

МИНОБРНАУКИ РОССИИ



Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«Российский государственный гуманитарный университет»
(ФГБОУ ВО «РГГУ»)**

ОТДЕЛЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ В ГУМАНИТАРНОЙ СФЕРЕ

Кафедра математики, логики и интеллектуальных систем в гуманитарной сфере

МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА

45.03.04 Интеллектуальные системы в гуманитарной сфере

Разработка и программирование интеллектуальных систем в гуманитарной сфере

Уровень высшего образования: бакалавриат

Форма обучения: очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2024

Математическая лингвистика

Рабочая программа дисциплины

Составитель:

Доктор физико-математических наук, доцент

М. Р. Пентус

.....

УТВЕРЖДЕНО

Протокол заседания кафедры МЛиИС

№ 6 от 08.02.2024

ОГЛАВЛЕНИЕ

1. Error! Bookmark not defined.	
1.1. Error! Bookmark not defined.	
1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций	4
1.3. Место дисциплины в структуре образовательной программы	6
2. Error! Bookmark not defined.	
3. Error! Bookmark not defined.	
4. Error! Bookmark not defined.	
5. Error! Bookmark not defined.	
5.1 Error! Bookmark not defined.	
5.2 Error! Bookmark not defined.	
5.3 Error! Bookmark not defined.	
6. Error! Bookmark not defined.	
6.1 Error! Bookmark not defined.	
6.2 Error! Bookmark not defined.	
7. Error! Bookmark not defined.	
8. Error! Bookmark not defined.	
9. Error! Bookmark not defined.	
9.1 Error! Bookmark not defined.	
9.2 Error! Bookmark not defined.	20
Приложение 1. Аннотация дисциплины	21

1. Пояснительная записка

1.1. Цели и задачи дисциплины

Цель дисциплины: подготовка специалиста, знающего основные понятия математической лингвистики, владеющего основными методами математической лингвистики, а также основными методами и алгоритмами решения задач математической лингвистики.

Задачи дисциплины:

- знакомство с основными понятиями математической лингвистики, необходимых специалисту в области программирования и разработки информационных и интеллектуальных систем;
- изучение теоретического аппарата математической лингвистики;
- приобретение навыков решения задач математической лингвистики;
- развитие навыков применения изученного математического аппарата к решению практических задач в области программирования и разработки информационных и интеллектуальных систем.

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Компетенция (код и наименование)	Индикаторы компетенций (код и наименование)	Результаты обучения
ОПК-1. Способен применять в профессиональной деятельности методы математического анализа, логики и моделирования, теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках	ОПК -1.2. Способен использовать математические методы для построения моделей в информатике, лингвистике и некоторых гуманитарных дисциплинах.	Знать: определение классов иерархии Хомского; основные свойства классов иерархии Хомского; основные методы категориальной грамматики. Уметь: доказывать основные теоремы, входящие в данный курс. Владеть: навыками детерминизации конечного

		автомата; определения, к какому классу данный язык принадлежит; преобразования конечного автомата в регулярное выражение.
ОПК-2. Способен к профессиональному росту и самосовершенствованию в области гуманитарных, социальных и лингвистических наук, а также в сфере техники и технологии информатики.	ОПК-2.2. Пользуется современными справочными и библиотечными системами и системами дистанционного образования.	Знать: базовую математическую терминологию данной области. Уметь: использовать методы теории формальных языков и грамматик при решении математических и лингвистических проблем как в научно-исследовательской деятельности, так и в рамках прикладных задач по лингвистическому обеспечению автоматических систем обработки текстов. Владеть: навыками построения и исследования простых математических моделей естественных языков с использованием методов математической лингвистики.

1.3. Место дисциплины в структуре основной образовательной программы

Дисциплина «Математическая лингвистика» относится к обязательной части, формируемой участниками образовательных отношений базовой части блока дисциплин учебного плана.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик:

- знание основных понятий математической логики и алгебры;
- умение использовать методы математической логики;
- владение основными методами линейной алгебры.

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин и прохождения практик: логическое программирование, интеллектуальные системы, базы данных, интеллектуальный анализ данных и машинное обучение, введение в робототехнику.

2. Структура дисциплины

Общая трудоемкость дисциплины составляет 3 з.е., 108 академических часов.

