

ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ
Кафедра программного обеспечения информационных технологий

Факультет НиДО
Специальность ПОИТ

Индивидуальная практическая работа № 1
по дисциплине «Информационное обеспечение финансовых структур»

Выполнил студент: Михаленя А.Н.
группа 801021
Зачетная книжка № 801021-23

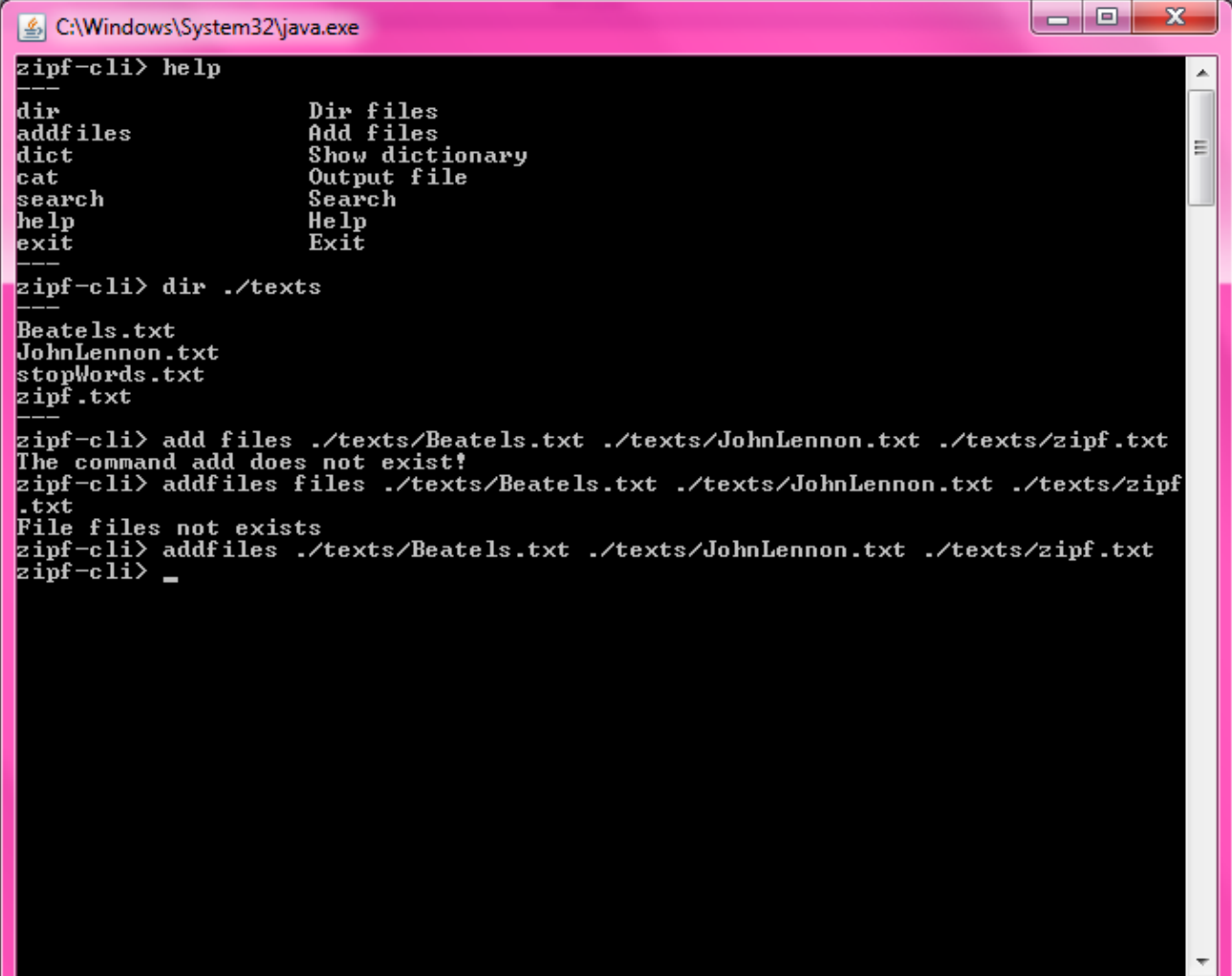
Минск 2013

ПРИМЕНЕНИЕ ЗАКОНОВ ЗИПФА ДЛЯ ОБРАБОТКИ ТЕКСТОВ И ПОИСК ТЕКСТОВОЙ ИНФОРМАЦИИ ПО ДОКУМЕНТУ-ОБРАЗЦУ

Цели работы:

