

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт вычислительной математики и информационных технологий



УТВЕРЖДАЮ
Проректор по образовательной деятельности КФУ
Проф. Д.А. Таюрский

_____» _____ 20__ г.

подписано электронно-цифровой подписью

Программа дисциплины

Математическая лингвистика

Направление подготовки: 09.03.03 - Прикладная информатика

Профиль подготовки: не предусмотрено

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
 - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
 - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и) доцент, к.н. Гусенков А.М. (кафедра Интеллектуальные технологии поиска, Высшая школа информационных технологий и интеллектуальных систем), Alexandr.Gusenkov@kpfu.ru

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-2	способностью анализировать социально-экономические задачи и процессы с применением методов системного анализа и математического моделирования
ПК-2	способностью разрабатывать, внедрять и адаптировать прикладное программное обеспечение

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

основные понятия, структурное строение языка, задачи МЛ и пути их решения с использованием компьютерных систем;

Должен уметь:

-ориентироваться в области смежных наук - лингвистики, компьютерных систем и технологий и, отчасти, математики

Должен владеть:

теоретическими знаниями о прагматически-ориентированных лингвистических формальных и концептуальных моделях, этапах и уровнях формализации в лингвистике, о современном состоянии и о наиболее перспективных разработках в области математическая лингвистики;

Должен демонстрировать способность и готовность:

Студенты, завершившие изучение данной дисциплины, должны: знать: содержание программы курса, принципы функционирования автоматизированных лингвистических систем, формулировки задач, условия применимости и характеристики рассмотренных в курсе методов; уметь: применять методы математической лингвистики для анализа текстовой информации на естественном языке; владеть: методами математической лингвистики для анализа текстовой информации на естественном языке.

- Обладать знаниями об основных методах обработки естественного языка.

- Иметь представление о технологии лингвистического моделирования

- Получить навыки разработки инструментальных средств обработки текстов на естественном языке

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.5 Дисциплины (модули)" основной профессиональной образовательной программы 09.03.03 "Прикладная информатика (не предусмотрено)" и относится к дисциплинам по выбору.

Осваивается на 4 курсе в 7, 8 семестрах.

3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 8 зачетных(ые) единиц(ы) на 288 часа(ов).

Контактная работа - 102 часа(ов), в том числе лекции - 42 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 60 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 150 часа(ов).

Контроль (зачёт / экзамен) - 36 часа(ов).

Форма промежуточного контроля дисциплины: зачет в 7 семестре; экзамен в 8 семестре.

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

4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. ОПРЕДЕЛЕНИЕ БАЗОВЫХ ПОНЯТИЙ ФИЛОЛОГИЯ, ЛИНГВИСТИКА. Квантитативная лингвистика. Когнитивная лингвистика. Психолингвистика. Математическая лингвистика. Компьютерная лингвистика (МЛ).	7	2	0	4	14