Структура дисциплины для очной формы обучения

Объем дисциплины в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Семестр	Тип учебных занятий	Количество часов
5	Лекции	10
5	Семинары	32
Всего:		42

Объем дисциплины в форме самостоятельной работы обучающихся составляет 66 академических часов.

3. Содержание дисциплины

№ п/п	Наименование раздела дисциплины	Содержание
1	Синтаксические теории и формальные языки	Формальный язык как множество цепочек над некоторым конечным алфавитом. Понятие о формальной грамматике как способе описания формального языка. Порождающие грамматики Хомского. Иерархия Хомского и иерархия распознающих автоматов.

2	Конечные автоматы и регулярные выражения	<p>Понятие о конечном автомате как способе задания формального языка. Детерминированные и недетерминированные конечные автоматы.</p> <p>Регулярные выражения как способ задания регулярных языков.</p> <p>Необходимое условие принадлежности языка классу автоматных языков.</p>
3	Контекстно-свободные языки	<p>Нормальные формы контекстно-свободных грамматик.</p> <p>Нормальные формы Хомского и Грейбах. Преобразование контекстно-свободной грамматики в нормальную форму Хомского.</p> <p>Необходимое условие принадлежности языка классу контекстно-свободных языков.</p>

4. Образовательные технологии

№ п/п	Наименование раздела	Виды учебных занятий	Образовательные технологии
1	2	3	4
1	Синтаксические теории и формальные языки	<p>Лекции 1-2.</p> <p>Семинар 1</p> <p>Семинар 2</p>	<p>Проблемная лекция.</p> <p>Обсуждение теоретического материала. Практикум по решению задач.</p> <p>Обсуждение теоретического материала. Практикум по решению задач.</p>
2	Конечные автоматы и	Лекция 3-4	Проблемная лекция.

	регулярные выражения	Семинар 3 Семинар 4 Семинар 5 Семинар 6	Обсуждение теоретического материала. Практикум по решению задач. Обсуждение теоретического материала. Практикум по решению задач. Обсуждение теоретического материала. Практикум по решению задач. Обсуждение теоретического материала. Практикум по решению задач.
3	Контекстно- свободные языки	Лекция 5-6 Семинар 7 Семинар 8	Проблемная лекция. Обсуждение теоретического материала. Практикум по решению задач. Обсуждение теоретического материала. Практикум по решению задач.

В период временного приостановления посещения обучающимися помещений и территории РГГУ для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

– видео-лекции;

- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;
- консультации с использованием телекоммуникационных средств.

5. Оценка планируемых результатов обучения

5.1. Система оценивания

<i>Форма контроля</i>	<i>Макс. количество баллов</i>	
	<i>За одну работу</i>	<i>Всего</i>
Текущий контроль:		
• Домашнее задание	3 балла	30 баллов
• Контрольная работа	30 баллов	30 баллов
Промежуточная аттестация (экзамен)		40 баллов
Итого за семестр (дисциплину)		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

5.2. Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ А,В	отлично/ зачтено	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	хорошо/ зачтено	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	удовлетво- рительно/ зачтено	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	неудовлет- ворительно/ не зачтено	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами.</p> <p>Демонстрирует фрагментарные знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.</p>

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

Примеры домашних заданий (первый блок)

1. Описать язык, порождаемый грамматикой $S \rightarrow aS, S \rightarrow bS, S \rightarrow cT, T \rightarrow bT, T \rightarrow a$.
2. Найти грамматику, порождающую язык $\{a^n b^m c^m \mid n > 0, m > 0\}$.
3. Какому классу принадлежит грамматика $F \rightarrow aFb, A \rightarrow bFa, F \rightarrow FF, F \rightarrow c$?
4. Найти праволинейную грамматику, эквивалентную грамматике $F \rightarrow FJ, J \rightarrow JF, F \rightarrow ab, J \rightarrow ba$.
5. Найти конечный автомат, распознающий язык $\{w \mid |w|_a < 3, |w|_b > 3\}$.
6. Существуют ли автоматный язык, который не распознаётся никаким конечным автоматом, содержащим только переходы с метками длины единица и имеющим ровно одно начальное состояние и ровно одно заключительное состояние?
7. Найти детерминированный конечный автомат, распознающий язык $\{vau \mid |v|_b \text{ чётно}, |u|_b \text{ нечётно}\}$.
8. Является ли автоматным язык $\{a^k b^m a^n \mid k < n \text{ или } m = 0\}$?
9. Является ли автоматным язык $\{a^k b^m a^n \mid k = 0 \text{ или } m = n\}$?
10. Найти детерминированный конечный автомат, распознающий язык $(a+b)^*(aab+abaa+abb)(a+b)^*$.
11. Упростить регулярное выражение $(a+ba+bb)^*bbb(a+b)^*$.

