

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

	Рабочая программа дисциплины (модуля)
по дисциплине:	Формальные языки и трансляции
по направлению:	Прикладная математика и информатика
профиль подготовки:	Проектирование и разработка комплексных бизнес-приложений Физтех-школа Прикладной Математики и Информатики кафедра алгоритмов и технологий программирования
курс:	2
квалификация:	бакалавр

Семестр, формы промежуточной аттестации: 3 (осенний) - Дифференцированный зачет

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 75 час.

Всего часов: 135, всего зач. ед.: 3

Программу составил: О.Н. Ивченко, заведующий кафедрой

Программа обсуждена на заседании кафедры алгоритмов и технологий программирования 10.02.2025

Аннотация

В курсе рассматриваются основные понятия теории формальных языков. Прежде всего, вводятся базовые понятия алфавита, слова и языка, после чего рассматриваются различные способы задания множеств слов (языков). Среди таких способов выделяются два основных: конечные автоматы и контекстно-свободные грамматики. Также рассматриваются эквивалентные перечисленным способы описания: регулярные выражения и праволинейные грамматики, задающие автоматные языки, и автоматы с магазинной памятью, эквивалентные контекстно-свободным грамматикам. Кроме того, в курсе подробно рассматриваются конечные преобразователи. Курс содержит большое количество математических теорем, описывающих свойства введенных объектов и соотношения между ними. Также в курсе подробно разбираются алгоритмы синтаксического анализа для контекстно-свободных грамматик, их алгоритмическая сложность и условия применения. Курс содержит большое количество теоретического материала, задач, а также практические задания на реализацию алгоритмов. Для успешного освоения курса слушателю необходимо изучить дисциплины «Математическая логика и теория алгоритмов» и «Алгоритмы и структуры данных», а также владеть основными методами доказательства математических утверждений.

1. Цели и задачи

Цель дисциплины

Знакомство студентов с основными разделами теории формальных языков для последующего использования в других дисциплинах естественнонаучного содержания и практической деятельности, развитие математической культуры, исследовательских и программистских навыков.

Задачи дисциплины

- Заложить базовые знания в области теории формальных языков.
- Развить общематематическую культуру: умение логически мыслить, формулировать и доказывать строгие математические утверждения.
- Научить выбирать алгоритм для решения задачи, обосновывать его правильность и реализовывать на требуемом языке программирования.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-1 Способен применять фундаментальные знания, полученные в области физико-математических и (или) естественных наук и использовать их в профессиональной деятельности	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.2 Способен строить математические модели, производить количественные расчеты и оценки
	ОПК-1.3 Способен определять границы применимости полученных результатов
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.1 Способен применять современные вычислительную технику и сервисы сети Интернет в области (сфере) профессиональной деятельности
	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
	ОПК-2.3 Знает основные требования информационной безопасности

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны знать:

- понятия формального языка и грамматики;
- классификацию формальных языков и грамматик;
- регулярные выражения;
- детерминированные и недетерминированные конечные автоматы;
- контекстно-свободные грамматики;
- основные алгоритмы синтаксического разбора для контекстно-свободных грамматик;
- автоматы с магазинной памятью;
- важнейшие подклассы контекстно-свободных грамматик (в т. ч. LR(k)-грамматики);
- конечные преобразователи;
- практическое применение автоматов и грамматик, в т. ч. в задачах компиляции.

уметь:

- делать правильные выводы из сопоставления результатов теории и эксперимента;
- определять класс грамматик, необходимый для решения задачи;
- формально доказывать соответствие формальных языков и грамматик поставленной задаче;
- строить конечный автомат/регулярное выражение по описанию автоматного языка;
- строить регулярное выражение по заданному конечному автомату;
- строить конечный автомат по заданному регулярному выражению;
- строить контекстно-свободную грамматику по описанию контекстно-свободного языка;
- строить автомат с магазинной памятью для заданной контекстно-свободной грамматики;
- строить синтаксический анализатор для заданной контекстно-свободной грамматики с помощью генераторов синтаксических анализаторов

владеть:

- математическим аппаратом теории формальных языков;
- навыками самостоятельной работы в современных программных комплексах;
- навыками освоения большого объема информации;
- навыками использования генераторов синтаксических анализаторов.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Задачи, цели и методы теории формальных языков.	5	5		10
2	Регулярные выражения.	5	5		10
3	Конечные автоматы.	5	5		10
4	Контекстно-свободные грамматики.	5	5		10
5	Алгоритмы синтаксического анализа.	5	5		10
6	Конечные преобразователи.	2	3		10
7	Приложения теории формальных языков.	3	2		15
Итого часов		30	30		75
Подготовка к экзамену		0 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 3 (Осенний)

1. Задачи, цели и методы теории формальных языков.

