

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО»

Институт компьютерных наук и кибербезопасности
Высшая школа технологий искусственного интеллекта
Направление **02.03.01** : Математика и компьютерные науки

Лабораторная работа «Построение контекстно-свободной
грамматики подмножества естественного языка: Past Perfect
английского языка»
по дисциплине «Теория алгоритмов»

Обучающийся: _____

Яшнова Дарья Михайловна
группа 5130201/20002

Руководитель: _____

Востров Алексей Владимирович

« ____ » _____ 2025г

Санкт-Петербург, 2025

Содержание

| | |
|---|-----------|
| Введение | 4 |
| 1 Описание грамматики Past Perfect | 5 |
| 1.1 Формы предложений | 5 |
| 1.1.1 Утвердительные предложения | 5 |
| 1.1.2 Отрицательные предложения | 5 |
| 1.1.3 Вопросительные предложения | 5 |
| 1.1.4 Отрицательные вопросительные предложения | 5 |
| 2 Математическое описание | 6 |
| 2.1 Грамматика и язык | 6 |
| 2.1.1 Формальное определение | 6 |
| 2.1.2 Алфавиты | 6 |
| 2.1.3 Продукции | 6 |
| 2.1.4 Примеры выводов | 6 |
| 2.2 БНФ задаваемого подмножества языка | 7 |
| 2.2.1 Особенности грамматики | 8 |
| 3 Особенности реализации | 9 |
| 3.1 Структуры данных | 9 |
| 3.1.1 Местоимения (Pronouns) | 9 |
| 3.1.2 Определители (Determiners) | 9 |
| 3.1.3 Прилагательные (Adjectives) | 9 |
| 3.1.4 Существительные (Nouns) | 9 |
| 3.1.5 Имена собственные (Proper Nouns) | 9 |
| 3.1.6 Даты (Dates) | 9 |
| 3.1.7 Глаголы (Verbs) | 9 |
| 3.1.8 Наречия (Adverbs) | 9 |
| 3.1.9 Предлоги и союзы | 10 |
| 3.2 Основные функции | 10 |
| 3.2.1 <code>validate_past_perfect(sentence)</code> | 10 |
| 3.2.2 <code>interactive_validation()</code> | 10 |
| 3.2.3 <code>generate_sentence()</code> | 10 |
| 3.3 Функции генерации частей предложения | 11 |
| 3.3.1 <code>generate_subject()</code> | 11 |
| 3.3.2 <code>generate_noun_phrase(exclude_date=False)</code> | 11 |
| 3.3.3 <code>generate_past_participle()</code> | 11 |
| 3.3.4 <code>generate_object()</code> | 11 |
| 3.3.5 <code>generate_adverbial()</code> | 12 |
| 3.3.6 <code>generate_adverbials()</code> | 12 |
| 3.4 Функции генерации типов предложений | 12 |
| 3.4.1 <code>generate_affirmative()</code> | 12 |
| 3.4.2 <code>generate_negative()</code> | 12 |
| 3.4.3 <code>generate_question()</code> | 13 |
| 3.4.4 <code>generate_simple_sentence()</code> | 13 |
| 3.5 Код программы | 13 |
| 4 Результаты работы программы | 18 |
| Заключение | 19 |

Введение

В данной лабораторной работе необходимо:

1. Реализовать распознавание вводимого предложения на принадлежность грамматике;
2. Реализовать генератор предложений в соответствии с выбранной грамматикой.

В качестве подмножества естественного языка был выбран *past perfect* английского языка.

1 Описание грамматики Past Perfect

Past Perfect (прошедшее совершенное время) используется для выражения действия, которое произошло до другого действия в прошлом или до определённого момента в прошлом. Оно подчёркивает завершённость действия и его связь с последующими событиями.

