

Zadania domowe – Referencje 2

Gotowe zadanie domowe należy przesłać na adres junczys@amu.edu.pl. Należy dołączyć wszystkie pliki potrzebne do poprawnego wyświetlenia wraz z krótkim **komentarzem** do każdego zadania.

Proszę umieścić w mailu Imię, Nazwisko, Specjalizację i Rok. Termin złożenia zadania domowego to 04.03.2008.

Zadanie 1

Proszę napisać program, który wyświetla tylko niepowtarzające się wiersze w takiej kolejności w jakiej zostały wczytane oraz liczbę wystąpień danego wiersza. Opisać wykorzystaną strukturę danych i uzasadnić jej wybór.

Wskazówka: Można rozszerzyć jeden z programów wzorcowych do zestawu PERL 4. Zastanowić się najpierw, jakie informacje trzeba skojarzyć z pojedynczym wierszem i który z programów wzorcowych się do tego najlepiej nadaje (tzn. który działa już teraz na najbardziej zbliżonej strukturze danych, która zawiera podobne dane).

Punkty: 3

Zadanie 2

Przy wykorzystaniu poprzedniego zadania domowego z referencji, rozszerzyć program do odmiany niemieckich przymiotników tak, aby pobierał rdzenia przymiotników nieregularnych z pliku tekstowego. Jeśli podamy przymiotnik, który znajduje się w tym słowniku form nieregularnych, to ma go odmienić na podstawie tych form. W przypadku gdy przymiotnik jest regularny, odmiana powinna się odbyć w sposób tradycyjny.

Punkty: 5

Zadanie 3

Rozszerzyć program ze skryptu służący do lematyzowania wyrazów tak, aby wczytywał słownik z pliku tekstowego. Opisać układ danych w pliku tekstowym. Wybór języka jak i zestawu słowoform oraz lematów jest dowolny. Można oczywiście wybrać język inny niż polski.

Uwaga: Tutaj lepiej wczytywać dane o frazach z ustalonego pliku z pomocą funkcji **open** np. *sownik.txt* a nie z wejścia standardowego. Z wejścia standardowego wczytujemy później tekst, który ma być lematyzowany.

Punkty: 3

Rozszerzyć program tak, aby uwzględniał dodatkowe informacje jak np. przypadek, liczba lub inne cechy odpowiednie dla danej części mowy, tzn. że dla wyrazu „kobietom” powinien zwrócić informacje typu „kobieta_R:cel:mn:z” (Rzeczownik w celowniku, liczba mnoga, rodzaj żeński). Jak zmieni się program? Jak struktura słownika?

Punkty: 2

Zadanie 4

Zmodyfikować program ze skryptu służący do wyszukiwania bi-gramów tak, aby wczytywał listę bi-gramów z pliku tekstowego i zbudował strukturę służącą do wyszukiwania na podstawie tych danych.

Uwaga: Tutaj lepiej wczytywać dane o frazach z ustalonego pliku z pomocą funkcji **open** np. *bigramy.txt* a nie z wejścia standardowego. Z wejścia standardowego wczytujemy później tekst, który ma być przeszukany.

Punkty: 3

Zadania 5 (trudne!)

Rozszerzyć powyższy program tak, aby działał dla fraz o dowolnej długości. Tzn. dane w pliku tekstowym określają głębokość struktury, nie kod programu.

Problem jest nietrywialny i proszę nie załamywać się, jeśli się nie uda. Program można podzielić na dwie części: a) Budowę odpowiedniej struktury danych na podstawie listy wyrazów o różnych długościach, b) Poruszanie się po strukturze w celu sprawdzenia, czy dany ciąg wyrazów należy do szukanych.

Proszę się nie krępować i przysłać też ewentualne próby. Może będą punkty za pomysł.

Punkty za cały program: 10