

МИНОБРАЗОВАНИЯ 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

Название дисциплины: Математическая лингвистика

Рабочая программа дисциплины

Название и код направления подготовки:

Интеллектуальные системы в гуманитарной сфере – 45.03.04

Название профиля: Разработка и программирование интеллектуальных систем

Квалификация выпускника: бакалавр

Форма обучения: очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2019

Название дисциплины: Математическая лингвистика
Рабочая программа дисциплины
Составитель:
Доктор физико-математических наук, доцент, профессор
М.Р. Пентус

УТВЕРЖДЕНО
Протокол заседания кафедры МЛиИС
№ 7 от 10.06.2019

ОГЛАВЛЕНИЕ

1. Пояснительная записка

1.1 Цель и задачи дисциплины

1.2. Перечень планируемых результатов обучения по дисциплине (*модулю*), соотнесенных с индикаторами достижения компетенций

1.3. Место дисциплины в структуре образовательной программы

2. Структура дисциплины

3. Содержание дисциплины

4. Образовательные технологии

5. Оценка планируемых результатов обучения

5.1. Система оценивания

5.2. Критерии выставления оценок

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

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

7. Материально-техническое обеспечение дисциплины

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья

9. Методические материалы

9.1. Планы семинарских занятий

9.2. Иные материалы

Приложения

Приложение 1. Аннотация дисциплины

Приложение 1. Лист изменений

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

1. Пояснительная записка

1.1. Цель и задачи дисциплины

Цель дисциплины: ознакомление студентов с математическими теориями, лежащими в основе формального описания синтаксиса и семантики естественных и искусственных языков, в том числе языков программирования, и формирование у студентов навыков математического моделирования с использованием средств современной теории формальных языков и грамматик.

Задачи дисциплины:

- формирование у студентов системы понятий и навыков, необходимых для дальнейшего углублённого изучения теоретических основ и практических методов построения систем искусственного интеллекта;
- изучение теоретического аппарата математической лингвистики;
- приобретение навыков решения задач математической лингвистики;
- развитие навыков применения изученного математического аппарата к решению практических задач.

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

| Компетенция (код и наименование) | Индикаторы компетенций (код и наименование) | Результаты обучения |
|--|---|---|
| ОПК-1. Способен использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа, логики и моделирования, теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках | ОПК-1.2. Умеет использовать математические методы для построения моделей в информатике, лингвистике и некоторых гуманитарных дисциплинах. | Знать: определение классов иерархии Хомского; основные свойства классов иерархии Хомского. 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 получать знания в области современных проблем науки, техники и технологии информатики, гуманитарных, лингвистических, и социальных наук | ОПК-2.2. Умеет пользоваться 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» является базовой частью Б1.О.13 блока Б1 дисциплин учебного плана по направлению подготовки 45.03.04 (интеллектуальные системы в гуманитарной сфере). Дисциплина реализуется на отделении интеллектуальных систем в гуманитарной сфере кафедрой математики, логики и интеллектуальных систем в гуманитарной сфере в пятом семестре.

Для освоения дисциплины необходимы компетенции, сформированные в ходе изучения следующих дисциплин и прохождения практик: дискретная математика в объёме 1 семестра, а также математика в объёме средней школы.

В результате освоения дисциплины формируются компетенции, необходимые для изучения следующих дисциплин и прохождения практик: вычислительная лингвистика.

2. Структура дисциплины (модуля) для очной формы обучения

Общая трудоёмкость дисциплины составляет 2 з.е., 72 ч., в том числе контактная работа обучающихся с преподавателем 28 ч., промежуточная аттестация 0 ч., самостоятельная работа обучающихся 44 ч.