1.1 Формы предложений

1.1.1 Утвердительные предложения

Образуются по схеме: **Subject + had + V3 (Past Participle) + ...**

| Subject | had | V3 | ... |
|---------|-----|----------|-------------------|
| She | had | finished | her work. |
| They | had | left | before we called. |

1.1.2 Отрицательные предложения

Образуются по схеме: **Subject + had + not + V3 + ...**

| Subject | had | not | V3 | ... |
|---------|-----|-----|---------|------------------------|
| He | had | not | seen | that movie before. |
| We | had | not | visited | Paris until last year. |

1.1.3 Вопросительные предложения

Образуются по схеме: **Had + Subject + V3 + ... ?**

| Had | Subject | V3 | ... |
|-----|---------|----------|---------------|
| Had | you | been | there before? |
| Had | she | finished | her report? |

1.1.4 Отрицательные вопросительные предложения

Образуются по схеме: **Had + Subject + not + V3 + ... ?**

| Had | Subject | not | V3 | ... |
|-----|---------|-----|----------|--------------|
| Had | they | not | received | the message? |

2 Математическое описание

2.1 Грамматика и язык

В данной работе строится *контекстно-свободная грамматика*, задающая язык предложений в Past Perfect английского языка.

2.1.1 Формальное определение

Порождающая грамматика Хомского определяется как четверка $G = \langle T, N, S, R \rangle$, где:

- T - конечное множество терминалов (слова языка).
- N - конечное множество нетерминалов ($T \cap N = \emptyset$).
- $S \in N$ - начальный нетерминал.
- R - множество правил вывода вида $\alpha \rightarrow \beta$.

Для Past Perfect определим следующие компоненты:

2.1.2 Алфавиты

$$T = \{\text{had, not, ?, ., verbs_V3, nouns, pronouns, adverbs}\}$$

$$N = \{\text{S, Affirm, Neg, Quest, Subj, VerbPhrase, TimeAdv}\}$$

2.1.3 Продукции

$$\begin{aligned} S &\rightarrow \text{Affirm} \mid \text{Neg} \mid \text{Quest} \\ \text{Affirm} &\rightarrow \text{Subj had V3 TimeAdv} \\ \text{Neg} &\rightarrow \text{Subj had not V3 TimeAdv} \\ \text{Quest} &\rightarrow \text{Had Subj V3 TimeAdv?} \\ \text{Subj} &\rightarrow \text{nouns} \mid \text{pronouns} \\ \text{V3} &\rightarrow \text{verbs_V3} \\ \text{TimeAdv} &\rightarrow \text{adverbs} \mid \epsilon \end{aligned}$$

2.1.4 Примеры выводов

1. Утвердительное предложение:

$$S \Rightarrow \text{Affirm} \Rightarrow \text{Subj had V3 TimeAdv} \Rightarrow \text{She had finished already.}$$

2. Отрицательное предложение:

$$S \Rightarrow \text{Neg} \Rightarrow \text{Subj had not V3 TimeAdv} \Rightarrow \text{They had not seen the film before.}$$

3. Вопросительное предложение:

$$S \Rightarrow \text{Quest} \Rightarrow \text{Had Subj V3 TimeAdv?} \Rightarrow \text{Had you visited London previously?}$$

2.2 БНФ задаваемого подмножества языка

Форма Бэкуса-Наура (БНФ) — форма представления контекстно-свободных формальных грамматик, в которой правые части продукций с одинаковой левой частью объединены. Расширенная БНФ включает «регулярные» конструкции: итерацию { a }, необязательность [a], группировку (a | b).

Грамматика подмножества языка в Past Perfect в расширенной форме Бэкуса-Наура:

```
<sentence>:= (affirmative | negative | question) [ "." | "?" | "!" ]

<affirmative>:= <subject> "had" <past_participle> <object> <adverbial>*
<negative>:= <subject> "had" "not" <past_participle> <object> <adverbial>*
<question>:= "had" <subject past_participle> <object>? <adverbial>*

<subject>:= <pronoun> | <noun_phrase>
<pronoun>:= "i" | "you" | "he" | "she" | "it" | "we" | "they"

<noun_phrase>:= (<determiner>)? <adjective>* (<noun> | <proper_noun> | <date>)
<determiner>:= "the" | "a" | "an" | "my" | "your" | "his" | "her" | "its"
| "our" | "their"

<adjective>:= "angry" | "happy" | "sad" | "excited" | "tired"
<noun>:= "movie" | "book" | "picture" | "performance" | "letter" |
"report" | "cat" | "dogs" | "bird" | "home" | "school" | "answer"

<proper_noun>:= "xuxa" | "bobik" | "tuzik" | "margo" | "pushok" |
"funtik" | "teacher" | "student" | "woman" | "girl" | "man"

<date>:= "2021", "2022", "2020" | "2023" | "2024"

<past_participle>:= <regular_past> | <irregular_past>
<regular_past>:= "killed" | "finished" | "saved" | "counted" | "painted" |
"visited" | "read" | "written" | "started" |
"returned" | "arrived" | "come" | "explained"

irregular_past: "eaten" | "seen" | "gone" | "taken" | "made" | "known"

<object>:= <noun_phrase>
<adverbial>:= (<adverb> | <prepositional_phrase> | <clause> | <date_phrase>)+
<adverb>:= "quickly" | "happily" | "yesterday"
<prepositional_phrase>:= <preposition> <noun_phrase>
<preposition>:= "before" | "after" | "in" | "on" | "at" | "by" | "from" | "to"
<clause>:= <sub_conj> <simple_sentence>
sub_conj: "before" | "after" | "when" | "while" | "because"

<date_phrase>:= ("by" | "after") <date>

<simple_sentence>:= (<simple_affirmative> | <simple_negative> | <simple_question>)
<simple_affirmative>:= <subject> <simple_verb object>?
<simple_negative>:= <subject> "did" "not" <simple_verb> <object>?
<simple_question>:= "did" <subject> <simple_verb> <object>?
```

```
<simple_verb>:= "explained" | "left" | "eaten" | "gone" | "visited" | "started"  
| "returned" | "arrived" | "finished" | "read"
```