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
2.	Тема 2. ОСНОВНЫЕ ЗАДАЧИ МЛ. Исследование ЕЯ и разработка лингвистических моделей исходя из принципа минимальной достаточности. Разработка технологий и средств обработки ЕЯ. Создание машиночитаемых лингвистических ресурсов. МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА. Математические модели и инструментарий из математической логики, теории множеств, теории алгоритмов.	7	2	0	4	14
3.	Тема 3. ПОНЯТИЕ NLP. ОПРЕДЕЛЕНИЕ И ЗАДАЧИ NLP. Описание функционирования языка. Разные определения языка (включая из теории формальных языков). NLP как научно-прикладная дисциплина по исследованию языков и построению систем обработки ЕЯ (СМП, СПЕЯ)	7	2	0	4	14
4.	Тема 4. ЛИНГВИСТИЧЕСКИЕ МОДЕЛИ. ПРАГМАТИЧЕСКИ-ОРИЕНТИРОВАННЫЙ ПОДХОД К ЯЗЫКОВОМУ МОДЕЛИРОВАНИЮ. Когнитивные модели. Интерактивные модели. Формально-концептуальные модели. Структурно-функциональные модели.	7	2	0	4	14
5.	Тема 5. ФОРМАЛЬНЫЕ ГРАММАТИКИ. Порождающие и распознающие грамматики. Главные задачи в рамках ФГ.	7	2	0	4	16
6.	Тема 6. БАЗОВАЯ КЛАССИФИКАЦИЯ ЯЗЫКОВ И ГРАММАТИК ПО ХОМСКОМУ. Четыре класса грамматик и языков. Грамматики типа "0". Неукорачивающие грамматики (НГ) и свойство легкораспознаваемости языков, порожденных НГ. КЗ-грамматики и НС-грамматики. КС-грамматики, А-грамматки.	7	2	0	4	8
7.	Тема 7. ОЦЕНКА СЛОЖНОСТИ ВЫВОДА. Временная и емкостная сложности и соответствующие сигнализирующие функции для всех классов грамматик.	7	2	0	4	8
8.	Тема 8. МОДЕЛЬ "СМЫСЛ-ТЕКСТ". Раскрывается на примере программы ЭТАП.	7	2	0	4	8
9.	Тема 9. ЛИНГВИСТИЧЕСКИЙ ПОДХОД. Генеративные грамматики, функциональные грамматики, грамматики на ограничениях.	7	2	0	4	12

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
10.	Тема 10. ЭКСПЕРИМЕНТАЛЬНЫЕ МОДЕЛИ. Семантические роли Филлмора, семантические сценарии (КЗ-грамматики) Шенка.	8	2	0	8	4
11.	Тема 11. ПРАГМАТИЧЕСКИЙ ПОДХОД. Жесткие ограничения на контекст и язык.	8	2	0	8	4
12.	Тема 12. РС-КИММО КАК СРЕДСТВО РАЗРАБОТКИ МОРФОЛОГИЧЕСКИХ АНАЛИЗАТОРОВ.	8	2	0	8	4
13.	Тема 13. ФАЙЛ ПРАВИЛ. ФАЙЛ МОРФОТАКТИЧЕСКИХ ПРАВИЛ. Пример из татарской морфологии.	8	2	0	0	4
14.	Тема 14. ПРОБЛЕМА ПРЕДСТАВЛЕНИЯ И ИЗВЛЕЧЕНИЯ ЗНАНИЙ И ОБРАБОТКИ ЕЯ-ТЕКСТОВ. Онтологии.	8	2	0	0	6
15.	Тема 15. СЕМИОТИЧЕСКИЕ СИСТЕМЫ. ТРЕУГОЛЬНИК ФРЕГГЕ. ЧЕТЫРЕХУГОЛЬНИК ПОСПЕЛОВА.	8	2	0	0	8
16.	Тема 16. ПРИОБРЕТЕНИЕ ЗНАНИЙ ПРИ ОБРАБОТКЕ ТЕКСТА. Использование знаний. Общая схема интеллектуальной системы, понимающей текст. Крупноблочный алгоритм общего механизма вывода.	8	2	0	0	4
17.	Тема 17. ЛИНГВИСТИЧЕСКИЙ ПРОЦЕССОР. Лингвистический процессор вопросно-ответных текстов. Особенности. Нерешенные проблемы. Описание ЛП1: разработки СНИЛ ПИИ АНТ и КГУ.	8	2	0	0	2
18.	Тема 18. ШЕСТЬ ПРИНЦИПОВ ПОСТРОЕНИЯ ЛП В ВОПРОСНО-ОТВЕТНОМ КОНТЕКСТЕ.	8	4	0	0	3
19.	Тема 19. СТАТИСТИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ В ЛИНГВИСТИКЕ. ЯЗЫК И РЕЧЬ. Динамика и статика языка. Аспекты речевой деятельности. Языковая схема. Языковая компетенция. Речевой процесс. Речевой продукт.	8	4	0	0	3
	Итого		42	0	60	150

4.2 Содержание дисциплины (модуля)

Тема 1. ОПРЕДЕЛЕНИЕ БАЗОВЫХ ПОНЯТИЙ ФИЛОЛОГИЯ, ЛИНГВИСТИКА. Квантитативная лингвистика. Когнитивная лингвистика. Психолингвистика. Математическая лингвистика. Компьютерная лингвистика (МЛ).