| № п/п | Раздел дисциплины/темы | С е м е с тр | Виды учебной работы (в часах) | | | | | Промежуточная аттестация | С а м о с то я т е л ь н а я р а б о т а | Формы текущего контроля успеваемости, форма промежуточной аттестации <i>(по семестрам)</i> |
|-------|--|--------------|-------------------------------|----------|----------------------|--|----------------------|--------------------------|--|--|
| | | | Контактная | | | | Лабораторные занятия | | | |
| | | | Лекции | Семинары | Практические занятия | | | | | |
| 1 | Синтаксические теории и формальные языки | 5 | 4 | 4 | | | | 8 | Оценка выполнения практических заданий | |
| 2 | Конечные | 5 | 6 | 6 | | | | 12 | Оценка | |

| | | | | | | | | | |
|---|---------------------------------|---|----|----|--|--|--|----|--|
| | автоматы и регулярные выражения | | | | | | | | выполнения практических заданий |
| 3 | Контекстно-свободные языки | 5 | 4 | 4 | | | | 8 | Оценка выполнения практических заданий |
| | Зачет | 5 | | | | | | 16 | Решение задач контрольных работ |
| | Итого | | 14 | 14 | | | | 44 | |

3. Содержание дисциплины

| № | Наименование раздела дисциплины | № | Содержание |
|---|--|-----|---|
| 1 | Синтаксические теории и формальные языки | | |
| | | 1.1 | Формальный язык как множество цепочек над некоторым конечным алфавитом. Понятие о формальной грамматике как способе описания формального языка. |
| | | 1.2 | Порождающие грамматики Хомского. |
| | | 1.3 | Иерархия Хомского и иерархия распознающих автоматов. |
| 2 | Конечные автоматы и регулярные выражения | | |
| | | 2.1 | Понятие о конечном автомате как способе задания формального языка. Детерминированные и недетерминированные конечные автоматы. |
| | | 2.2 | Регулярные выражения как способ задания регулярных языков. |
| | | 2.3 | Необходимое условие принадлежности языка классу автоматных языков. |
| 3 | Контекстно-свободные языки | | |
| | | 3.1 | Нормальные формы контекстно-свободных грамматик. Нормальные формы Хомского и Грейбах. Преобразование контекстно-свободной грамматики в нормальную форму Хомского. |
| | | 3.2 | Необходимое условие принадлежности языка классу контекстно-свободных языков. |

4. Образовательные технологии

Образовательные технологии

| № п/п | Наименование раздела | Виды учебной работы | Информационные и образовательные технологии |
|-------|----------------------|---------------------|---|
| 1 | 2 | 3 | 5 |
| 1 | Синтаксические | Лекция 1. | Вводная лекция-беседа. |

| | | | |
|---|--|---|---|
| | теории и формальные языки | Семинар 1. Лекция 2 Семинар 2 Самостоятельная работа | Семинар-обсуждение Теоретическая лекция. Семинар-обсуждение. Практикум по решению задач. Работа с электронным конспектом и интернет-ресурсами. Консультирование посредством электронной почты |
| 2 | Конечные автоматы и регулярные выражения | Лекция 3 Семинар 3 Лекция 4 Семинар 4 Лекция 5 Семинар 5 Самостоятельная работа | Теоретическая лекция. Семинар-обсуждение. Практикум по решению задач Теоретическая лекция. Семинар-обсуждение. Практикум по решению задач. Теоретическая лекция. Семинар-обсуждение. Работа с электронным конспектом, электронным задачником и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты |
| 3 | Контекстно-свободные языки | Лекция 6 Семинар 6 Лекция 7 Семинар 7 Самостоятельная работа | Теоретическая лекция. Семинар-обсуждение. Практикум по решению задач. Теоретическая лекция. Семинар-обсуждение. Практикум по решению задач. Работа с электронным конспектом, электронным задачником и интернет-ресурсами. Консультирование и приём домашних заданий посредством электронной почты |