Примеры предложений

1. Утвердительное:

He had returned an angry teacher.

2. Отрицательное:

She had not started his sad performance.

3. Вопросительное:

Had they not written movie after boy went to Paris?

2.2.1 Особенности грамматики

- Грамматика не может быть упрощена до автоматной из-за рекурсии в правилах для `<DeclarativePast>` и `<PlaceAdjunct>`.
- Поддерживает сложные временные конструкции с `before/after/by`.
- Позволяет выражать завершенность действия относительно другого прошлого события.
- Включает варианты с наречиями `already/never/yet`.

Пример дерева разбора представлен на рис.1.

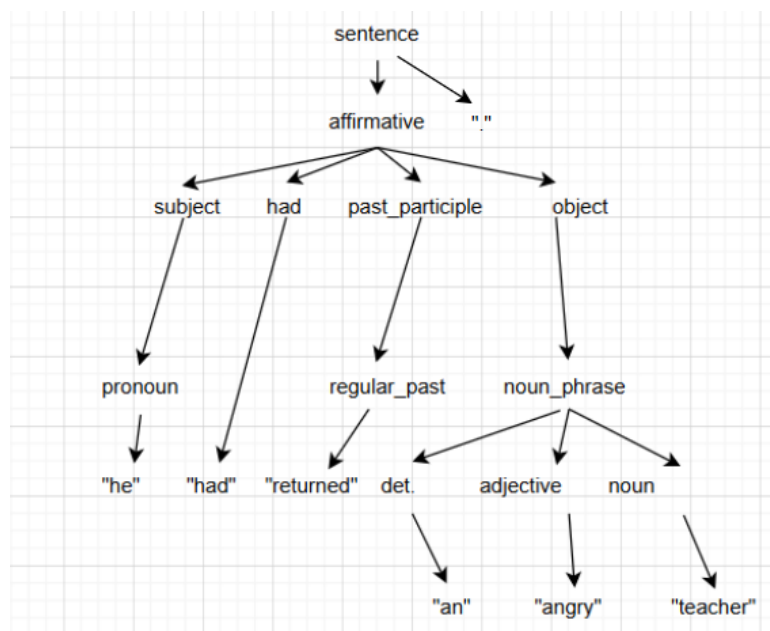


Рис. 1: Дерево разбора

3 Особенности реализации

3.1 Структуры данных

Следующие структуры данных представляют терминальные символы грамматики:

3.1.1 Местоимения (Pronouns)

```
pronoun = ["I", "you", "he", "she", "it", "we", "they"]
```

3.1.2 Определители (Determiners)

```
determiner = ["the", "a", "an", "my", "your", "his",  
"her", "its", "our", "their"]
```

3.1.3 Прилагательные (Adjectives)

```
adjective = ["angry", "happy", "sad", "excited", "tired"]
```

3.1.4 Существительные (Nouns)

```
noun = ["movie", "book", "picture", "performance", "letter",  
"report", "cat", "dogs", "bird", "home", "school", "answer"]
```