Теория формальных языков как научная дисциплина. История теории формальных языков, её историческое значение для лингвистики и программирования. Философское значение теории формальных языков. Краткое описание современных задач.

2. Регулярные выражения.

Понятие регулярного выражения и регулярного языка. Замкнутость класса регулярных языков относительно основных операций над языками. Отдельные примеры регулярных языков

3. Конечные автоматы.

Недетерминированные конечные автоматы. Различные варианты их определения. Эквивалентность этих вариантов. Эпсилон-переходы и их устранение. Детерминированные конечные автоматы. Совпадение классов языков, задаваемых детерминированными и недетерминированными автоматами. Алгоритм детерминизации конечного автомата. Теорема Клини о совпадении классов регулярных и автоматных языков. Алгоритм построения конечного автомата по заданному регулярному выражению. Понятие минимального автомата. Минимизация детерминированных конечных автоматов. Алгоритм проверки равенства регулярных выражений. Лемма о разрастании для автоматных языков. Примеры неавтоматных языков.

4. Контекстно-свободные грамматики.

Понятие порождающей грамматики. Иерархия Хомского. Праволинейные грамматики. Совпадение классов языков, задаваемых праволинейными грамматиками и конечным автоматами. Понятие контекстно-свободной грамматики и контекстно-свободного языка. Вывод в контекстно-свободной грамматике и деревья вывода. Понятие однозначной грамматики. Замкнутость класса контекстно-свободных языков относительно отдельных операций. Нормальные формы для контекстно-свободных языков. Устранение эпсилон-правил в контекстно-свободных грамматиках. Нормальные формы Хомского и Грейбах. Автоматы с магазинной памятью, варианты определения. Эквивалентность различных вариантов. Автомат с магазинной памятью в нормальной форме. Совпадение классов языков, задаваемых автоматами с магазинной памятью и контекстно-свободными грамматиками. Лемма о разрастании для контекстно-свободных языков. Примеры языков, не являющихся контекстно-свободными. Детерминированные автоматы с магазинной памятью, различные варианты определения. Детерминированные контекстно-свободные языки. Однозначность детерминированных контекстно-свободных языков. Свойства класса детерминированных контекстно-свободных языков. Линейные и полулинейные множества, теорема Парика.

5. Алгоритмы синтаксического анализа.

Алгоритм Кока-Янгера-Касами для контекстно-свободной грамматики в нормальной форме Хомского, алгоритм Эрли синтаксического разбора, алгоритм восходящего синтаксического анализа «перенос-свёртка», LR-таблица и алгоритм синтаксического разбора по LR-таблице, понятие LR-ситуации, состояния LR-анализатора, алгоритм построения LR-таблицы

6. Конечные преобразователи.

Понятие конечного преобразователя. Различные варианты его определения, их равносильность. Примеры преобразований, задаваемых конечными преобразователями. Теорема Нива об общем виде конечного преобразования. Замкнутость автоматных и контекстно-свободных языков относительно конечных преобразований. Теорема Хомского-Шютценберже, слабая и сильная формы.

7. Приложения теории формальных языков.

Современные приложения теории формальных языков. Применение конечных преобразователей в задачах автоматического морфологического анализа.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система), а так же персональными компьютерами для обучающихся.

6. Перечень рекомендуемой литературы

Основная литература

1. Математическая теория формальных языков, Электрон. версия печ. публикации / А. Е. Пентус, М. Р. Пентус. — Москва, ИНТУИТ, 2016

Дополнительная литература

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На лекционных занятиях используются мультимедийные технологии, включая демонстрацию презентаций.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

Успешное освоение дисциплины требует:

- посещения студентом всех видов аудиторных занятий;
- ведения конспекта в ходе аудиторных занятий;
- качественной самостоятельной подготовки к практическим занятиям, активной работы на них;
- активной самостоятельной и аудиторной работы студента;
- своевременной сдачи преподавателю заданий по аудиторным видам работ.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Прикладная математика и информатика
профиль подготовки:	Проектирование и разработка комплексных бизнес-приложений Физтех-школа Прикладной Математики и Информатики кафедра алгоритмов и технологий программирования
курс:	<u>2</u>
квалификация:	бакалавр
Семестр, формы промежуточной аттестации: 3 (осенний) - Дифференцированный зачет	
Разработчик:	О.Н. Ивченко, заведующий кафедрой