5. Оценка планируемых результатов обучения

5.1. Система оценивания

| <i>Форма контроля</i> | <i>Срок отчетности</i> | <i>Макс. количество баллов</i> | |
|----------------------------------|------------------------|--------------------------------|--------------|
| | | <i>За одну работу</i> | <i>Всего</i> |
| Текущий контроль: | | | |
| • опрос | 3—15 недели | 5 баллов | 20 баллов |
| • дом. задание (темы 2.1—2.3) | 6—10 недели | 20 баллов | 20 баллов |
| • дом. задание (темы 3.1—3.3) | 11—13 недели | 20 баллов | 20 баллов |
| Промежуточная аттестация (зачет) | 14 неделя | | 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 (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. 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 (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 | «удовлетвори-тельно»/ «зачтено (удовлетвори-тельно)»/ «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</p> |

| | | |
|---------------|--------------------------------------|---|
| | | <p>аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p> |
| 49-0/ F,FX | «неудовлетворительно»/ 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

Типовые практические задания к зачёту

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 $(1+(a+b)^*b)b(a+b)^*$.
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-свободным язык $\{w \mid |w|_a = |w|_b\}$?

Список теоретических вопросов

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.

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 $(a+ba+bba)^*bbb(a+b)^*$.
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 $\{a^m u b^n \mid |u| \text{ нечётно}\}$.
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. К какому классу принадлежит язык $\{a^n b^m c^{n+m} \mid n > 1, m > 2\}$?
22. Найти автомат с магазинной памятью, распознающий язык, порожаемый грамматикой $F \rightarrow aFb, F \rightarrow aFa, F \rightarrow a$.
23. Является ли детерминированным контекстно-свободный язык $\{a^n b^{2m} a^m \mid n > 0, m \geq 0\}$?
24. Является ли детерминированным контекстно-свободный 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, порожающую язык $\{w \mid |w|_a < |w|_b < |w|_c\}$.
28. Является ли автоматным язык, порожаемый грамматикой $S \rightarrow RT, R \rightarrow TR, R \rightarrow a, T \rightarrow TR, T \rightarrow b$?

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. *Пентус А. Е., Пентус М. Р.* Математическая теория формальных языков. М., 2006.
3. *Пентус А. Е., Пентус М. Р.* Конечные автоматы и регулярные выражения. Сборник задач. М., 2015.
4. *Пентус А. Е., Пентус М. Р.* Контекстно-свободные языки. Сборник задач. М., 2016.

6.2.Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимый для освоения дисциплины

<http://lpcs.math.msu.su/~pentus/ftp/tvi/tfyaza.pdf>

https://ru.wikipedia.org/wiki/Иерархия_Хомского

Перечень БД и ИСС

| №п /п | Наименование |
|-------|--|
| 1 | Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2019 г. Web of Science Scopus |
| 2 | Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2019 г. |

| | |
|---|---|
| | Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis |
| 3 | Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru |
| 4 | Компьютерные справочные правовые системы Консультант Плюс, Гарант |

7. Материально-техническое обеспечение дисциплины

Академическая аудитория с большой доской.

1. Перечень ПО

| №п /п | Наименование ПО | Производитель | Способ распространения (лицензионное или свободно распространяемое) |
|-------|-----------------------------|------------------|---|
| 1 | Adobe Master Collection CS4 | Adobe | лицензионное |
| 2 | ОС «Альт Образование» 8 | ООО «Базальт СПО | лицензионное |
| 3 | Windows 10 Pro | Microsoft | лицензионное |
| 4 | Kaspersky Endpoint Security | Kaspersky | лицензионное |
| 5 | Microsoft Office 2016 | Microsoft | лицензионное |

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья

При необходимости рабочая программа дисциплины может быть адаптирована для обеспечения образовательного процесса лицам с ограниченными возможностями здоровья, в том числе для дистанционного обучения. Для этого от студента требуется представить заключение психолого-медико-педагогической комиссии (ПМПК) и личное заявление (заявление законного представителя).

В заключении ПМПК должно быть прописано:

- рекомендуемая учебная нагрузка на обучающегося (количество дней в неделю, часов в день);
- оборудование технических условий (при необходимости);
- сопровождение и (или) присутствие родителей (законных представителей) во время учебного процесса (при необходимости);
- организация психолого-педагогического сопровождение обучающегося с указанием специалистов и допустимой нагрузки (количества часов в неделю).