1. Ознакомиться с законами Зипфа и научиться их применять для определения заданных характеристик текстовой информации.
2. Выполнить поиск текстовой информации по документу-образцу на основе строки ключевых слов.

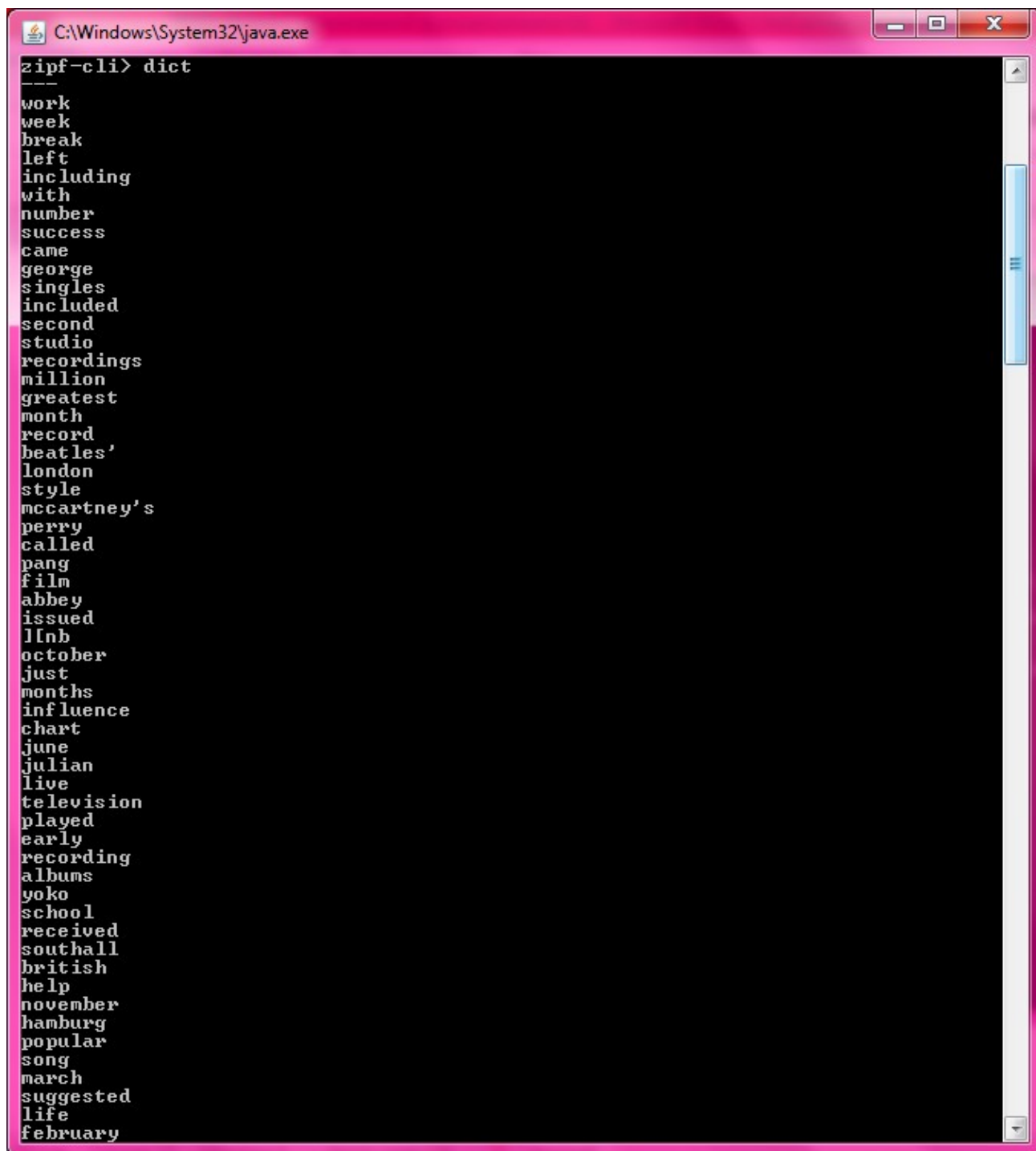
Для проверки законов Зипфа было разработано CLI приложение.
Для вывода списка команд на экран можно воспользоваться командой help.
Для вывода списка txt файлов в указанной директории(путь указывается относительный) можно воспользоваться командой dir (по умолчанию выводится список в текущей директории).
Для добавления файлов в коллекцию необходимо воспользоваться командой addfiles (через пробел указать пути к файлам). В случае если файла не существует – выведется соответствующее сообщение. В случае если будет введена ошибочная команда или не указан какой-либо параметр – система укажет ошибки. Стоп слова расположены в файле (`__ProgrammPath__`)/texts/stopWords.txt



```
C:\Windows\System32\java.exe
zipf-cli> help
---
dir                Dir files
addfiles           Add files
dict               Show dictionary
cat                Output file
search             Search
help               Help
exit               Exit
---
zipf-cli> dir ./texts
---
Beatels.txt
JohnLennon.txt
stopWords.txt
zipf.txt
---
zipf-cli> add files ./texts/Beatels.txt ./texts/JohnLennon.txt ./texts/zipf.txt
The command add does not exist!
zipf-cli> addfiles files ./texts/Beatels.txt ./texts/JohnLennon.txt ./texts/zipf
.txt
File files not exists
zipf-cli> addfiles ./texts/Beatels.txt ./texts/JohnLennon.txt ./texts/zipf.txt
zipf-cli> _
```

