

Zadanie domowe - Perl: Zastosowanie modułów

Marcin Junczys-Dowmunt

14 marca 2008

Zadanie 1

Zainstalować w domu moduł `Lingua::Identify` zgodnie z instrukcją z wykładu. Na podstawie dokumentacji ograniczyć działanie pakietu do dwóch języków, *polskiego* i *angielskiego*. Żadne inne języki nie mają być aktywne.

Proszę podać prawdopodobieństwa zdań *"Wyborcza ujawnia: Putin rozgrywa NATO"* i *"Chaos in Tibet: Protests in Lhasa become violent"* dla obu języków.

Punkty: 3

Zadanie 2

Wybrać dowolny moduł (inny niż `Lingua::Identify`) z archiwum CPAN, dział "String Language Text Processing", i opisać jego działanie, podstawowe funkcje i zastosowanie do językoznawstwa.

Punkty: 3

Zadanie 3

Napisać mały program wykorzystujący powyżej opisany moduł.

Punkty: 4

Na następnej stronie: połowa rozwiązania wzorcowego zadania 5 z poprzedniego zestawu.

Zadanie 4 – Pierwsza połowa wzorcowego rozwiązania Zd. 5 PerlRef2

```
1 use strict;
  use Data::Dumper;

  my $ngrams = {};
5
  open(NGRAMS, "<ngrams.txt");
  while(<NGRAMS>) {
    my $act = $ngrams;
    my @words = split(/\s+/, $_);
10    foreach my $word (@words) {
      if(exists($act->{$word})) {
        $act->{$word} = {};
      }
      $act = $act->{$word};
15    }
    $act->{"#STOP#"} = 1;
  }

  print STDERR Dumper($ngrams);
```

Proszę dokończyć program.

Punkty: 6