

Учреждение образования
«Белорусский государственный университет информатики и
радиоэлектроники»
Кафедра интеллектуальных информационных технологий

ОТЧЁТ
по лабораторной работе №6 по дисциплине «ЕЯзИИС»
на тему: «Разработка систем машинного перевода документов»

Выполнили студенты группы 821701:

Поживилко П.С.
Витушко Л. Д.

Проверил:

Крапивин Ю.Б.

Минск, 2021

Цель работы: освоить на практике основные принципы машинного перевода документов.

Вариант 1:

Направление перевода: англо-русский.

Предметная область: научные статьи по computer science, сочинения по литератур

Ход работы.

Сперва формируется базовый словарь, на основании которого будет осуществляться разбор.

```
start_grammar_str = ""S -> NP VP | S P S | S S
VP -> V | V NP | V NP PP | V ADJ | V RB | V VP | P V | V NP | P V PRP | V RB V NP
PP -> P NP
NP -> Det N | Det N PP | N | Det ADJ N | Det ADJ ADJ N | Det ADJ N PP | Det ADJ ADJ N PP | ADJ N PP | ADJ ADJ N PP | PRP | N PP | ADJ N ""
```

Рисунок 1. Словарь синтаксических правил

Далее для каждого предложения из текста данный словарь индивидуально дорабатывается.

Предложение разбивается на слова и для каждого слова ставится в соответствии часть речи, которой оно является в данном предложении (существительное, прилагательное, глагол и т.п.).

```
VP -> V | V NP | V NP PP | V ADJ | V RB | V VP | P V | V NP | P V PRP | V RB V NP
PP -> P NP
NP -> Det N | Det N PP | N | Det ADJ N | Det ADJ ADJ N | Det ADJ N PP | Det ADJ ADJ N PP | ADJ N PP | ADJ ADJ N PP | PRP | N PP | ADJ N
V -> "must" | "touch"
N -> "wires"
RB -> "not"
ADJ -> "opened"
PRP -> "You"
```

Рисунок 2. Дополненный словарь синтаксических правил для конкретного предложения

Основываясь на сформированном словаре, строим синтаксическое дерево зависимостей.

```
Введите текст для перевода:
You must not touch opened wires
Количество слов во входном тексте: 6
Дерево синтаксического разбора:
(S
  (NP (PRP You))
  (VP (V must) (RB not) (V touch) (NP (ADJ opened) (N wires))))
```

Слово	Перевод	Сколько раз встречается	Часть речи	Род	Число
You	ты	1	Местоимение	None	Единственное число
wires	провод	1	Имя существительное	Мужской род	Единственное число
touch	прикасаться	1	Инфинитив	None	None
not	не	1	Частица	None	None
must	должен	1	Имя прилагательное (краткое)	Мужской род	Единственное число

Переведено слов: 5
Перевод текста на русский язык: ты должен не прикасаться провод

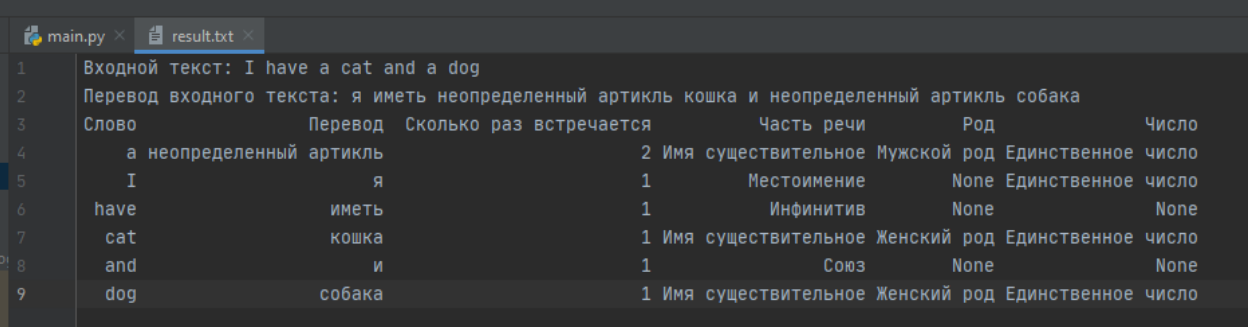
Рисунок 3. Результат работы программы

```
main.py x result.txt x
1 Входной текст: You must not touch opened wires
2 Перевод входного текста: ты должен не прикасаться провод
3 Слово    Перевод    Сколько раз встречается    Часть речи    Род    Число
4 You      ты              1                      Местоимение   None   Единственное число
5 wires    провод          1                      Имя существительное Мужской род   Единственное число
6 touch    прикасаться     1                      Инфинитив     None    None
7 not      не              1                      Частица       None    None
8 must     должен          1                      Имя прилагательное (краткое) Мужской род   Единственное число
9 opened   N/A             1                      N/A           N/A     N/A
```

Рисунок 4. Содержимое файла Result.txt

```
-----Меню-----
1. Перевод фразы
2. Печать файла с результатами
3. Справка
0. Выход
]
Введите текст для перевода:
I have a cat and a dog
Количество слов во входном тексте: 7
Дерево синтаксического разбора:
(S
  (NP (PRP I))
  (VP
    (V have)
    (NP (Det a) (N cat))
    (PP (P and) (NP (Det a) (N dog))))
(S
  (NP (PRP I))
  (VP
    (V have)
    (NP (Det a) (N cat) (PP (P and) (NP (Det a) (N dog))))))
+-----+-----+-----+-----+-----+-----+
| Слово | Перевод | Сколько раз встречается | Часть речи | Род | Число |
+-----+-----+-----+-----+-----+-----+
| a | неопределенный артикль | 2 | Имя существительное | Мужской род | Единственное число |
| I | я | 1 | Местоимение | None | Единственное число |
| have | иметь | 1 | Инфинитив | None | None |
| cat | кошка | 1 | Имя существительное | Женский род | Единственное число |
| and | и | 1 | Союз | None | None |
+-----+-----+-----+-----+-----+-----+
Переведено слов: 7
Перевод текста на русский язык: я иметь неопределенный артикль кошка и неопределенный артикль собака
```

Рисунок 5. Результат работы программы



```
main.py x result.txt x
1 Входной текст: I have a cat and a dog
2 Перевод входного текста: я иметь неопределенный артикль кошка и неопределенный артикль собака
3 Слово Перевод Сколько раз встречается Часть речи Род Число
4 a неопределенный артикль 2 Имя существительное Мужской род Единственное число
5 I я 1 Местоимение None Единственное число
6 have иметь 1 Инфинитив None None
7 cat кошка 1 Имя существительное Женский род Единственное число
8 and и 1 Союз None None
9 dog собака 1 Имя существительное Женский род Единственное число
```

Рисунок 6. Содержимое файла Result.txt