В этом разделе рассматриваются базовые понятия филологии и лингвистики. Даются определения количественной и когнитивной лингвистики, психолингвистики, математической лингвистики, компьютерной лингвистики. Определяется значение качественных и количественных данных для изучения языка. Указывается отличие количественной лингвистики от математической лингвистики. **ОСНОВНЫЕ ЗАДАЧИ МЛ.** Исследование ЕЯ и разработка лингвистических моделей исходя из принципа минимальной достаточности. Разработка технологий и средств обработки ЕЯ. Создание машиночитаемых лингвистических ресурсов. **МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА.** Математические модели и инструментарий из математической логики, теории множеств, теории алгоритмов. NLP как научно-прикладной дисциплины по исследованию языков и построению систем обработки ЕЯ (СМП, СПЕЯ). Когнитивные модели. Интерактивные модели. Формально-концептуальные модели. Структурно-функциональные модели.

Тема 2. ОСНОВНЫЕ ЗАДАЧИ МЛ. Исследование ЕЯ и разработка лингвистических моделей исходя из принципа минимальной достаточности. Разработка технологий и средств обработки ЕЯ. Создание машиночитаемых лингвистических ресурсов. МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА. Математические модели и инструментарий из математической логики, теории множеств, теории алгоритмов.

ФОРМАЛЬНЫЕ ГРАММАТИКИ. Набор правил синтаксиса образует грамматику языка. Правила синтаксиса могут описывать либо процедуру получения правильных предложений, либо процедуру распознавания "правильности" предложений (т. е. их принадлежности данному языку). В первом случае грамматику называют порождающей, во втором - распознающей. **БАЗОВАЯ КЛАССИФИКАЦИЯ ЯЗЫКОВ И ГРАММАТИК ПО ХОМСКОМУ.** Четыре класса грамматик и языков. Грамматики типа "0". Неукорачивающие грамматики (НГ) и свойство легкораспознаваемости языков, порожденных НГ. КЗ-грамматики и НС-грамматики. КС-грамматики, А-грамматики. Временная и емкостная сложности и соответствующие сигнализирующие функции для всех классов грамматик. **МОДЕЛЬ "СМЫСЛ-ТЕКСТ".** Генеративные грамматики, функциональные грамматики, грамматики на ограничениях.

Тема 3. ПОНЯТИЕ NLP. ОПРЕДЕЛЕНИЕ И ЗАДАЧИ NLP. Описание функционирования языка. Разные определения языка (включая из теории формальных языков). NLP как научно-прикладная дисциплина по исследованию языков и построению систем обработки ЕЯ (СМП, СПЕЯ)

Семантические роли Филлмора, семантические сценарии (КЗ-грамматики) Шенка. **ПРАГМАТИЧЕСКИЙ ПОДХОД.** Жесткие ограничения на контекст и язык. РС-KIMMO КАК СРЕДСТВО РАЗРАБОТКИ МОРФОЛОГИЧЕСКИХ АНАЛИЗАТОРОВ. **ФАЙЛ ПРАВИЛ. ФАЙЛ МОРФОТАКТИЧЕСКИХ ПРАВИЛ.** Пример из татарской морфологии. **ПРОБЛЕМА ПРЕДСТАВЛЕНИЯ И ИЗВЛЕЧЕНИЯ ЗНАНИЙ И ОБРАБОТКИ ЕЯ-ТЕКСТОВ.** Онтологии.