Для осуществления процедур текущего контроля успеваемости и промежуточной аттестации обучающихся при необходимости могут быть созданы фонды оценочных средств, адаптированные для лиц с ограниченными возможностями здоровья и позволяющие оценить достижение ими запланированных в основной образовательной программе результатов обучения и уровень сформированности всех компетенций, заявленных в образовательной программе.

Форма проведения текущей и итоговой аттестации для лиц с ограниченными возможностями здоровья устанавливается с учетом индивидуальных психофизических особенностей (устно, письменно (на бумаге, на компьютере), в форме тестирования и т.п.). При необходимости студенту предоставляется дополнительное время для подготовки ответа на зачете или экзамене.

9. Методические материалы

9.1. Планы семинарских занятий

7.1.

Тема 1 (2 ч.) Порождающие грамматики Хомского

Цель занятия: проиллюстрировать порождающие грамматики Хомского.

Форма проведения – обсуждение, опрос.

Разделы 1.1, 1.2, 1.4 из «Математическая теория формальных языков».

Задачи 1.1.1-1.2.10 из «Контекстно-свободные языки. Сборник задач».

Вопросы для обсуждения:

1. Как можно задать бесконечный язык?

Контрольные вопросы:

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

Список источников и литературы:

1. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. М., 2006.
2. Гросс М., Ланген 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.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Иерархия_Хомского

Материально-техническое обеспечение занятия: доска.

Тема 2 (2 ч.) Иерархия Хомского

Цель занятия: освоение иерархии Хомского.

Форма проведения – опрос, решение задач.

Разделы 1.5, 15.1 из «Математическая теория формальных языков».

Задачи 1.5.1-1.5.20 из «Контекстно-свободные языки. Сборник задач».

Вопросы для обсуждения:

1. Какие свойства языков важны для компьютерных приложений?

Контрольные вопросы:

1. К какому классу принадлежит данная грамматика?

Список источников и литературы:

1. Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. М., 2006.
2. Пентус А. Е., Пентус М. Р. Контекстно-свободные языки. Сборник задач. М., 2016.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Иерархия_Хомского

Материально-техническое обеспечение занятия: доска.

Тема 3 (2 ч.) Детерминированные и недетерминированные конечные автоматы

Цель занятия: отработать навыки детерминизации конечного автомата.

Форма проведения – опрос, решение задач.

Задачи 2.1.1-3.2.10 из «Конечные автоматы и регулярные выражения. Сборник задач».

Вопросы для обсуждения:

1. Какие свойства конечных автоматов важны для компьютерных приложений?

Контрольные вопросы:

1. Найти детерминированный конечный автомат для 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.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Иерархия_Хомского

Материально-техническое обеспечение занятия: доска.

Тема 4 (2 ч.) Лемма о разрастании для автоматных языков

Цель занятия: выработать навыки доказательства неавтоматности.

Форма проведения – опрос, решение задач.

Задачи 3.3.1-3.4.14 из «Конечные автоматы и регулярные выражения. Сборник задач».

Вопросы для обсуждения:

1. Можно ли доказать автоматность с помощью леммы о разрастании?

Контрольные вопросы:

1. Является ли автоматным данный язык?

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.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Регулярная_грамматика

Материально-техническое обеспечение занятия: доска.

Тема 5 (2 ч.) Регулярные выражения

Цель занятия: отработать навыки построения регулярного выражения по конечному автомату и наоборот.

Форма проведения – опрос, решение задач.

Задачи 5.1.1-5.3.38 из «Конечные автоматы и регулярные выражения. Сборник задач».

Вопросы для обсуждения:

1. Какие критерий автоматности доказаны в данном курсе?

Контрольные вопросы:

1. 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. 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.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Регулярная_грамматика

Материально-техническое обеспечение занятия: доска.

Тема 6 (2 ч.) Нормальные формы контекстно-свободных грамматик

Цель занятия: выработать навыки преобразования контекстно-свободной грамматики к

нормальной форме Хомского.

Форма проведения – опрос, решение задач.

