

Algorytmy Tekstowe

Laboratorium 4 - Odległość edycyjna

28.04.2021

Magdalena Pastuła

Zadanie 3.

Wynik wizualizacji dla pary 'los'-'kloc':

```
'los '  
'*k*los '  
'klo*c* '  
'kloc '
```

Wynik wizualizacji dla pary 'Lodz'-'Łódź':

```
'Lodz '  
'*Ł*odz '  
'Ł*ó*dz '  
'Łód*ź* '  
'Łódź '
```

Wynik wizualizacji dla pary 'quintessence'-'kwintesencja':

```
'quintessence '  
'*k*uintessence '  
'k*w*intessence '  
'kwintes**ence '  
'kwintesenc*j*e '  
'kwintesencj*a* '  
'kwintesencja '
```

Wynik wizualizacji dla pary 'ATGAGGCTCTGGCCCCTG'-'ATGAATCTTACCGCCTCG':

```
'ATGAGGCTCTGGCCCCTG '  
'ATGA*A*GCTCTGGCCCCTG '  
'ATGAA*T*CTCTGGCCCCTG '  
'ATGAATCT**TGGCCCCTG '  
'ATGAATCTT*A*GCCCCTG '  
'ATGAATCTTA*C*CCCCTG '  
'ATGAATCTTACC*G*CCTG '
```

```
ATGAATCTTACCGCCT*C*G'
'ATGAATCTTACCGCCTCG'
```

Każda wizualizacja zawiera na początku słowo wejściowe i na końcu słowo wyjściowe. Usunięcie znaku zaznaczane jest poprzez '*'. Dodanie lub modyfikacja znaku są zaznaczane jako '*znak*', gdzie znak to litera dodana lub podstawiona na dane miejsce.

Zadanie 4.

Do obliczania najdłuższego wspólnego podciągu wykorzystano algorytm znajdowania sekwencji edycji jednego łańcucha w drugi. Po uzyskaniu tablicy takich sekwencji następuje iteracja po niej i w przypadku braku akcji dany znak jest dodawany do części wspólnej, w pozostałych przypadkach nie ma żadnej akcji.

Zadanie 7.

Otrzymana długość najdłuższego podciągu dwóch plików z usuniętą częścią tokenów wyniosła 2485, między oryginalnym tekstem i pierwszym zmodyfikowanym wyniosła 2597, a między oryginalnym i drugim zmodyfikowanym 2582.

Zadanie 8.

Napisane narzędzie najpierw oblicza najdłuższy wspólny podciąg, następnie porównuje oba teksty z tym podciągiem i przypadku tokena, który się nie zgadza ze wspólnym podciągiem jest dodawany do dotychczasowych różnic z oznaczeniem '<<<<<<<<' jeśli różnica jest z pierwszego pliku i '>>>>>>>>' jeśli różnica jest z drugiego pliku, a do tego no koniec numer linii z danego pliku, w której ta zmiana występuje.

Zadanie 9.