12. Найти регулярное выражение для языка, порождаемого грамматикой $S \rightarrow cT, S \rightarrow aU, T \rightarrow bS, T \rightarrow cU, U \rightarrow bS, U \rightarrow aT, U \rightarrow c$.
13. Найти регулярное выражение для дополнения языка $(ab+ba)^*(1+a+b)$ над алфавитом $\{a,b\}$.
14. Найти минимальный полный детерминированный конечный автомат, распознающий язык $\{a^m ub^n \mid |u| \text{ нечётно}\}$.
15. Равны ли регулярные выражения $(aa+b+ab)^*$ и $((a+b)^*b+1)(aa)^*$?
16. Сколько элементов в синтаксическом моноиде языка $a^*b^*+b^*a^*$ над алфавитом $\{a,b\}$?

Примеры домашних заданий (второй блок)

17. Найти однозначную контекстно-свободную грамматику, эквивалентную грамматике $S \rightarrow aSbb, S \rightarrow aSb, S \rightarrow c$.
18. Описать язык, порождаемый грамматикой $F \rightarrow ab, F \rightarrow aFb, F \rightarrow FF$.
19. Найти контекстно-свободную грамматику в нормальной форме Хомского, эквивалентную грамматике $F \rightarrow a, F \rightarrow bF, F \rightarrow cFF$.
20. Является ли контекстно-свободным язык $\{a^k b^m a^n \mid k < m < n\}$?
21. Какому классу принадлежит язык $\{a^n b^m c^{n+m} \mid n > 1, m > 2\}$?
22. Найти автомат с магазинной памятью, распознающий язык, порождаемый грамматикой $F \rightarrow aFb, F \rightarrow aFa, F \rightarrow a$.
23. Является ли детерминированным контекстно-свободным язык $\{a^n b^{2m} ab^m a^n \mid n > 0, m \geq 0\}$?
24. Является ли детерминированным контекстно-свободным язык $\{a^k b^m c^n \mid k < \max(m,n)\}$?
25. Существует ли решение у постовской системы соответствия $((babc, b, ab, cababa), (ab, c, abab, cab))$?
26. Существует ли решение у постовской системы соответствия $((a, ca, bcc, cccacc), (ab, ccc, ccca, a))$?
27. Найти неукорачивающую грамматику, порождающую язык $\{w \mid |w|_a < |w|_b < |w|_c\}$.
28. Является ли автоматным язык, порождаемый грамматикой $S \rightarrow RT, R \rightarrow TR, R \rightarrow a, T \rightarrow TR, T \rightarrow b$?

Пример итоговой работы

1. Найти грамматику, порождающую язык $\{a^n b^n a^n b^n \mid n > 0\}$.
2. Найти минимальный полный детерминированный конечный автомат, распознающий язык $(1+(a+b)^*b)(a+b)^*$.

3. Является ли автоматным язык $\{a^k b^m a^n \mid k = n \text{ или } m > 0\}$?
4. Является ли автоматным язык $\{a(cb)^n c(bc)^n \mid n > 0\}$?
5. Упростить регулярное выражение $((bc)^*+(b+ca)^*c)^*$.
6. Найти регулярное выражение для дополнения языка $b^*a^*+b^*a(aa)^*(ba)^*(b+ba)$ над алфавитом $\{a,b\}$.
7. Равны ли регулярные выражения $(a+b)^*aa(a+b)^*bb(a+b)^*$ и $(ab+b)^*aa(a+b)^*bb(a+ba)^*$?
8. Сколько элементов в синтаксическом моноиде языка $(ab)^*(ba)^*+a^*$ над алфавитом $\{a,b\}$?
9. Является ли контекстно-свободным язык $\{a^k b^i a^m b^n \mid k < n, m < i\}$?
10. К какому классу принадлежит язык $\{a^n b^m c^{nm} \mid n > 0, m < 4\}$?
11. Является ли детерминированным контекстно-свободный язык $\{w \mid |w|_a = |w|_b\}$?