Задачи 3.1.1-3.3.9 из «Контекстно-свободные языки. Сборник задач».

Вопросы для обсуждения:

1. Какие свойства контекстно-свободных языков важны для компьютерных приложений?

Контрольные вопросы:

1. Привести данную грамматику к нормальной форме Хомского.

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.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Контекстно-свободная_грамматика

Материально-техническое обеспечение занятия: доска.

Тема 7 (2 ч.) Лемма о разрастании для контекстно-свободных языков

Цель занятия: выработать навыки доказательства того, что данный язык не является контекстно-свободным.

Форма проведения – опрос, решение задач.

Задачи 4.1.1-4.1.69 из «Контекстно-свободные языки. Сборник задач».

Вопросы для обсуждения:

1. Можно ли доказать контекстную свободу с помощью леммы о разрастании

Контрольные вопросы:

1. 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.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»:

https://ru.wikipedia.org/wiki/Контекстно-свободная_грамматика

Материально-техническое обеспечение занятия: доска.

9.2. Иные материалы

7.1.

| Вид работы | Содержание (перечень вопросов) | Трудоемкость Самост. работы (в часах) | Рекомендации |
|---|---|--|---|
| Раздел № 1: Синтаксические теории и формальные языки | | | |
| Освоение материала лекции № 1 | Формальный язык как множество цепочек над | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 1.1-1.2 |

| | | | |
|---|---|----|--|
| | некоторым конечным алфавитом. Понятие о формальной грамматике как способе описания формального языка. | | |
| Освоение материала лекции № 2 | Иерархия Хомского и иерархия распознающих автоматов. | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 1.4-1.5 |
| <i>Итого по разделу</i> | | 8 | |
| Раздел № 2: Конечные автоматы и регулярные выражения | | | |
| Освоение материала лекции № 3 | Детерминированные и недетерминированные конечные автоматы. | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 2.1, 2.3, 2.6, 2.7, 3.1, 3.2 |
| Освоение материала лекции № 4 | Лемма о разрастании для автоматных языков | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 3.3, 3.4 |
| Освоение материала лекции № 5 | Регулярные выражения как способ задания регулярных языков. | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 5.1, 5.3 |
| <i>Итого по разделу</i> | | 12 | |
| Раздел № 3: Контекстно-свободные языки | | | |
| Освоение материала лекции № 6 | Нормальные формы контекстно-свободных грамматик. | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 8.1-8.3 |
| Освоение материала лекции № 7 | Необходимое условие принадлежности языка классу контекстно-свободных языков. | 4 | Пентус А. Е., Пентус М. Р. Математическая теория формальных языков. 9.1, 9.4-9.6 |
| Подготовка к зачету | | 16 | Все материалы курса |
| <i>Итого по</i> | | 24 | |



| | | | |
|----------------------------|--|----|--|
| <i>разделу</i> | | | |
| <i>Итого по дисциплине</i> | | 44 | |

АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина «Математическая лингвистика» является базовой частью Б1.О.13 блока Б1 дисциплин учебного плана по направлению подготовки 45.03.04 (интеллектуальные системы в гуманитарной сфере). Дисциплина реализуется на отделении интеллектуальных систем в гуманитарной сфере кафедрой математики, логики и интеллектуальных систем в гуманитарной сфере в пятом семестре.

Цель дисциплины: ознакомление студентов с математическими теориями, лежащими в основе формального описания синтаксиса и семантики естественных и искусственных языков, в том числе языков программирования, и формирование у студентов навыков математического моделирования с использованием средств современной теории формальных языков и грамматик.

Задачи:

- формирование у студентов системы понятий и навыков, необходимых для дальнейшего углублённого изучения теоретических основ и практических методов построения систем искусственного интеллекта;
- изучение теоретического аппарата математической лингвистики;
- приобретение навыков решения задач математической лингвистики;
- развитие навыков применения изученного математического аппарата к решению практических задач.

Дисциплина направлена на формирование следующих компетенций:

- ПК-15 – способность применять новые информационные технологии в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления знаний.