3.1.5 Имена собственные (Proper Nouns)

```
proper_noun = ["Xuxa", "Bobik", "Tuzik", "Margo", "Pushok",  
"Funtik", "teacher", "student", "woman", "girl", "man"]
```

3.1.6 Даты (Dates)

```
date = ["2020", "2021", "2022", "2023", "2024"]
```

3.1.7 Глаголы (Verbs)

- Правильные формы Past Participle:

```
regular_past = ["killed", "finished", "saved", "counted",  
"painted", "visited", "read", "written", "started", "returned",  
"arrived", "come", "explained"]
```

- Неправильные формы Past Participle:

```
irregular_past = ["eaten", "seen", "gone",  
"taken", "made", "known"]
```

- Простые формы глаголов:

```
simple_verb = ["explained", "left", "eaten",  
"gone", "visited", "started", "returned", "arrived",  
"finished", "read"]
```

3.1.8 Наречия (Adverbs)

```
adverb = ["quickly", "happily", "yesterday", "suddenly", "quietly"]
```

3.1.9 Предлоги и союзы

- Предлоги:

```
preposition = ["before", "after", "in", "on", "at", "by", "from", "to"]
```

- Подчинительные союзы:

```
sub_conj = ["before", "after", "when", "while", "because"]
```

3.2 Основные функции

3.2.1 validate_past_perfect(sentence)

| | |
|-----------------|---|
| Назначение | Проверяет, соответствует ли предложение грамматике Past Perfect |
| Входные данные | <ul style="list-style-type: none">• sentence: str - предложение для проверки |
| Выходные данные | <ul style="list-style-type: none">• bool - True если предложение валидно, False если нет |
| Описание | Использует парсер Lark для проверки соответствия грамматике. При ошибке парсинга возвращает False . |

3.2.2 interactive_validation()

| | |
|-----------------|---|
| Назначение | Интерактивный режим проверки предложений |
| Входные данные | Предложения введенные пользователем |
| Выходные данные | Результат валидации - строка |
| Описание | Бесконечный цикл, запрашивающий предложения для проверки до ввода 'q'. Для каждого предложения вызывает validate_past_perfect() и выводит результат. |

3.2.3 generate_sentence()

| | |
|-----------------|---|
| Назначение | Генерирует случайное предложение в Past Perfect |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - сгенерированное предложение с правильной пунктуацией |
| Описание | Случайно выбирает между утвердительным, отрицательным и вопросительным типом, вызывает соответствующую функцию генерации и добавляет правильную пунктуацию. |

3.3 Функции генерации частей предложения

3.3.1 generate_subject()

| | |
|-----------------|--|
| Назначение | Генерирует подлежащее |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - подлежащее (местоимение или именная группа) |
| Описание | С вероятностью 70% выбирает местоимение, иначе генерирует именную группу (без дат). |

3.3.2 generate_noun_phrase(exclude_date=False)

| | |
|-------------------|--|
| Назначение | Генерирует именную группу |
| Входные параметры | <ul style="list-style-type: none">• exclude_date: bool - исключать ли даты из генерации |
| Выходные данные | <ul style="list-style-type: none">• str - именная группа |
| Описание | Генерирует группу с возможными: определителем (с правильным артиклем), 0-2 прилагательными и существительным/именем собственным/датой. |

3.3.3 generate_past_participle()

| | |
|-----------------|---|
| Назначение | Выбирает форму Past Participle |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - глагол в форме Past Participle |
| Описание | Случайно выбирает из списков правильных и неправильных глаголов. |

3.3.4 generate_object()

| | |
|-----------------|--|
| Назначение | Генерирует дополнение |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - дополнение (именная группа) |
| Описание | Вызывает <code>generate_noun_phrase()</code> для генерации дополнения. |

3.3.5 generate_adverbial()

| | |
|-----------------|---|
| Назначение | Генерирует обстоятельство |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - обстоятельство |
| Описание | Случайно выбирает между: наречием, предложной группой, придаточным предложением или указанием даты. |

3.3.6 generate_adverbials()

| | |
|-----------------|--|
| Назначение | Генерирует несколько обстоятельств |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - последовательность обстоятельств |
| Описание | С 50% вероятностью добавляет хотя бы одно обстоятельство, затем с 30% вероятностью добавляет дополнительные. |