Тема 4. ЛИНГВИСТИЧЕСКИЕ МОДЕЛИ. ПРАГМАТИЧЕСКИ-ОРИЕНТИРОВАННЫЙ ПОДХОД К ЯЗЫКОВОМУ МОДЕЛИРОВАНИЮ. Когнитивные модели. Интерактивные модели. Формально-концептуальные модели. Структурно-функциональные модели.

СЕМИОТИЧЕСКИЕ СИСТЕМЫ. ТРЕУГОЛЬНИК ФРЕГГЕ. ЧЕТЫРЕХУГОЛЬНИК ПОСПЕЛОВА. ПРИОБРЕТЕНИЕ ЗНАНИЙ ПРИ ОБРАБОТКЕ ТЕКСТА. Использование знаний. Общая схема интеллектуальной системы, понимающей текст. Крупноблочный алгоритм общего механизма вывода. **ЛИНГВИСТИЧЕСКИЙ ПРОЦЕССОР.** Лингвистический процессор вопросно-ответных текстов. Особенности. Нерешенные проблемы. **ШЕСТЬ ПРИНЦИПОВ ПОСТРОЕНИЯ ЛП В ВОПРОСНО-ОТВЕТНОМ КОНТЕКСТЕ. СТАТИСТИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ В ЛИНГВИСТИКЕ. ЯЗЫК И РЕЧЬ.** Динамика и статика языка. Аспекты речевой деятельности. Языковая схема. Языковая компетенция. Речевой процесс. Речевой продукт.

Тема 5. ФОРМАЛЬНЫЕ ГРАММАТИКИ. Порождающие и распознающие грамматики. Главные задачи в рамках ФГ.

ФОРМАЛЬНЫЕ ГРАММАТИКИ. Порождающие и распознающие грамматики. Главные задачи в рамках формальных.

Теоретическая часть. Порождение языков формальными грамматиками. Построение цепочек вывода языков из формальных грамматик. Построение цепочек распознавания принадлежности языку из формальных грамматик.

Тема 6. БАЗОВАЯ КЛАССИФИКАЦИЯ ЯЗЫКОВ И ГРАММАТИК ПО ХОМСКОМУ. Четыре класса грамматик и языков. Грамматики типа "0". Неукорачивающие грамматики (НГ) и свойство легкораспознаваемости языков, порожденных НГ. КЗ-грамматики и НС-грамматики. КС-грамматики, А-грамматики.

БАЗОВАЯ КЛАССИФИКАЦИЯ ЯЗЫКОВ И ГРАММАТИК ПО ХОМСКОМУ. Четыре класса грамматик и языков. Грамматики типа "0". Неукорачивающие грамматики (НГ) и свойство легкораспознаваемости языков, порожденных неукорачивающимися грамматиками. Построение цепочек вывода языков из формальных грамматик. НГ. КЗ-грамматики и НС-грамматики. КС-грамматики, А-грамматики.

Теоретическая часть.

Тема 7. ОЦЕНКА СЛОЖНОСТИ ВЫВОДА. Временная и емкостная сложности и соответствующие сигнализирующие функции для всех классов грамматик.

ОЦЕНКА СЛОЖНОСТИ ВЫВОДА. Временная и емкостная сложности и соответствующие сигнализирующие функции для всех классов грамматик. Теоретическая часть. Особенности вывода временной сложности. Методы, приемы, техника получения временной сложности. Особенности вывода емкостная сложности. Методы, приемы, техника получения емкостная сложности.

Тема 8. МОДЕЛЬ "СМЫСЛ-ТЕКСТ". Раскрывается на примере программы ЭТАП.

МОДЕЛЬ "СМЫСЛ-ТЕКСТ". Раскрывается на примере программы ЭТАП. Теоретическая часть. Теория "Смысл - текст" (название более точно записывается как "теория "Смысл \leftrightarrow Текст") - лингвистическая концепция, созданная И. А. Мельчуком и представляющая язык как многоуровневую модель преобразований смысла в текст и обратно (модель "Смысл \leftrightarrow Текст"); отличительной особенностью этой теории является также использование синтаксиса зависимостей. Значительна роль, отводимая лексическому компоненту модели

Тема 9. ЛИНГВИСТИЧЕСКИЙ ПОДХОД. Генеративные грамматики, функциональные грамматики, грамматики на ограничениях.