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-1 Способен применять фундаментальные знания, полученные в области физико-математических и (или) естественных наук и использовать их в профессиональной деятельности	ОПК-1.1 Способен анализировать поставленную задачу, намечать пути ее решения
	ОПК-1.2 Способен строить математические модели, производить количественные расчеты и оценки
	ОПК-1.3 Способен определять границы применимости полученных результатов
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.1 Способен применять современные вычислительную технику и сервисы сети Интернет в области (сфере) профессиональной деятельности
	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
	ОПК-2.3 Знает основные требования информационной безопасности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Формальные языки и трансляции» обучающийся должен:

знать:

- понятия формального языка и грамматики;
- классификацию формальных языков и грамматик;
- регулярные выражения;
- детерминированные и недетерминированные конечные автоматы;
- контекстно-свободные грамматики;
- основные алгоритмы синтаксического разбора для контекстно-свободных грамматик;
- автоматы с магазинной памятью;
- важнейшие подклассы контекстно-свободных грамматик (в т. ч. LR(k)-грамматики);
- конечные преобразователи;
- практическое применение автоматов и грамматик, в т. ч. в задачах компиляции.

уметь:

- делать правильные выводы из сопоставления результатов теории и эксперимента;
- определять класс грамматик, необходимый для решения задачи;
- формально доказывать соответствие формальных языков и грамматик поставленной задаче;
- строить конечный автомат/регулярное выражение по описанию автоматного языка;
- строить регулярное выражение по заданному конечному автомату;
- строить конечный автомат по заданному регулярному выражению;
- строить контекстно-свободную грамматику по описанию контекстно-свободного языка;
- строить автомат с магазинной памятью для заданной контекстно-свободной грамматики;
- строить синтаксический анализатор для заданной контекстно-свободной грамматики с помощью генераторов синтаксических анализаторов

владеть:

- математическим аппаратом теории формальных языков;
- навыками самостоятельной работы в современных программных комплексах;
- навыками освоения большого объёма информации;
- навыками использования генераторов синтаксических анализаторов.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

1. Построить регулярное выражение для языка слов, в которых число букв a чётно, а букв b – нечётно.
2. Построить конечный автомат для языка слов, в которых число букв a чётно, а букв b – нечётно.

3. Построить минимальный детерминированный конечный автомат для языка, задаваемого выражением $(a(ab+ba)^*b(a+ba)^*)^*$.
4. Найти регулярное выражение для дополнения языка $(a(ab+ba)^*b(a+ba)^*)$.
5. Проверить, является ли автоматным множество слов вида $anbncn$.
6. Построить контекстно-свободную грамматику для множества слов в алфавите $\{a,b,c\}$, где букв a больше, чем букв b .
7. Построить автомат с магазинной памятью для множества слов в алфавите $\{a,b,c\}$, где букв a больше, чем букв b .
8. Построить грамматику в нормальной форме Хомского для множества слов в алфавите $\{a,b,c\}$, где букв a больше, чем букв b .
9. Доказать, что язык слов вида $anbnc^{2n}$ не является контекстно-свободным.
10. Проверить выводимость заданного слова в заданной грамматике с помощью алгоритма Кока-Янгера-Касами.
11. Проверить выводимость заданного слова в заданной грамматике с помощью алгоритма Эрли.
12. Построить LR-анализатор для заданной грамматики и проверить выводимость заданного слова с его помощью.
13. Построить конечный преобразователь, осуществляющий преобразование $a \rightarrow b \parallel a+c_d$.
14. Построить конечный преобразователь, умножающий число на 5 в двоичной системе счисления.

Контрольная работа состоит из 5 задач. Оценка «отлично» ставится, если обучающийся способен выбрать верный алгоритм во всех задачах, после чего довести их до корректного ответа. При этом решение должно быть полным и теоретически обоснованным. Оценка «хорошо» ставится, если не более, чем в одной задаче обучающийся выбрал неверный алгоритм, допустил ошибки в рассуждениях или обоснованиях. Оценка «удовлетворительно» ставится, если существенные недостатки имеются в двух или трёх задачах. Если число полностью решенных задач равно 1 и менее, ставится оценка «неудовлетворительно».