3.4 Функции генерации типов предложений

3.4.1 generate_affirmative()

| | |
|-----------------|--|
| Назначение | Генерирует утвердительное предложение |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - предложение в форме Past Perfect (без пунктуации) |
| Описание | Генерирует структуру: Подлежащее + "had" + Past Participle + Дополнение + Обстоятельства. |

3.4.2 generate_negative()

| | |
|-----------------|---|
| Назначение | Генерирует отрицательное предложение |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - отрицательное предложение (без пунктуации) |
| Описание | Генерирует структуру: Подлежащее + "had not" + Past Participle + Дополнение + Обстоятельства. |

3.4.3 generate_question()

| | |
|-----------------|--|
| Назначение | Генерирует вопросительное предложение |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - вопрос (без вопросительного знака) |
| Описание | Генерирует структуру: "had" + Подлежащее + Past Participle + (Дополнение с 70% вероятностью) + Обстоятельства. |

3.4.4 generate_simple_sentence()

| | |
|-----------------|--|
| Назначение | Генерирует простое предложение для придаточных |
| Входные данные | Псевдослучайное число |
| Выходные данные | <ul style="list-style-type: none">• str - простое предложение в Past Simple |
| Описание | Случайно выбирает между утвердительной и отрицательной формой, использует простые глаголы. |

3.5 Код программы

```
1  from lark import Lark
2  import random
3
4  GRAMMAR = """
5      start: sentence
6
7      sentence: question_punct | affirmative_punct | negative_punct
8
9      question_punct: question "?"
10     affirmative_punct: affirmative "."
11     negative_punct: negative "."
12
13     question: "had" subject past_participle object? adverbial*
14     affirmative: subject "had" past_participle object adverbial*
15     negative: subject "had" "not" past_participle object adverbial*
16
17     subject: pronoun | noun_phrase
18     pronoun: "i" | "you" | "he" | "she" | "it" | "we" | "they"
19
20     noun_phrase: (determiner)? adjective* (noun | proper_noun | date)
21     determiner: "the" | "a" | "an" | "my" | "your" | "his" | "her" | "its"
22                | "our" | "their"
23     adjective: "angry" | "happy" | "sad" | "excited" | "tired"
24     noun: "movie" | "book" | "picture" | "performance" | "letter" | "
25            report" | "cat" | "dogs" | "bird" | "home" | "school" | "answer"
26     proper_noun: "xuxa" | "bobik" | "tuzik" | "margo" | "pushok" | "
27                funtik" | "teacher" | "student" | "woman" | "girl" | "man"
28
29     date: "2021" | "2022" | "2020" | "2023" | "2024"
```

```

27
28     past_participle: regular_past | irregular_past
29     regular_past: "killed" | "finished" | "saved" | "counted" | "painted"
        | "visited" | "read" | "written" | "started" | "returned" | "
        arrived" | "come"|"explained"
30     irregular_past: "eaten" | "seen" | "gone" | "taken" | "made" | "known
        "
31
32     object: noun_phrase
33     adverbial: (adverb | prepositional_phrase | clause | date_phrase)+
34     adverb: "quickly" | "happily" | "yesterday" | /[a-z]+ly/
35     prepositional_phrase: preposition noun_phrase
36     preposition: "before" | "after" | "in" | "on" | "at" | "by" | "from"
        | "to"
37     clause: sub_conj simple_sentence
38     sub_conj: "before" | "after" | "when" | "while" | "because"
39
40     date_phrase: ("by" | "after") date
41
42     simple_sentence: (simple_affirmative | simple_negative | simple_
        question)
43     simple_affirmative: subject simple_verb object?
44     simple_negative: subject "did" "not" simple_verb object?
45     simple_question: "did" subject simple_verb object?
46     simple_verb: "explained" | "left" | "eaten" | "gone" | "visited" | "
        started" | "returned" | "arrived" | "finished" | "read"
47
48     %import common.WS
49     %ignore WS
50
51     """
52     parser = Lark(GRAMMAR, parser='earley')
53
54     def interactive_validation():
55         print("Past Perfect Tense Validator")
56         print("Enter sentences to validate (press 'q' to exit):")
57         print("=" * 50)
58
59         while True:
60             sentence = input("> ").strip()
61             if sentence.lower() == 'q':
62                 print("Exiting...")
63                 break
64
65             result = validate_past_perfect(sentence)
66             print(sentence)
67             print(f"Result: {'VALID' if result else 'INVALID'}")
68             print("=" * 50)
69
70     def validate_past_perfect(sentence):
71         try:
72             parser.parse(sentence.lower())
73             return True
74         except Exception as e:
75             print(f"Error parsing: {e}")
76             return False
77
78     pronoun = ["I", "you", "he", "she", "it", "we", "they"]
79     determiner = ["the", "a", "an", "my", "your", "his", "her", "its", "our",
        "their"]
80     adjective = ["angry", "happy", "sad", "excited", "tired"]