Рисунок 1.

Для вывода ключевых слов в результате обработки по законам Зипфа необходимо воспользоваться командой dict



```
C:\Windows\System32\java.exe
zipf-cli> dict
work
week
break
left
including
with
number
success
came
george
singles
included
second
studio
recordings
million
greatest
month
record
beatles'
london
style
mccartney's
perry
called
pang
film
abbey
issued
llnb
october
just
months
influence
chart
june
julian
live
television
played
early
recording
albums
yoko
school
received
southall
british
help
november
hamburg
popular
song
march
suggested
life
february
```

Рисунок 2.

Для поиска информации в документах необходимо воспользоваться командой search. Вывести содержимое файла на экран можно с помощью команды cat. Для выхода из системы введите exit

```
C:\Windows\System32\java.exe
zipf-cli> search zipf
----
./texts/zipf.txt
----
zipf-cli> cat zipf
File zipf not exists
zipf-cli> cat ./texts/zipf.txt
File ./texts/zipf.txt not exists
zipf-cli> cat ./texts/zipf.txt
-----
Wikimedia Commons has media related to: Zipf's law
Steven, Strogatz (2009-05-29). "Guest Column: Math and the City". The New York Times. Retrieved 2009-05-29 -- An article on Zipf's law applied to city populations
Seeing Around Corners (Artificial societies turn up Zipf's law)
PlanetMath article on Zipf's law
Distributions de type ?fractal parabolique? dans la Nature (French, with English summary)
An analysis of income distribution
Zipf List of French words
Zipf list for English, French, Spanish, Italian, Swedish, Icelandic, Latin, Portuguese and Finnish from Gutenberg Project and online calculator to rank words in texts
Citations and the Zipf-Mandelbrot's law
Zipf's Law for U.S. Cities by Fiona Maclachlan, Wolfram Demonstrations
.
Weisstein, Eric W., "Zipf's Law", MathWorld.
Zipf's Law examples and modelling (1985)
Complex systems: Unzipping Zipf's law (2011)
zipf-cli> help
----
dir                Dir files
addfiles           Add files
dict               Show dictionary
cat                Output file
search             Search
help               Help
exit               Exit
----
zipf-cli> exit_
```

Рисунок 3.