Otrzymane wyniki po porównaniu dwóch plików z częścią usuniętych tokenów:

```
<<<<<<<<< Paszkowski, line 4
>>>>>>>>> PaszkowskiISBN, line 4
<<<<<<<<<

, line 4
<<<<<<<<< ISBN, line 6
<<<<<<<<< 29039, line 6
>>>>>>>>> 2903, line 4
>>>>>>>>> -, line 4
>>>>>>>>> 9, line 4
<<<<<<<<< ESKALUS, line 11
>>>>>>>>> młody, line 10
<<<<<<<<< dwóch, line 13
>>>>>>>>>

, line 11
>>>>>>>>> TYBALT, line 16
<<<<<<<<< *, line 19
<<<<<<<<< *, line 21
```

```
<<<<<<<<< GRZEGORZ, line 21
>>>>>>>> ABRAHAM, line 21
>>>>>>>> PAŹ, line 24
<<<<<<<<<
, line 25
<<<<<<<<< się, line 32
<<<<<<<<<

, line 32
<<<<<<<<< Kasprowicz, line 43
>>>>>>>> KasprowiczDwa, line 36
<<<<<<<<<

, line 43
<<<<<<<<< Dwa, line 45
<<<<<<<<< -, line 45
>>>>>>>> łon, line 41
<<<<<<<<< bowiem, line 50
<<<<<<<<< przygód, line 52
>>>>>>>> zgon, line 48
>>>>>>>> Wchodzą, line 63
>>>>>>>> Dalipan, line 68
>>>>>>>> ., line 68
>>>>>>>> Ma, line 73
<<<<<<<<< byli, line 82
<<<<<<<<< namibędzie, line 92
>>>>>>>> nami, line 83
>>>>>>>> ,, line 83
>>>>>>>> będzie, line 83
<<<<<<<<< SAMSON, line 95
>>>>>>>> SAMSONMam, line 86
<<<<<<<<<

, line 95
<<<<<<<<< Mam, line 97
<<<<<<<<< zwykłeś, line 102
>>>>>>>> rozruchać, line 96
>>>>>>>> co, line 101
<<<<<<<<< twego, line 112
>>>>>>>> mogą, line 106
<<<<<<<<< pokazuje, line 122
>>>>>>>> się, line 111
<<<<<<<<< najsłabszetulą, line 127
>>>>>>>> najsłabsze, line 116
>>>>>>>> ,, line 116
>>>>>>>> tulą, line 116
>>>>>>>> tylko, line 121
<<<<<<<<< ., line 137
>>>>>>>> kobiet, line 131
<<<<<<<<<
```

```
, line 142
>>>>>>>>
```

```
, line 131
>>>>>>>> SAMSON, line 134
>>>>>>>>
```

```
, line 134
<<<<<<<<< liczysz, line 152
<<<<<<<<< zwierząt, line 152
>>>>>>>> ich, line 148
<<<<<<<<< przechodząc, line 184
>>>>>>>> przechodzącniech, line 173
<<<<<<<<< ;, line 184
<<<<<<<<< niech, line 184
<<<<<<<<< chcą, line 184
<<<<<<<<< ale, line 189
<<<<<<<<< im, line 189
<<<<<<<<<
```

```
, line 189
>>>>>>>> mości, line 190
>>>>>>>> jest, line 197
<<<<<<<<<
```

```
, line 211
>>>>>>>>
```

```
, line 197
<<<<<<<<< GRZEGORZ, line 214
<<<<<<<<<
```

```
, line 214
<<<<<<<<< GRZEGORZ, line 224
>>>>>>>> GRZEGORZ/, line 210
<<<<<<<<<
```

```
, line 224
<<<<<<<<< /, line 226
>>>>>>>> ., line 232
<<<<<<<<< Powiedz, line 257
<<<<<<<<<
```

```
, line 267
```

```
<<<<<<<<< pamiętaj, line 272
<<<<<<<<< wiecie, line 277
<<<<<<<<< ich, line 279
<<<<<<<<< mieczem, line 279
>>>>>>>> mieczem/, line 260
<<<<<<<<< ., line 279
<<<<<<<<< /, line 279
>>>>>>>> mną, line 274
>>>>>>>> Z, line 279
>>>>>>>> Tego, line 280
<<<<<<<<< tchórze, line 301
>>>>>>>> tchórze./, line 282
<<<<<<<<< ., line 301
<<<<<<<<<

, line 301
<<<<<<<<< /, line 303
<<<<<<<<< pałkami, line 303
>>>>>>>> pałkami/, line 282
<<<<<<<<< ., line 303
<<<<<<<<< /, line 303
>>>>>>>> Hola, line 287
<<<<<<<<< z, line 309
>>>>>>>> nadchodzi, line 306
>>>>>>>> ., line 306
<<<<<<<<< Monteki, line 330
>>>>>>>> Monteki/, line 309
<<<<<<<<< ., line 330
<<<<<<<<< /, line 330
<<<<<<<<< krok, line 344
>>>>>>>> krokgdy, line 323
<<<<<<<<< ,, line 344
<<<<<<<<< gdy, line 344
<<<<<<<<< z, line 346
>>>>>>>> natychmiast, line 335
