

ZAAWANSOWANE TECHNOLOGIE JAVY

WZORZEC W TEKŚCIE

Instytut Informatyki Uniwersytetu Wrocławskiego

Paweł Rzechonek

Termin *bioinformatyka* pojawił się po raz pierwszy w 1970 roku i odnosił się do badania procesów informatycznych w systemach biotycznych. Jego twórcami są Paulien Hogeweg i Ben Hesper. W ten sposób wyróżnili oni bioinformatykę jako osobną dziedzinę nauki obok biochemii i biofizyki.

Z bioinformatyką związana jest ściśle biologia molekularna. Na rozwój tej drugiej dziedziny miały wpływ przede wszystkim dwa dokonania w latach 50 XX wieku. Pierwszym z nich było podanie w 1953 roku przez J. Watsona i F. Cricka modelu podwójnej helisy łańcucha DNA przechowującej informację genetyczną we wszystkich organizmach żywych. Drugim było określenie przez F. Sangera w 1952 roku całkowitej sekwencji aminokwasów insuliny. Eksperymenty biologiczne zaczęły generować ogromne ilości danych, które okazały się trudne do przechowywania i ręcznego przetwarzania. Z tego powodu wykorzystanie komputerów i informatyki stało się niezbędne przy pracy biologów.

<http://pl.wikipedia.org/wiki/Bioinformatyka>

W 2001 roku został zsekwencjonowany ludzki genom.

* * *

Napisz w języku C/C++ funkcję, która wylicza ile razy w zadanym tekście występuje określony wzorzec:

```
int sprawdz(const char *tekst, const char *zworzec);
```

Zaimplementuj algorytm Knutha–Morrisa–Pratta:

http://pl.wikipedia.org/wiki/Algorytm_Knutha-Morrisa-Pratta

Potem skompiluj ten kod i stwórz z niego bibliotekę dynamiczną (.so pod Linuxem albo .dll pod Windowsem).

Następnie napisz aplikację okienkową w technologii *Swing*, która będzie testowała działanie wcześniej napisanej i skompilowanej funkcji `sprawdz()` w kodzie rodzimym — technologia JNI. Sama aplikacja powinna zawierać pole tekstowe `JButton` na wzorzec i obszar tekstowy `JTextField` na tekst oraz z przycisk do testowania wystąpienia wzorca w tekście. Do aplikacji możesz dołożyć przycisk, który będzie umożliwiał wczytanie tekstu z pliku.