ЛИНГВИСТИЧЕСКИЙ ПОДХОД. Генеративные грамматики, функциональные грамматики, грамматики на ограничениях. Теоретическая часть. Порождающая грамматика (генеративная грамматика, англ. generative grammar) - формализм генеративной лингвистики, связанный с изучением синтаксиса. В рамках подхода порождающей грамматики формулируется система правил, при помощи которых можно определить, какая комбинация слов оформляет грамматически правильное предложение.

Тема 10. ЭКСПЕРИМЕНТАЛЬНЫЕ МОДЕЛИ. Семантические роли Филлмора, семантические сценарии (КЗ-грамматики) Шенка.

ЭКСПЕРИМЕНТАЛЬНЫЕ МОДЕЛИ. Семантические роли Филлмора, семантические сценарии (КЗ-грамматики) Шенка. Теоретическая часть. Семантическая роль имени при предикате - в языкознании: часть семантики предиката, отражающая общие свойства аргумента предиката - участника называемой предикатом ситуации [1]. Описание в терминах семантических ролей отражает сходства моделей управления различных предикатных слов].

Понятие семантической роли предложено Ч. Филлмором, вначале использовавшим термин "глубинный падеж". В порождающей грамматике ему соответствует термин "тета-роль".

Тема 11. ПРАГМАТИЧЕСКИЙ ПОДХОД. Жесткие ограничения на контекст и язык.

ПРАГМАТИЧЕСКИЙ ПОДХОД. Жесткие ограничения на контекст и язык. Теоретическая часть.

ПРАГМАТИЧЕСКИЙ ПОДХОД. Жесткие ограничения на контекст и язык. Теоретическая часть. Основными понятиями прагматики являются субъект (отправитель текста), адресат (получатель текста), предмет коммуникации и коммуникативная установка. При этом прагматика охватывает следующие вопросы:

в обыденной речи, с одной стороны, отношение говорящего к тому, что и как он говорит: истинность, объективность, предположительность речи, её искренность или неискренность, её приспособленность к социальной среде и к социальному положению слушающего и т. д., с другой стороны, интерпретация речи слушателем: как истинной, объективной, искренней или, напротив, ложной, сомнительной, вводящей в заблуждение.

Тема 12. РС-KIMMO КАК СРЕДСТВО РАЗРАБОТКИ МОРФОЛОГИЧЕСКИХ АНАЛИЗАТОРОВ.

РС-KIMMO КАК СРЕДСТВО РАЗРАБОТКИ МОРФОЛОГИЧЕСКИХ АНАЛИЗАТОРОВ. Теоретическая часть. Морфологический анализатор РС-KIMMO <http://www-01.sil.org/pckimmo/>

Для чего нужен автоматический морфологический разбор слов? Парсинг (синтаксический разбор) - это стандартный метод, используемый в АОЕЯ. Но сначала каждому входящему в предложение слову должна быть приписана определенная информация. Чтобы разобрать предложение The cat chased the rat, анализатор должен знать, что слово cat - это существительное в ед.ч., chased - форма прошедшего времени глагола и т.д.

Тема 13. ФАЙЛ ПРАВИЛ. ФАЙЛ МОРФОТАКТИЧЕСКИХ ПРАВИЛ. Пример из татарской морфологии.

ФАЙЛ ПРАВИЛ. ФАЙЛ МОРФОТАКТИЧЕСКИХ ПРАВИЛ. Пример из морфологии.

Прагматически-ориентированный подход к построению лингвистических моделей это, прежде всего, концептуально-инструментальная технология, которая позволяет, с одной стороны, осуществлять адекватный подбор средств эффективной обработки ЕЯ-текста, с другой стороны, детерминировать контекст и управлять формированием образа генерируемого или распознаваемого ЕЯ-текста.