<<<<<<<<< tę, line 357
>>>>>>>> oraz, line 340
>>>>>>>> wiekiem, line 342
>>>>>>>> zasługą, line 342
>>>>>>>> /, line 358
<<<<<<<<< obywatel, line 379
>>>>>>>> zwadę, line 363
<<<<<<<<< się, line 385
<<<<<<<<< BENWOLIO, line 388
>>>>>>>> BENWOLIONieprzyjaciela, line 367
<<<<<<<<<

, line 388
<<<<<<<<< Nieprzyjaciela, line 390
>>>>>>>> w, line 371
<<<<<<<<< zbiegł, line 398
<<<<<<<<<
```

```
, line 400
<<<<<<<<<<

, line 406
>>>>>>>>>

, line 380
<<<<<<<<< Godziną, line 411
>>>>>>>>> BENWOLIO, line 383
>>>>>>>>>

, line 383
>>>>>>>>> z, line 387
<<<<<<<<< tak, line 416
>>>>>>>>> Ledwiem, line 391
>>>>>>>>> natychmiast, line 392
>>>>>>>>> ukrył, line 393
>>>>>>>>> ), line 396
>>>>>>>>> chętnie, line 398
>>>>>>>>> unikał, line 399
<<<<<<<<<

, line 425
<<<<<<<<< chmurami, line 432
<<<<<<<<< wschodzie, line 433
<<<<<<<<< kotarę, line 435
>>>>>>>>> uciekając, line 407
>>>>>>>>> zajdzie, line 411
>>>>>>>>> ,, line 432
<<<<<<<<< daleki, line 464
>>>>>>>>> dalekiJak, line 435
<<<<<<<<<

, line 464
<<<<<<<<< Jak, line 465
<<<<<<<<< /, line 471
>>>>>>>>> /BENWOLIO, line 441
<<<<<<<<<

, line 471
<<<<<<<<< BENWOLIO, line 474
>>>>>>>>> piersi, line 444
>>>>>>>>> rani, line 449
<<<<<<<<< ROMEOJeszcze, line 493
>>>>>>>>> ROMEO, line 460
>>>>>>>>>

, line 460
>>>>>>>>> Jeszcze, line 462
```

```
<<<<<<<<< jest, line 510
<<<<<<<<<

, line 510
<<<<<<<<< tobrak, line 530
>>>>>>>> to, line 496
>>>>>>>> ?, line 496
>>>>>>>> brak, line 496
<<<<<<<<< tyranką, line 541
<<<<<<<<< NiestetyCzemuz, line 546
>>>>>>>> Niestety, line 512
>>>>>>>> !, line 512
>>>>>>>> Czemuż, line 512
>>>>>>>> na, line 513
>>>>>>>> swój, line 513
<<<<<<<<< dziś, line 548
>>>>>>>> Gdzież, line 514
>>>>>>>> Był, line 514
>>>>>>>> lecz, line 516
<<<<<<<<< OdpychającaPoważna, line 554
>>>>>>>> !, line 520
>>>>>>>> Poważna, line 520
<<<<<<<<< !, line 554
>>>>>>>> snu, line 523
<<<<<<<<< ?, line 559
<<<<<<<<< ?, line 581
>>>>>>>> twoje, line 550
<<<<<<<<< cierpieniem, line 584
>>>>>>>> cierpieniemNie, line 550
<<<<<<<<<
, line 584
>>>>>>>> ulgąale, line 550
<<<<<<<<< Nie, line 585
<<<<<<<<< ulgą, line 585
<<<<<<<<< ,, line 585
<<<<<<<<< ale, line 585
>>>>>>>> co, line 553
<<<<<<<<< ,, line 590
>>>>>>>> trawiącą, line 556
>>>>>>>> ., line 556
<<<<<<<<<

, line 600
<<<<<<<<< ROMEO, line 603
<<<<<<<<< nie, line 606
<<<<<<<<< Romeo, line 606
>>>>>>>> kochasz, line 573
>>>>>>>> ., line 578
<<<<<<<<< Pisać, line 630
<<<<<<<<< tego, line 631
>>>>>>>> tegokto, line 593
<<<<<<<<< ,, line 631
```

```

<<<<<<<<< kto, line 631
<<<<<<<<<

, line 632
>>>>>>>> piękny, line 608
>>>>>>>> twardą, line 615
<<<<<<<<< biedna, line 661
<<<<<<<<< zstąpi, line 662
>>>>>>>> zstąpiCałe, line 621
<<<<<<<<<
, line 662
<<<<<<<<< Całe, line 663
>>>>>>>> Wiecznie, line 626
>>>>>>>> strawia, line 632
<<<<<<<<< ona, line 676
>>>>>>>> zarazem, line 634
<<<<<<<<< pięknastąd, line 677
>>>>>>>> piękna, line 635
>>>>>>>> :, line 635
>>>>>>>> stąd, line 635
>>>>>>>> jest, line 635
>>>>>>>> wieczne, line 637
<<<<<<<<< Mógł, line 690
<<<<<<<<<

, line 690
>>>>>>>>

, line 648
>>>>>>>> BENWOLIO, line 651
>>>>>>>>

, line 651

