|  |  |
| --- | --- |
|  | ы |
| МИНОБРНАУКИ РОССИИ |  |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА – Российский технологический университет»**  **РТУ МИРЭА** |  |

**Институт информационных технологий**

КАФЕДРА ИНСТРУМЕТНАЛЬНОГО И ПРИКЛАДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ (ИППО)

пРАКТИЧЕСКАЯ РАБОТА № 1

по дисциплине «Теория автоматов и формальных языков»

«ЛЕКСИЧЕСКИЙ АНАЛИЗАТОР»

Выполнил студент группы ИКБО-12-18 Копотов М.А.

Принял Алпатов А.Н.

Практическая работа выполнена «25» сентября 2020г.

«Зачтено» «\_\_\_»\_\_\_\_\_\_\_2020г.

Москва 2020

**Оглавление**

[Цель практической работы 3](#_Toc52228313)

[Теоретические сведения 3](#_Toc52228314)

[Порядок выполнения работы 4](#_Toc52228315)

[Вывод 4](#_Toc52228316)

# Цель практической работы

Целью данной практической работы является ознакомление с базовыми понятиями теории формальных языков, грамматики и их свойств; изучение правил построения цепочек, распознавание принадлежности цепочки к определенному алфавиту по имеющимся правилам.

# Теоретические сведения

Тексты на любом языке (естественном или формальном) представляют собой цепочки символов некоторого алфавита. Цепочкой символов называют произвольную упорядоченную конечную последовательность символов, записанных один за другим.

Понятие символа является базовым в теории формальных языков. Для цепочки символов имеют значение три фактора: состав входящих в цепочку символов, их количество и порядок символов в цепочке. Цепочки символов α и β равны (α = β), если они имеют один и тот же состав символов, одно и то же их количество и одинаковый порядок следования символов в цепочке. Количество символов в цепочке определяет её длину. Длина цепочки α обозначается как |α |. Можно выделить следующие операции над цепочками символов: Можно выделить следующие операции над цепочками символов: • конкатенация (объединение, сложение двух цепочек) – это дописывание второй цепочки в конец первой. Конкатенация цепочек α и β обозначается как αβ. Например: α = аб, β = вг, тогда αβ = абвг. При этом 5 αβ ≠ βα, так как в цепочке важен порядок символов. Но конкатенация обладает свойством ассоциативности: (αβ)γ = α(βγ); • замена (подстановка) – замена подцепочки символов на любую произвольную цепочку символов. В результате получается новая цепочка символов. Например: γ = абвг, разобьём эту цепочку символов на подцепочки: α = а, ω = б, β = вг и выполним подстановку цепочки υ = аба вместо подцепочки ω. Получим новую цепочку γ' = аабавг. Таким образом, подстановка выполняется путём разбиения исходной цепочки на подцепочки и конкатенации; • обращение – запись символов цепочки в обратном порядке. Эта операция обозначается как αR. Если α = абвг, то αR = гвба. Пустая цепочка символов – это цепочка, не содержащая ни одного символа. Основой любого языка является алфавит, определяющий набор допустимых символов языка. Алфавит – это счётное множество допустимых символов языка, обозначаемое как V. Грамматика – это описание способа построения предложений некоторого языка. Она относится ко второму способу определения языков – порождению цепочек символов. Грамматику языка можно описать различными способами. Например, можно использовать формальное описание грамматики, построенное на основе системы правил (или продукций).

# Порядок выполнения работы

# Задание 1. Дана грамматика. Построите вывод заданной̆ цепочки

*a) S → T | T+S | T–S*

*T → F | F\*T*

*F → a | b*

*Цепочка a–b\*a+b*

Решение: **S** *→ T –* ***S*** *→ T – T +* ***S*** *→*

***T*** *– T + T → F –* ***T*** *+ T → F – F\*****T*** *+ T → F – F\*F +* ***T*** *→*

***F*** *– F\*F + F → a –* ***F****\*F + F → a – b\*****F*** *+ F → a – b\*a +* ***F*** *→*

*a – b\*a + b*

*b) S →aSBC | abC*

*CB → BC*

*bB → bb*

*bC →bc*

*cC →cc*

*Цепочка aaabbbccc*

Решение: ***S*** *→ a****S****BC → aa****S****BCBC → aaab****CB****CBC →aaabBC****CB****C → aaabB****CB****CC → aaa****bB****BCCC → aaab****bB****CCC → aaabb****bC****CC → aaabbb****cC****C → aaabbbc****cC*** *→ aaabbbccc*

# Задание 2. Какой̆ язык порождается грамматикой̆ с правилами

*a) S → aaCFD*

*AD → D*

*F → AFB | AB*

*Cb → bC*

*AB→ bBA*

*CB →С*

*Ab→ bA*

*bCD →ε*

*aaC****AB****D → aa Количество b: 0 = 1^2 - 1*

*aaC****AABB****D → aabbb Количество b: 3 = 2^2 - 1*

*aaC****AAABBB****D → aabbbbbbbb Количество b: 8 = 3^2 - 1*

*aaC****AAAABBBB****D → aabbbbbbbbbbbbbbb Количество b: 15 = 4^2 – 1*

*Ответ:*

*b) S→A*⊥*|B*⊥

*A → a | Ba*

*B → b | Bb | Ab*

*Решение: получается язык, состоящий их слов, чередующих буквы* ***a,b****, причем так, что начинаться и заканчиваться слово может с любой из букв, между любыми буквами* ***a*** *должны стоять* ***bn****при n > 0, а на конце слов всегда стоит* ⊥

# Задание 3. Построить грамматику, порождающую язык:

*а)* **L** = {*anbmck* | *n*, *m*, *k* > 0}

Решение:

S *→ ABC*

*A → a | aA*

*B → b | bB*

*C → c | cC*

*б)* **L** = {0*n*(10)*m* |*n*, *m* ≥ 0}

Решение:

S *→ A | B | AB*

*A → 0 | 0A*

*B → 10 | 10B*

*в)* **L** = {*a*1*a*2...*anan*...*a*2*a*1 | *ai* Î {0, 1}}

Пример при n = 4; a1a2a3a4a4a3a2a1

Решение:

S *→ AA | ASA*

*A → 0 | 1*

**5. Эквивалентны ли грамматики с правилами:**

*S → aSL | aL L → Kc*

*cK → Kc*

*K → b\*

*S → aSBc | abc*

*cB → Bc*

*bB → bb*

Решение:

*первая грамматика определяет язык ancnbn, где n >= 0*

*Вторая грамматика определяет язык язык ancnbn, где n >= 0*

*Пример второй грамматики: aaabcBcBc -> aaabBBccc -> aaabbbccc.*

*Следовательно Ответ: грамматики эквивалентны, так как определяют один язык*

# Вывод

В ходе выполнения практической работы были произведено ознакомление с базовыми понятиями теории формальных языков, грамматики и их свойств, также были изучены правила построения цепочек, распознавание принадлежности цепочки к определенному алфавиту по имеющимся правилам.

# Список литературы

# А.А.Малявко. Формальные языки и компиляторы. Учебное пособие для вузов. – М.:ACT, 2017 – 433c.

# А.Саломаа. Жемчужины теории формальных языков. Учебное пособие для вузов. – М.:МИР, 1986 – 160c.

# Е.Н.Ешакова. Теория языков программирования и методов трансляции. – Спб.:БИБКОМ, 2007 – 137с.