```

```

81 noun = ["movie", "book", "picture", "performance", "letter", "report", "
    cat", "dogs", "bird", "home", "school", "answer"]
82 proper_noun = ["Xuxa", "Bobik", "Tuzik", "Margo", "Pushok", "Funtik", "
    teacher", "student", "woman", "girl", "man"]
83 date = ["2020", "2021", "2022", "2023", "2024"]
84 regular_past = ["killed", "finished", "saved", "counted", "painted", "
    visited", "read", "written", "started", "returned", "arrived", "come",
    "explained"]
85 irregular_past = ["eaten", "seen", "gone", "taken", "made", "known"]
86 adverb = ["quickly", "happily", "yesterday", "suddenly", "quietly"]
87 preposition = ["before", "after", "in", "on", "at", "by", "from", "to"]
88 sub_conj = ["before", "after", "when", "while", "because"]
89 simple_verb = ["explained", "left", "eaten", "gone", "visited", "started
    ", "returned", "arrived", "finished", "read"]
90
91 def generate_sentence():
92     sentence_type = random.choice(['affirmative', 'negative', 'question
        '])
93
94     if sentence_type == 'affirmative':
95         sentence = generate_affirmative()
96         return sentence[0].upper() + sentence[1:] + '.'
97     elif sentence_type == 'negative':
98         sentence = generate_negative()
99         return sentence[0].upper() + sentence[1:] + '.'
100    else:
101        sentence = generate_question()
102        return sentence[0].upper() + sentence[1:] + '?'
103
104    def generate_subject():
105        if random.random() < 0.7:
106            return random.choice(pronoun)
107        else:
108            return generate_noun_phrase(exclude_date=True)
109
110    def generate_noun_phrase(exclude_date=False):
111        parts = []
112        if random.random() < 0.6:
113            next_word = ""
114            if random.random() < 0.5:
115                next_word = random.choice(adjective + noun + proper_noun)
116            else:
117                next_word = random.choice(noun + proper_noun)
118
119            if next_word[0].lower() in ['a', 'e', 'i', 'o', 'u'] and random.
                choice(determiner) in ['a', 'an']:
120                parts.append('an' if next_word[0].lower() in ['a', 'e', 'i',
                    'o', 'u'] else 'a')
121            else:
122                parts.append(random.choice([d for d in determiner if d not in
                    ['a', 'an']]))
123
124            parts += random.sample(adjective, k=random.randint(0, 2))
125
126            choices = ['noun', 'proper_noun']
127            if not exclude_date and random.random() < 0.2:
128                choices.append('date')
129
130            noun_type = random.choice(choices)
131
132            if noun_type == 'noun':

```

```

133     parts.append(random.choice(noun))
134 elif noun_type == 'proper_noun':
135     parts.append(random.choice(proper_noun))
136 else:
137     parts.append(random.choice(date))
138
139     return ' '.join(parts)
140
141 def generate_past_participle():
142     return random.choice(regular_past + irregular_past)
143
144 def generate_object():
145     return generate_noun_phrase()
146
147 def generate_adverbial():
148     adverbial_type = random.choice(['adverb', 'prepositional_phrase', '
149                                     clause', 'date_phrase'])
150
151     if adverbial_type == 'adverb':
152         return random.choice(adverb)
153     elif adverbial_type == 'prepositional_phrase':
154         return f"{random.choice(preposition)} {generate_noun_phrase()}"
155     elif adverbial_type == 'clause':
156         return f"{random.choice(sub_conj)} {generate_simple_sentence().
157                     lower()}"
158     else:
159         return f"{random.choice(['by', 'after'])} {random.choice(date)}"
160
161 def generate_adverbials():
162     adverbials = []
163     if random.random() < 0.5:
164         adverbials.append(generate_adverbial())
165         while random.random() < 0.3:
166             adverbials.append(generate_adverbial())
167     return ' '.join(adverbials)
168
169 def generate_affirmative():
170     parts = [
171         generate_subject().lower(),
172         "had",
173         generate_past_participle(),
174         generate_object(),
175         generate_adverbials()
176     ]
177     return ' '.join([p for p in parts if p]).strip()
178
179 def generate_negative():
180     parts = [
181         generate_subject().lower(),
182         "had not",
183         generate_past_participle(),
184         generate_object(),
185         generate_adverbials()
186     ]
187     return ' '.join([p for p in parts if p]).strip()
188
189 def generate_question():
190     parts = [
191         "had",
192         generate_subject().lower(),
193         generate_past_participle(),