В результате освоения дисциплины обучающийся должен:

Знать:

- определение классов иерархии Хомского;
- основные свойства классов иерархии Хомского;
- 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.

Рабочей программой предусмотрены следующие виды контроля: текущий контроль успеваемости в форме устных ответов у доски и выполнения письменных домашних заданий, промежуточная аттестация в форме зачета.



*Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы,
72 часа.*

ЛИСТ ИЗМЕНЕНИЙ

| № | Текст актуализации или прилагаемый к РПД документ, содержащий изменения | Дата | № протокола |
|---|---|-------------|-------------|
| 1 | Приложение №1 | 08.06.2020г | 6 |

1. Структура дисциплины (к п. 2 РПД на 2020)

Структура дисциплины (модуля) для очной формы обучения

Общая трудоёмкость дисциплины составляет 2 з.е., 76 ч., в том числе контактная работа обучающихся с преподавателем 28 ч., промежуточная аттестация 0 ч., самостоятельная работа обучающихся 48 ч.

| № п/п | Раздел дисциплины/темы | С е м е с тр | Виды учебной работы (в часах) | | | | | Формы текущего контроля успеваемости, форма промежуточной аттестации (по семестрам) | |
|-------|--|--------------|-------------------------------|----------|----------------------|----------------------|--------------------------|---|--|
| | | | Контактная | | | | Промежуточная аттестация | | Самостоятельная работа |
| | | | Лекции | Семинары | Практические занятия | Лабораторные занятия | | | |
| 1 | Синтаксические теории и формальные языки | 5 | 4 | 4 | | | | 8 | Оценка выполнения практических заданий |
| 2 | Конечные автоматы и регулярные выражения | 5 | 6 | 6 | | | | 12 | Оценка выполнения практических заданий |
| 3 | Контекстно-свободные языки | 5 | 4 | 4 | | | | 8 | Оценка выполнения практических заданий |
| | Зачет | 5 | | | | | | 20 | Решение задач контрольных работ |
| | Итого | | 14 | 14 | | | | 48 | |

2. Образовательные технологии (к п.4 на 2020 г.)

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.

3. Перечень БД и ИСС (к п. 6.2 на 2020 г.)

| № п/п | Наименование |
|-------|--|
| 1 | Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus |
| 2 | Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis |
| 3 | Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru |
| 4 | Компьютерные справочные правовые системы Консультант Плюс, Гарант |

4. Состав программного обеспечения (ПО) (к п. 7 на 2020 г.)

| № п/п | Наименование ПО | Производитель | Способ распространения (лицензионное или свободно распространяемое) |
|-------|-----------------------------|---------------|---|
| 1 | Adobe Master Collection CS4 | Adobe | лицензионное |
| 2 | Microsoft Office 2010 | Microsoft | лицензионное |
| 3 | Windows 7 Pro | Microsoft | лицензионное |
| 4 | AutoCAD 2010 Student | Autodesk | свободно распространяемое |

| | | | |
|----|-----------------------------|------------------|---------------------------|
| 5 | Archicad 21 Rus Student | Graphisoft | свободно распространяемое |
| 6 | SPSS Statistics 22 | IBM | лицензионное |
| 7 | Microsoft Share Point 2010 | Microsoft | лицензионное |
| 8 | SPSS Statistics 25 | IBM | лицензионное |
| 9 | Microsoft Office 2013 | Microsoft | лицензионное |
| 10 | ОС «Альт Образование» 8 | ООО «Базальт СПО | лицензионное |
| 11 | Microsoft Office 2013 | Microsoft | лицензионное |
| 12 | Windows 10 Pro | Microsoft | лицензионное |
| 13 | Kaspersky Endpoint Security | Kaspersky | лицензионное |
| 14 | Microsoft Office 2016 | Microsoft | лицензионное |
| 15 | Visual Studio 2019 | Microsoft | лицензионное |
| 16 | Adobe Creative Cloud | Adobe | лицензионное |
| 17 | Zoom | Zoom | лицензионное |