```

Puste linie oznaczają znak nowej linii. Dla porównania poniżej znajdują się pierwszych 30 linii każdego z plików.

Plik pierwszy:

William Shakespeare

Romeo i Julia

tłum. Józef Paszkowski

ISBN 978-83-288-29039

OSOBY:

- * ESKALUS – księżę panujący w Weronie
- * PARYS – Weronieńczyk szlacheckiego rodu, krewny księcia
- * MONTEKI, KAPULET – naczelnicy dwóch domów nieprzyjacielskich sobie* STARZEC – stryjeczny brat Kapuleta
- * ROMEO – syn Montekiego
- * MERKUCJO – krewny księcia
- * BENWOLIO – synowiec Montekiego
- * – krewny Pani Kapulet
- * LAURENTY – ojciec franciszkanin
- * JAN – brat z tegoż zgromadzenia
- * BALTAZAR – służący Romea
- * SAMSON, GRZEGORZ – słudzy Kapuleta
- * – służący Montekiego
- * APTEKARZ
- * TRZECH MUZYKANTÓW
- * PARYSA
- * PIOTR
- * DOWÓDCA WARTY
- * PANI MONTEKI – małżonka Montekiego
- * PANI KAPULET – małżonka Kapuleta
- * JULIA – córka Kapuletów

Plik drugi:

William Shakespeare

Romeo i Julia

tłum. Józef Paszkowski ISBN 978-83-288-2903-9

OSOBY:

- * – księżę panujący w Weronie
- * PARYS – młody Weronieńczyk szlacheckiego rodu, krewny księcia
- * MONTEKI, KAPULET – naczelnicy domów nieprzyjacielskich sobie
- * STARZEC – stryjeczny brat Kapuleta
- * ROMEO – syn Montekiego
- * MERKUCJO – krewny księcia
- * BENWOLIO – synowiec Montekiego
- * TYBALT – krewny Pani Kapulet
- * LAURENTY – ojciec franciszkanin
- JAN – brat z tegoż zgromadzenia
- * BALTAZAR – służący Romea
- SAMSON, – słudzy Kapuleta
- * ABRAHAM – służący Montekiego
- * APTEKARZ
- * TRZECH MUZYKANTÓW
- * PAŹ PARYSA* PIOTR
- * DOWÓDCA WARTY

- * PANI MONTEKI – małżonka Montekiego
- * PANI KAPULET – małżonka Kapuleta
- * JULIA – córka Kapuletów
- * MARTA – mamka Julii
- * Obywatele weroneńscy, różne osoby płci obojej, liczący do przyjaciół obu domów, maski, straż wojskowa i inne osoby. Rzecz odbywa się przez większą część sztuki w Weronie, przez część piątego aktu w Mantui.