Вопросы для самоконтроля

1. Формальные языки.
2. Порождающие грамматики.
3. Классы грамматик.
4. Недетерминированные конечные автоматы.
5. Конечные автоматы с однобуквенными переходами.
6. Детерминированные конечные автоматы.
7. Преобразование конечного автомата к детерминированному виду.
8. Свойства замкнутости класса автоматных языков.
9. Пересечение и дополнение автоматных языков.
10. Лемма о разрастании для автоматных языков.
11. Примеры неавтоматных языков.
12. Определение регулярного выражения.
13. Свойства регулярных выражений.
14. Теорема Клини.
15. Множества правых контекстов.
16. Минимизация детерминированных конечных автоматов.
17. Множества двусторонних контекстов.
18. Деревья вывода.

19. Однозначные контекстно-свободные грамматики.
20. Языки Дика и Лукасевича.
21. Устранение бесполезных символов.
22. Устранение эpsilon-правил.
23. Нормальная форма Хомского.
24. Лемма о разрастании для контекстно-свободных языков.
25. Свойства замкнутости класса контекстно-свободных языков.
26. Пересечение и дополнение контекстно-свободных языков.
27. Пересечение контекстно-свободного языка с автоматным языком.

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

а) Основная литература

1. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции: в 2 т. М.: Мир, 1978. Т.1: Синтаксический анализ / пер. с англ. В. Н. Агафонова; под ред. В. М. Курочкина. 1978. 611 с.
2. Гросс М., Лантен А. Теория формальных грамматик. М., 1971.
3. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. М., 2006.
4. Пентус А. Е., Пентус М. Р. Конечные автоматы и регулярные выражения. Сборник задач. М., 2015.
5. Пентус А. Е., Пентус М. Р. Контекстно-свободные языки. Сборник задач. М., 2016.
6. Хопкрофт Дж. Э., Мотвани Р., Ульман Дж. Д. Введение в теорию автоматов, языков и вычислений, 2-е изд. М., 2002.

б) Дополнительная литература

1. Гинзбург С. Математическая теория контекстно-свободных языков. М., 1970.

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Национальная электронная библиотека (НЭБ) www.rusneb.ru

ELibrary.ru Научная электронная библиотека www.elibrary.ru

Электронная библиотека Grebennikon.ru www.grebennikon.ru

Cambridge University Press

ProQuest Dissertation & Theses Global

SAGE Journals

Taylor and Francis

JSTOR

<http://www.wolframalpha.com>

6.3. Профессиональные базы данных и информационно-справочные системы

Доступ к профессиональным базам данных: <https://liber.rsuh.ru/ru/bases>

Информационные справочные системы:

1. Консультант Плюс
2. Гарант

7. Материально-техническое обеспечение дисциплины

Для обеспечения дисциплины используется материально-техническая база образовательного учреждения: учебные аудитории, оснащённые доской, а также компьютером и проектором для демонстрации учебных материалов.

Состав программного обеспечения:

1. Windows
2. Microsoft Office

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

□ для слепых и слабовидящих: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением или могут быть заменены устным ответом; обеспечивается индивидуальное равномерное освещение не менее 300 люкс; для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств; письменные задания оформляются увеличенным шрифтом; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

□ для глухих и слабослышащих: лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования; письменные задания выполняются на компьютере в письменной форме; экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.

□ для лиц с нарушениями опорно-двигательного аппарата: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих: в печатной форме увеличенным шрифтом, в форме электронного документа, в форме аудиофайла.
- для глухих и слабослышащих: в печатной форме, в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата: в печатной форме, в форме электронного документа, в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих: устройством для сканирования и чтения с камерой SARA CE; дисплеем Брайля PAC Mate 20; принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих: автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих; акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата: передвижными, регулируемые эргономическими партами СИ-1; компьютерной техникой со специальным программным обеспечением.

9. Методические материалы

9.1. Планы семинарских занятий

Тема 1 (4 ч) Синтаксические теории и формальные языки

Цель занятия: освоение иерархии Хомского

Форма проведения – обсуждение, решение упражнений

Вопросы для обсуждения

1. Как можно задать бесконечный язык?
2. Какие свойства языков важны для компьютерных приложений?

Контрольные вопросы

1. Все ли языки порождаются грамматиками?
2. Описать язык, порождаемый данной грамматикой.

3. К какому классу принадлежит данная грамматика?

