
Zadania z podstaw OO oraz kolekcji w Javie

1. Napisz program, który dla pliku tekstowego o podanej nazwie wyznaczy statystyki dotyczące wystąpień słów:

- 20 najczęstszych słów (słowo - liczba wystąpień), np:

```
and - 6109
the - 4867
...
```

- 20 najrzadszych słów, np:

```
bricked - 1
...
```

- 20 najdłuższych słów,
- wszystkie słowa, które są palindromami.

Uwaga – za słowo przyjmujemy tylko ciągi złożone z liter, cyfr, myślnika oraz podkreślenia. Jeżeli ciąg zawiera inny symbol nie jest liczony jako słowo. Do testów można wykorzystać, np. „Adventures of Huckleberry Finn” M. Twaina dostępnego pod adresem:

<http://www.gutenberg.org/dirs/7/76/76.txt>

Wskazówka – do wygodnego zliczania słów warto użyć

```
java.util.HashMap<String, Integer>
```

2. Na podstawie częstości słów z Zad. 1 wygenerować „chmurę słów” 400 najczęściej używanych słów. Wielkość czcionki słowa powinna być proporcjonalna do liczby wystąpień, przy czym minimalna wielkość czcionki to 20 pikseli, a maksymalna to 70 pikseli.

Rysowanie można zrealizować generując stronę HTML o następującym szablonie:

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
  </head>
```

```
<body>
  <-- Tutaj znaczniki dla słów -->
</body>
</html>
```

Dla każdego należy wygenerować ciąg postaci:

```
<div style="position:absolute;
  left: 12px; top: 45px;
  font-size: 30px">Słowo</div>
```

gdzie

- `left: 12px` oznacza współrzędną x ,
- `top: 45px` oznacza współrzędną y ,
- `font-size: 30px` to wielkość czcionki.

Współrzędne (x, y) słów należy wygenerować losowo z zakresu $[0, 800]$, co da prostokątny obszar. Można też wygenerować je losowo w obrębie koła o promieniu $r = 400$ oraz środka w punkcie $(400, 400)$.

Przykładowa chmura dla tekstu „Adventures of Huckleberry Finn” M. Twaina pokazana jest na rys. 1.1.