Тема 14. ПРОБЛЕМА ПРЕДСТАВЛЕНИЯ И ИЗВЛЕЧЕНИЯ ЗНАНИЙ И ОБРАБОТКИ ЕЯ-ТЕКСТОВ. Онтологии.

обсуждается возможность использования знаний и концепций, накопленных к настоящему моменту лингвистической наукой, при решении задач прикладного характера, возникающих в контексте автоматической обработки естественного языка. Речь идет о необходимости восприятия обрабатываемого текста как целостного дискурса. Рассматривается ряд технологических приемов, использование которых в процессе разработки ЕЯ-систем позволяет говорить о дискурсивности автоматического анализа. Углубленное изучение на конкретных примерах.

Тема 15. СЕМИОТИЧЕСКИЕ СИСТЕМЫ. ТРЕУГОЛЬНИК ФРЕГГЕ. ЧЕТЫРЕХУГОЛЬНИК ПОСПЕЛОВА.

СЕМИОТИЧЕСКИЕ СИСТЕМЫ. ТРЕУГОЛЬНИК ФРЕГГЕ. ЧЕТЫРЕХУГОЛЬНИК ПОСПЕЛОВА. Углубленное изучение на конкретных примерах. Развивая идеи Пирса, немецкий математик и логик Готлоб Фреге (1848-1925) предложил собственную треугольную логическую модель функционирования знака. Фреге писал: "Употребляя знак, мы хотим сказать что-то не о знаке, но главным, как правило, является его значение".

В этой модели, также имеющей треугольную форму, как и модель Пирса, представлена зависимость между тремя другими феноменами: знаком, объективно существующим элементом реальности, который он отражает, и представлением об этом элементе (всем классе элементов), функционирующем в сознании субъекта знаковой деятельности

Тема 16. ПРИОБРЕТЕНИЕ ЗНАНИЙ ПРИ ОБРАБОТКЕ ТЕКСТА. Использование знаний. Общая схема интеллектуальной системы, понимающей текст. Крупноблочный алгоритм общего механизма вывода.

ПРИОБРЕТЕНИЕ ЗНАНИЙ ПРИ ОБРАБОТКЕ ТЕКСТА. Использование знаний. Общая схема интеллектуальной системы, понимающей текст. Крупноблочный алгоритм общего механизма вывода. Углубленное изучение на конкретных примерах. ри использовании продукционной модели представления знаний БЗ состоит из набора правил, а программа, управляющая перебором правил, называется интерпретатором правил, или машиной вывода. Интерпретатор правил - это программа, имитирующая логический вывод эксперта, пользующегося продукционной БЗ для интерпретации поступивших в систему данных

Тема 17. ЛИНГВИСТИЧЕСКИЙ ПРОЦЕССОР. Лингвистический процессор вопросно-ответных текстов. Особенности. Нерешенные проблемы. Описание ЛП1: разработки СНИЛ ПИИ АНТ и КГУ.

ЛИНГВИСТИЧЕСКИЙ ПРОЦЕССОР. Лингвистический процессор вопросно-ответных текстов. Особенности. Нерешенные проблемы. Описание ЛП1: разработки СНИЛ ПИИ АНТ и КГУ. Углубленное изучение на конкретных примерах. ри использовании продукционной модели представления знаний БЗ состоит из набора правил, а программа, управляющая перебором правил, называется интерпретатором правил, или машиной вывода. Интерпретатор правил - это программа, имитирующая логический вывод эксперта, пользующегося продукционной БЗ для интерпретации поступивших в систему данных

Тема 18. ШЕСТЬ ПРИНЦИПОВ ПОСТРОЕНИЯ ЛП В ВОПРОСНО-ОТВЕТНОМ КОНТЕКСТЕ.

