Министерство образования Республики Беларусь

Учреждение образования

«Брестский государственный технический университет»

Кафедра ИИТ

Лабораторная работа № 2

По дисциплине ЕЯИИС

Тема: «Разработка системы автоматического реферирования документов»

Выполнил:

Студент 4 курса

Группы ИИ-16

Пилипук З.А.

Проверил:

Слинко Е.В.

Брест, 2021

**Цель работы:** освоить на практике основные принципы автоматического реферирования документов.

**Код программы:**

#include "search-engine.h"

int main()

{

SearchEngine e;

e.SetPath("D:\\c++\\labs\\lab2\\text.txt");

e.Analyze();

return 0;

}

#pragma once

#include <fstream>

#include <map>

#include <string>

#include <vector>

#include <algorithm>

#include <iostream>

struct SearchEngine

{

SearchEngine() = default;

void SetPath(const std::string& path)

{

path\_ = path;

}

void Analyze()

{

std::ifstream file(path\_);

while (!file.eof())

{

std::string line;

getline(file, line);

if (line == "")

{

result\_.push\_back(local\_);

local\_.clear();

continue;

}

auto data = Parse(line);

for (const auto& value : data)

{

if (!CheckWord(value))

{

continue;

}

addValue(local\_, value);

addValue(stats\_, value);

}

}

key\_word\_ = GetMax(stats\_).first;

std::cout << key\_word\_ << "\n";

for (const auto& value : result\_)

{

std::cout << "\t" << GetMaxOther(value).first << "\n";

}

}

private:

std::vector<std::string> Parse(const std::string& input)

{

std::vector<std::string> res;

std::string word;

for (size\_t i = 0; i < input.size(); ++i)

{

if (isalpha(input[i]))

{

word += input[i];

}

if (input[i] == ' ')

{

res.push\_back(word);

word.clear();

}

}

res.push\_back(word);

return res;

}

std::pair<std::string, size\_t> GetMax(const std::map<std::string, size\_t>& stats)

{

size\_t max = 0;

std::pair<std::string, size\_t> local\_max;

for (const auto& value : stats)

{

if (value.second > max)

{

local\_max = value;

max = value.second;

}

}

return local\_max;

}

std::pair<std::string, size\_t> GetMaxOther(const std::map<std::string, size\_t>& stats)

{

size\_t max = 0;

std::pair<std::string, size\_t> local\_max;

for (const auto& value : stats)

{

if (value.first == "" || value.first == key\_word\_)

continue;

if (value.second > max)

{

local\_max = value;

max = value.second;

}

}

return local\_max;

}

void addValue(std::map<std::string, size\_t>& stats, const std::string& value)

{

auto it = std::find\_if(stats.begin(), stats.end(), [&value](std::pair<std::string, size\_t> p)

{ return p.first == value; });

if (it != stats.cend())

{

++it->second;

}

else

{

stats.insert({ value, 1 });

}

}

bool CheckWord(const std::string& word)

{

return word != "and" && word != "a" && word != "the" && word != "of" && word != "is"

&& word != "for" && word != "to" && word != "on" && word != "or" && word != "in"

&& word != "that" && word != "as" && word != "are" && word != "such" && word != "have";

}

std::map<std::string, size\_t> stats\_;

std::map<std::string, size\_t> local\_;

std::vector<std::map<std::string, size\_t>> result\_;

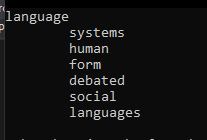
std::string path\_;

std::string key\_word\_;

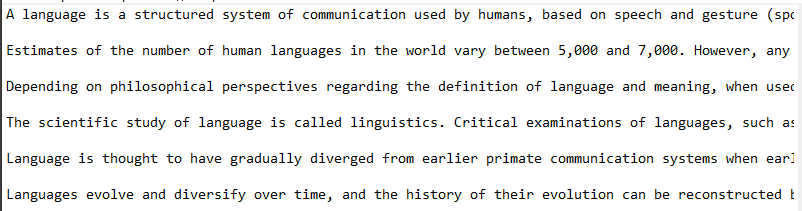
};

**Тестирование:**

Реферат в виде ключевых слов:



Классический реферат:



**Вывод:** освоил на практике основные принципы автоматического реферирования документов.