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Недетерминированные конечные автоматы (НКА). Различные варианты определений. Автоматные языки.
2. Детерминированные конечные автоматы (ДКА). Эквивалентность ДКА и НКА.
3. Свойства класса автоматных языков. Замкнутость относительно булевых операций.
4. Регулярные выражения. Теорема Клини о совпадении классов регулярных и автоматных языков.
5. Лемма о разрастании для автоматных языков. Примеры неавтоматных языков.
6. Минимальные ДКА. Алгоритм минимизации ДКА.
7. Алгоритм проверки равенства регулярных выражений. Теорема Майхилла-Нерода.
8. Праволинейные грамматики. Праволинейные языки. Теорема о совпадении классов автоматных и праволинейных языков.
9. Контекстно-свободные грамматики (КС-грамматики). Контекстно-свободные языки (КС-языки). Вывод в КС-грамматике, деревья вывода. Однозначные КС-грамматики. Пример существенно-неоднозначного языка (без доказательства).
10. Устранение бесполезных вспомогательных символов и ϵ -правил для КС-грамматик.
11. Нормальная форма Хомского для КС-грамматик. Алгоритм приведения к нормальной форме Хомского.
12. Алгоритм Кока-Янгера-Касами синтаксического разбора для КС-грамматик.
13. Автоматы с магазинной памятью (МП-автоматы). Различные варианты определений. Языки, распознаваемые МП-автоматами.
14. Совпадение классов КС-языков и языков, распознаваемых МП-автоматами.
15. Лемма о разрастании для КС-языков. Примеры языков, не являющихся КС-языками.
16. Алгоритм Эрли синтаксического разбора для КС-грамматик.
17. Алгоритм синтаксического разбора по LR-таблице.

18. Понятие LR-ситуации, состояния LR-анализатора. Операции GOTO и CLOSURE.
 19. Алгоритм построения LR-таблицы. Корректность LR-анализатора.
 20. Нормальная форма Грейбах для КС-грамматик.
 21. МП-автоматы без ϵ -переходов и распознаваемые ими языки.
 22. Детерминированные МП-автоматы. Детерминированные КС-языки. Различные варианты определения.
 23. Детерминированные КС-языки и однозначные КС-грамматики. Свойства класса детерминированных КС-языков (без доказательства).
 24. Конечные преобразователи и задаваемые ими преобразования. Различные варианты определения. Примеры конечных преобразований. Теорема Нива.
 25. Замкнутость конечных преобразований относительно композиции.
 26. Замкнутость классов автоматных и КС-языков относительно конечных преобразований.
 27. Лемма о разрастании для конечных преобразований. Примеры соответствий, не задаваемых конечными преобразованиями.
 28. Теорема Хомского-Шютценберже.
 29. Линейные и полулинейные множества. Теорема Парика.
- Пример экзаменационного билета:
30. Минимальный детерминированный автомат. Его единственность.
 31. Алгоритм Кока-Янгера-Касами. Доказательство корректности. Сложность алгоритма Кока-Янгера-Касами.

Пример экзаменационного билета:

1. Минимальный детерминированный автомат. Его единственность.
2. Алгоритм Кока-Янгера-Касами. Доказательство корректности. Сложность алгоритма Кока-Янгера-Касами.

Пример экзаменационной задачи:

Доказать, что в определении автомата с магазинной памятью можно не требовать очистки стека.

Пример задачи повышенной сложности:

Привести пример неавтоматного языка, удовлетворяющего лемме о разрастании для автоматного языка.

Критерии оценивания

отлично (10) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений;

отлично (9) - выставляется студенту, показавшему свободное оперирование знаниями учебной программы дисциплины, выполнение заданий творческого характера;

отлично (8) - выставляется студенту, показавшему владение программным учебным материалом с наличием несущественных ошибок в действиях, самостоятельно исправляемых учащимся;

хорошо (7) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;

хорошо (6) - выставляется студенту если он осознает воспроизведение программного учебного материала, в том числе и различной степени сложности, с несущественными ошибками, затруднения в применении отдельных навыков;

хорошо (5) - выставляется студенту если теоретическое содержание освоено не полностью, некоторые практические навыки сформированы недостаточно, в некоторых случаях были допущены ошибки;

удовлетворительно (4) - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;

удовлетворительно (3) - выставляется студенту в случае большого количества недочетов и неправильных ответов, а также пассивной работе в ходе занятий, многие учебные задания не выполнены;

неудовлетворительно (2) - выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач;

неудовлетворительно (1) - выставляется студенту, который не освоил теоретическое и практическое содержание курса, все выполненные учебные задания содержат грубые ошибки.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Дифференцированный зачет проводится в устной форме. Во время проведения обучающиеся могут пользоваться программой дисциплины, а также справочной литературой и вычислительной техникой.