ШЕСТЬ ПРИНЦИПОВ ПОСТРОЕНИЯ ЛП В ВОПРОСНО-ОТВЕТНОМ КОНТЕКСТЕ. Углубленное изучение на конкретных примерах. Вопросно-ответная система (QA-система; от англ. QA - англ. Question-answering system) - информационная система, способная принимать вопросы и отвечать на них на естественном языке, другими словами, это система с естественно-языковым интерфейсом.

Тема 19. СТАТИСТИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ В ЛИНГВИСТИКЕ. ЯЗЫК И РЕЧЬ. Динамика и статика языка. Аспекты речевой деятельности. Языковая схема. Языковая компетенция. Речевой процесс. Речевой продукт.

СТАТИСТИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ В ЛИНГВИСТИКЕ. ЯЗЫК И РЕЧЬ. Динамика и статика языка. Аспекты речевой деятельности. Языковая схема. Языковая компетенция. Речевой процесс. Речевой продукт. Углубленное изучение на конкретных примерах. Под речью современное языкознание понимает не только устную речь, но также речь письменную. В понятие речи в широком смысле включается и "внутренняя" речь, то есть мышление с помощью языковых средств (слов и др.), осуществляемое про себя. Проявлением речевой деятельности служит речевой акт. Отдельный акт речи, речевой акт представляет собой двусторонний процесс, охватывающий при устном говорении не только говорение, но и слуховое восприятие, а также понимание услышанного. При письменном общении речевой акт охватывает соответственно писание и чтение (зрительное восприятие и понимание написанного), причем участники могут быть отделены друг от друга во времени.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства образования и науки Российской Федерации от 5 апреля 2017 года №301)

Письмо Министерства образования Российской Федерации №14-55-996ин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемым результатам обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;
- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы. Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;
- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

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

Википедия - <http://ru.wikipedia.org>

Интернет-журнал по ИТ - <http://www.rsdn.ru>

Интернет-портал образовательных ресурсов по ИТ - <http://www.intuit.ru>

Интернет-портал образовательных ресурсов по ИТ - <http://algolist.manual.ru/>

Портал математических интернет-ресурсов - <http://www.math.ru/>

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

Вид работ	Методические рекомендации
лекции	Посещение и активная работа студента на лекции позволяет сформировать базовые теоретические понятия по дисциплине, овладеть общей логикой построения дисциплины, усвоить закономерности и тенденции, которые раскрываются в данной дисциплине. При этом студенту рекомендуется быть достаточно внимательным на лекции, стремиться к пониманию основных положений лекции, а при определенных трудностях и вопросах, своевременно обращаться к лектору за пояснениями, уточнениями или при дискуссионности рассматриваемых вопросов, получения от лектора собственной научной точки зрения как ученого.

Вид работ	Методические рекомендации
лабораторные работы	Посещение и работа студента на лабораторном занятии позволяет в процессе решения практических задач и коллективного обсуждения результатов их решения глубже усвоить теоретические положения, сформировать отдельные практические умения и навыки, научиться правильно обосновывать методику выполнения расчетов, четко и последовательно проводить расчеты, формулировать выводы и предложения. Работа на практическом занятии дает возможность студенту всесторонне изучить дисциплину и подготовиться для самостоятельной работы. В процессе выполнения аудиторных практических работ студент подтверждает полученные знания, умения и навыки, которые формируют соответствующие компетенции.
самостоятельная работа	Работа над материалами лекции во вне аудиторное время предполагает более глубокое рассмотрение вопросов темы с учетом того, что на лекции не возможно полно осветить все вопросы темы. Для глубокой проработки темы студент должен: а) внимательно прочитать лекцию (возможно несколько раз); б) рассмотреть вопросы темы или проблемы по имеющейся учебной, учебно-методической литературе, ознакомиться с подходами по данной теме, которые существуют в современной научной литературе (посмотреть монографии, статьи в журналах, тезисы научных докладов и выступлений). Кроме того, студент может при глубокой проработке темы пользоваться материалами, которые представляют эксперты, различные научные дискуссии и т.п.
зачет	В ходе подготовки к зачету обучающимся доводятся заранее подготовленные вопросы по дисциплине. В преддверии зачета преподаватель заблаговременно проводит групповую консультацию и, в случае необходимости, индивидуальные консультации с обучающимися. При проведении консультации обобщается пройденный материал, раскрывается логика его изучения, привлекается внимание к вопросам, представляющим наибольшие трудности для всех или большинства обучающихся, рекомендуется литература, необходимая для подготовки к зачету. При подготовке к зачету обучающиеся внимательно изучают конспект, рекомендованную литературу и делают краткие записи по каждому вопросу. Такая методика позволяет получить прочные и систематизированные знания, необходимые на зачете.
экзамен	В ходе подготовки к экзамену обучающимся доводятся заранее подготовленные вопросы по дисциплине. В преддверии экзамена преподаватель заблаговременно проводит групповую консультацию и, в случае необходимости, индивидуальные консультации с обучающимися. При проведении консультации обобщается пройденный материал, раскрывается логика его изучения, привлекается внимание к вопросам, представляющим наибольшие трудности для всех или большинства обучающихся, рекомендуется литература, необходимая для подготовки к экзамену. При подготовке к экзамену обучающиеся внимательно изучают конспект, рекомендованную литературу и делают краткие записи по каждому вопросу. Такая методика позволяет получить прочные и систематизированные знания, необходимые на экзамене.