Список источников и литературы:

1. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. М., 2006.
2. Гросс М., Лантен А. Теория формальных грамматик. М., 1971.
3. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции: в 2 т. М.: Мир, 1978. Т.1: Синтаксический анализ / пер. с англ. В. Н. Агафонова; под ред. В. М. Курочкина. 1978. 611 с.
4. Пентус А. Е., Пентус М. Р. Контекстно-свободные языки. Сборник задач. М., 2016.

Материально-техническое обеспечение занятия: академическая аудитория с доской.

Тема 2 (8 ч) Конечные автоматы и регулярные выражения

Цель занятий: отработать навыки детерминизации конечного автомата, выработать навыки доказательства неавтоматности, отработать навыки построения регулярного выражения по конечному автомату и наоборот.

Форма проведения – обсуждение, решение упражнений

Вопросы для обсуждения

1. Какие свойства конечных автоматов важны для компьютерных приложений?
2. Можно ли доказать автоматность с помощью леммы о разрастании?
3. Какие критерий автоматности доказаны в данном курсе?

Контрольные вопросы

1. Найти детерминированный конечный автомат для данного языка?
2. Является ли автоматным данный язык?
3. Найти регулярное выражение для дополнения данного языка?

Список источников и литературы:

1. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. М., 2006.
2. Пентус А. Е., Пентус М. Р. Конечные автоматы и регулярные выражения. Сборник задач. М., 2015.
3. Хопкрофт Дж. Э., Мотвани Р., Ульман Дж. Д. Введение в теорию автоматов, языков и вычислений, 2-е изд. М., 2002.

Материально-техническое обеспечение занятия: академическая аудитория с доской.

Тема 3 (4 ч) Контекстно-свободные языки

Цель занятий: выработать навыки преобразования контекстно-свободной грамматики к нормальной форме Хомского, выработать навыки доказательства того, что данный язык не является контекстно-свободным.

Форма проведения – обсуждение, решение упражнений

Вопросы для обсуждения

1. Какие свойства контекстно-свободных языков важны для компьютерных приложений?
2. Можно ли доказать контекстную свободу с помощью леммы о разрастании?

Контрольные вопросы

1. Привести данную грамматику к нормальной форме Хомского.
2. Является ли контекстно-свободным данный язык?

Список источников и литературы:

1. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. М., 2006.
2. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции: в 2 т. М.: Мир, 1978. Т.1: Синтаксический анализ / пер. с англ. В. Н. Агафонова; под ред. В. М. Курочкина. 1978. 611 с.
3. Пентус А. Е., Пентус М. Р. Контекстно-свободные языки. Сборник задач. М., 2016.

Материально-техническое обеспечение занятия: академическая аудитория с доской.

9.2. Методические рекомендации по подготовке письменных работ

При подготовке к письменной контрольной работе и к итоговой работе необходимо:

- 1) изучить теоретический материал, который обсуждался на занятиях;
- 2) выполнить домашние задания по пройденным темам.

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Дисциплина реализуется кафедрой математики, логики и интеллектуальных систем в гуманитарной сфере в пятом семестре.

Цель дисциплины: подготовка специалиста, знающего основные понятия математической лингвистики, владеющего основными методами математической лингвистики, а также основными методами и алгоритмами решения задач математической лингвистики.

Задачи дисциплины: знакомство с основными понятиями математической лингвистики, необходимых специалисту в области программирования и разработки информационных и интеллектуальных систем; изучение теоретического аппарата математической лингвистики; приобретение навыков решения задач математической лингвистики; развитие навыков применения изученного математического аппарата к решению практических задач в области программирования и разработки информационных и интеллектуальных систем.

В результате освоения дисциплины обучающийся должен:

Знать:

- определение классов иерархии Хомского;
- основные свойства классов иерархии Хомского;
- основные методы категориальной грамматики;
- базовую математическую терминологию данной области.

Уметь:

- доказывать основные теоремы, входящие в данный курс,
- использовать методы теории формальных языков и грамматик при решении математических и лингвистических проблем как в научно-исследовательской деятельности, так и в рамках прикладных задач по лингвистическому обеспечению автоматических систем обработки текстов.

Владеть навыками:

- детерминизации конечного автомата;
- определения, к какому классу данный язык принадлежит;
- преобразования конечного автомата в регулярное выражение;
- построения и исследования простых математических моделей естественных языков с использованием методов математической лингвистики.

По дисциплине предусмотрена промежуточная аттестация в форме экзамена.

Общая трудоемкость освоения дисциплины составляет 3 зачетные единицы.