```



```

192         generate_object() if random.random() < 0.7 else "",
193         generate_adverbials()
194     ]
195     return ' '.join([p for p in parts if p]).strip()
196
197 def generate_simple_sentence():
198     sentence_type = random.choice(['simple_affirmative', 'simple_negative',
199                                   ''])
200
201     if sentence_type == 'simple_affirmative':
202         return f"{random.choice(pronoun).lower()} {random.choice(simple_verb)} {generate_object()}"
203     else:
204         return f"{random.choice(pronoun).lower()} did not {random.choice([v for v in simple_verb if v not in irregular_past])} {generate_object()}"
205
206 for _ in range(10):
207     sentence = generate_sentence()
208     is_valid = validate_past_perfect(sentence)
209     print(f"Generated: {sentence}")
210     print(f"Validation: {'VALID' if is_valid else 'INVALID'}")
211     print("=" * 60)

```

4 Результаты работы программы

Ниже представлены результаты работы функции генерации и валидации.

Генерация предложений:

Generated: I had known tired report.

Validation: VALID

Generated: They had finished our angry sad picture.

Validation: VALID

Generated: Had they made excited sad teacher?

Validation: VALID

Generated: We had not finished their tired excited performance.

Validation: VALID

Generated: It had not explained excited movie.

Validation: VALID

Generated: It had not killed Pushok.

Validation: VALID

Generated: She had started sad angry movie because it finished tired letter.

Validation: VALID

Проверка предложений:

They had not killed the sad happy woman.

Result: VALID

They had not killed the sad happy woman?

Result: INVALID

They had not killed the sad happy woman

Result: INVALID

They had killed the sad happy woman.

Result: VALID

They had killed the sad happy woman?

Result: INVALID

They had killed the sad happy woman

Result: INVALID

They had not killed the sad happy.

Result: INVALID

They not killed the sad happy woman.

Result: INVALID

Заключение

В ходе лабораторной работы были разработаны анализатор и генератор предложений подмножества английского языка в прошедшем совершенном времени (Past Perfect). Для этого была создана контекстно-свободная грамматика, соответствующая второму типу в иерархии Хомского. На основе этой грамматики была записана БНФ-нотация. Программа написана на языке Python.

Анализатор способен обрабатывать подмножество естественного языка. Однако у него есть ограничения: жесткая структура и отсутствие смысловой связи между словами в сгенерированных предложениях.

Достоинства

- Модульность кода. Каждая функция выполняет одну задачу (например, `generate_subject()`, `generate_object()`).
- Код поддерживает случайную генерацию разных типов предложений (утвердительные, отрицательные, вопросы). Учитывается вероятность появления разных элементов (например, обстоятельств с вероятностью 50%).

Недостатки

- Ограниченный словарный запас. Списки слов (noun, adjective, adverb и т. д.) фиксированы и небольшие. Нет возможности легко расширять их без изменения кода.
- Из-за случайной генерации могут получаться грамматически правильные, но странные фразы (например, "Had they eaten the happy 2023?").

Для масштабирования программы можно расширить словари и добавить новые слова. Также стоит рассмотреть методы подбора связанных слов для более естественного звучания предложений. Одним из возможных решений может быть использование эмбедингов, обученных на готовых текстах.

Список литературы

- [1] Электронный ресурс ВШТИИ
URL:<https://tema.spbstu.ru/compiler/>
(Дата обращения: 13.04.2025).
- [2] Карпов, Ю. Г. Теория автоматов, Санкт-Петербург : Питер, 2003.
URL:<https://djvu.online/file/eeLVKnyRZPXfl>
(Дата обращения: 17.04.2025).