)

2. [linux.bioinfo@gmail.com](mailto:linux.bioinfo@gmail.com)
3. [tux@free\\_penguins.org](mailto:tux@free_penguins.org)
4. [forza@milan.italy](mailto:forza@milan.italy)

Nie są 3 (bo \_ w części po znaku @) oraz 4 (bo italy ma więcej niż 4 znaki)

7. Jak należy zmodyfikować komendę aby dodatkowo wyświetliła linię zawierającą opis sekwencji?

```
grep '^..AG..A' oligos.fasta -B 1
```

8. Jakie polecenie należy wydać aby określić liczbę sekwencji zapisanych w pliku oligos.fasta.

```
grep '^..AG..A' oligos.fasta -c
```

9. W oparciu o powyższe informacje proszę spróbować przewidzieć (zawsze można poeksperymentować na stworzonych przez siebie plikach) efekt działania komendy:

- `sed -e 's/^[ \t]*//'` infile

Powoduje usunięcie wszystkich tabulatorów z pliku

```

otrryw@DESKTOP-57A3GCO:~/cw_05$ cat testowy2
aaaaaaaaaaaa
    aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa
    aaaaaaaaaaaaaa

otrryw@DESKTOP-57A3GCO:~/cw_05$ sed -e 's/^[ \t]*//' testowy2
aaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa
aaaaaaaaaaaaaa

```

10. W oparciu o przedstawione wcześniej informacje proszę zinterpretować wydaną komendę.

```
sed -e '/^>!/ s/T/U/g' s100b.fasta
```

Polecenie powoduje zamianę wszystkich T na U, pomijając pierwszą linię, bo ona zaczyna się od znaku “ > “.

11. Proszę przenieść powyższy skrypt do pliku, wykonać go na wcześniej pobranym pliku oligos.fasta oraz przedstawić wyniki.

```
piotr@DESKTOP-57A3GCO:~/cw_05$ awk -f input-file oligos.fasta
Seq: 1 Length: 40
Seq: 2 Length: 10
Seq: 3 Length: 16
Seq: 4 Length: 26
Seq: 5 Length: 47
Seq: 6 Length: 18
Seq: 7 Length: 15
Seq: 8 Length: 42
Seq: 9 Length: 42
Seq: 10 Length: 49
Seq: 11 Length: 9
Seq: 12 Length: 12
Seq: 13 Length: 42
Seq: 14 Length: 29
Seq: 15 Length: 33
Seq: 16 Length: 18
Seq: 17 Length: 30
Seq: 18 Length: 37
Seq: 19 Length: 35
Seq: 20 Length: 29
```

```
Seq: 151 Length: 49
Seq: 152 Length: 12
Seq: 153 Length: 16
Seq: 154 Length: 47
Seq: 155 Length: 22
Seq: 156 Length: 40
Seq: 157 Length: 23
Seq: 158 Length: 48
Seq: 159 Length: 43
Seq: 160 Length: 19
Seq: 161 Length: 12
Seq: 162 Length: 23
Seq: 163 Length: 28
Seq: 164 Length: 45

Total length of 164 sequences is 4688 nucleotides
```

Program zliczył długość wszystkich linii zaczynających się od A/G/C/T i dodał je do siebie.

12. Proszę przy pomocy programu grep zidentyfikować w pliku 1adb.pdb numery atomów odpowiadające dowolnym dwóm jonom cynku (znajdujące się we wpisach zawierających dane o

położeniu heteroatomów - znacznik HETATM) a następnie wykonać poniższy skrypt (w pierwszej linii należy zmiennym a oraz b przypisać odpowiednie wartości znalezione przy pomocy grep) i podać wynik

```
HETATM 6786 ZN      ZN A 375      -5.233 -10.509 -24.793  1.00 14.21      ZN
HETATM 6787 ZN      ZN A 376     -23.259  -2.638 -28.043  1.00 18.96      ZN
HETATM 6845 ZN      ZN B 375     -10.146   0.818 -67.555  1.00 16.12      ZN
HETATM 6846 ZN      ZN B 376     -5.191 -18.236 -65.594  1.00 16.63      ZN
```

```
piotr@DESKTOP-57A3GCO:~/cw_05$ awk -f skrypt.1 1adb.pdb
19.9362
```

13. Plik data zawiera wyniki stu pomiarów pewnych trzech wielkości. Proszę przygotować skrypt, który będzie wyliczał średnie wartości każdej z mierzonych wielkości. Dane wewnątrz pliku uporządkowane są w kolumnach: pierwsza z nich odpowiada kolejnemu numerowi pomiaru, trzy następne zawierają zmierzone wartości. Należy także wziąć pod uwagę fakt, iż w pliku występują puste linie oraz komentarze zawarte w liniach zaczynających się od znaku #. Wszystkie numery pomiarów zapisane są od pierwszego znaku każdej odpowiedniej linii.

Nie mogłem pobrać pliku dat ze strony z instrukcją