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

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

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

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Мультимедийная аудитория.

12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 09.03.03 "Прикладная информатика" и профилю подготовки "не предусмотрено".

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 09.03.03 - Прикладная информатика

Профиль подготовки: не предусмотрено

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Основная литература:

1. Микони, С.В. Дискретная математика для бакалавра: множества, отношения, функции, графы [Электронный ресурс] : учебное пособие / С.В. Микони. - Электрон. дан. - Санкт-Петербург : Лань, 2012. - 192 с. - Режим доступа: <https://e.lanbook.com/book/4316>
2. Попов, С.В. Прикладная логика [Электронный ресурс] : учебное пособие / С.В. Попов, Н.Л. Брошкова. - Электрон. дан. - Москва : Физматлит, 2011. - 216 с. - Режим доступа: <https://e.lanbook.com/book/5286>
3. Математика и информатика: Учебник / В.Я. Турецкий; Уральский государственный университет им. А.М. Горького. - 3-е изд., испр. и доп. - М.: ИНФРА-М, 2010. - 558 с. - (Высшее образование). ISBN 978-5-16-000171-5 - Режим доступа: <http://znanium.com/catalog/product/206346>

Дополнительная литература:

1. Информационные системы: Учебное пособие / О.Л. Голицына, Н.В. Максимов, И.И. Попов. - 2-е изд. - М.: Форум: НИЦ ИНФРА-М, 2014. - 448 с. - (Высшее образование). ISBN 978-5-91134-833-5 - Режим доступа: <http://znanium.com/catalog/product/435900>
2. Теоретические основы информационных процессов и систем / Душин В.К., - 5-е изд. - М.: Дашков и К, 2018. - 348 с.: ISBN 978-5-394-01748-3 - Режим доступа: <http://znanium.com/catalog/product/450784>
3. Информационные технологии и системы: Учебное пособие / Е.Л. Федотова. - М.: ИД ФОРУМ: НИЦ ИНФРА-М, 2014. - 352 с. - (Высшее образование). ISBN 978-5-8199-0376-6 - Режим доступа: <http://znanium.com/catalog/product/429113>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.5 Математическая лингвистика

Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Направление подготовки: 09.03.03 - Прикладная информатика

Профиль подготовки: не предусмотрено

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен обучающимся. ЭБС Издательства "Лань" включает в себя электронные версии книг издательства "Лань" и других ведущих издательств учебной литературы, а также электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. ЭБС Издательства "Лань" обеспечивает доступ к научной, учебной литературе и научным периодическим изданиям по максимальному количеству профильных направлений с соблюдением всех авторских и смежных прав.