

## Zadania domowe – Przetwarzanie plików wielojęzycznych

Termin złożenia zadania domowego to 09.04.2008.

### Zadanie 1

Przy wykorzystaniu pragmy `utf8` zmienić program z 14. slajdu z wykładu tak, aby wszystkie polskie znaki występowały w nim w sposób jawny. Tzn. jeśli ma wyświetlić „ę”, to w kodzie ma wystąpić „ę”.

Punkty: 2

### Zadanie 2

Zapisać program ze slajdu 16. z wykładu bez użycia pragmy `utf8` w taki sposób, by nie zawierał żadnego znaku wykraczającego poza standard ASCII a jednocześnie zachował pełną funkcjonalność.

Punkty: 2

### Zadanie 3

Napisać program w Perlu, który konwertuje pliki tekstowe z podanego kodowania źródłowego na wybrane kodowanie docelowe. Program ma działać z poziomu konsoli. Przykładowe wykonanie programu:

```
perl konwertuj.pl -f iso-8859-2 -t utf-8 <plik-latin2.txt >plik-utf8.txt
```

Opcje `-f` i `-t` to odpowiednio skróty od `from` i `to`. W celu obsługi opcji wykorzystać moduł `Getopt::Long`. Dokumentację do tego pakietu znajdują państwo pod <http://perldoc.perl.org/Getopt/Long.html>.

Proszę napisać dwie wersje tego programu. Jedna wersja ma korzystać z warstw kodowania, druga z modułu `Encode`. Która funkcja z pakietu `Encode` wykonuje czynność zamiany dwóch kodowania w jednym kroku?

Punkty: 2 za każdą wersję +1 punkt za poprawną odpowiedź na ostatnie pytanie.

### Zadania 4

Zmodyfikować lub napisać od nowa lematyzator z zadania 3 z zestawu „Perl – Referencje 2”. Tym razem ma w pełni obsługiwać polskie znaki zarówno na poziomie słownika jak i na poziomie przetwarzanego tekstu.

Czy kodowania słownika i przetwarzanego tekstu muszą być takie same? Jak ma się do tej kwestii nasza zasada „czarnej skrzynki”?

Punkty: 3 + 1 za odpowiedź na pytanie.