

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Факультет исторических и политических наук

УТВЕРЖДЕНО:

Декан

Ж. А. Рожнева

Оценочные материалы по дисциплине

Основы математической лингвистики

по направлению подготовки

43.03.02 Туризм

Направленность (профиль) подготовки:

«Организация и управление туристским и гостиничным бизнесом»

Форма обучения

Очная

Квалификация

Бакалавр

Год приема

2025

СОГЛАСОВАНО:

Руководитель ОП

И.В. Муравьев

1. Компетенции и индикаторы их достижения, проверяемые данными оценочными материалами

Целью освоения дисциплины является формирование следующих компетенций:

УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИУК-1.1 Анализирует задачу, выделяя ее базовые составляющие, осуществляет декомпозицию задачи

ИУК-1.2 Находит и критически анализирует информацию, необходимую для решения поставленной задачи

ИУК-1.3 Грамотно, логично, аргументированно формирует собственные суждения и оценки. Отличает факты от мнений, интерпретаций, оценок и т.д. в рассуждениях других участников деятельности

2. Оценочные материалы текущего контроля и критерии оценивания

Элементы текущего контроля:

Оценочные средства для текущего контроля

Примерные тесты для входного контроля

1. Сколько различных "слов" можно составить из четырех разных букв русского алфавита?

а) 4; б) 8; в) 16; г) 24.

2. Как называется совпадение в звучании и написании языковых единиц, значения которых не связаны друг с другом?

а) паронимия; б) омонимия; в) синонимия; г) антонимия.

3. Из лжи может следовать:

а) все, что угодно; б) только истина; в) только ложь; г) ничего.

4. Установите истинность высказывания: "В любом треугольнике все углы острые":

а) ложь; б) истина; в) это предложение не является высказыванием; г) иногда - ложь, иногда - истина;

5. Какова местность предиката "Река x впадает в море y"?

а) 1; б) 2; в) 3; г) 4.

6. Как можно получить высказывание из предиката?

а) только навешиванием кванторов; б) только подставив вместо логических переменных какие-то конкретные значения; в) и так, и так; г) никак.

7) Омонимия в математике:

а) вредна; б) полезна; в) встречается очень часто; г) приводит к новым математическим открытиям.

8) Основной метод добывания истины в математике - это:

а) доказательство; б) эксперимент; в) наблюдение; г) аналогия.

9) Какой союз в русском языке отвечает логической операции "конъюнкция"?

а) но; б) а; в) или; г) и.

10) Какой союз в русском языке отвечает логической операции "дизъюнкция"?

а) но; б) а; в) или; г) и.

Примерные тесты для текущего контроля №1

1. Информационные технологии в лингвистике - это
 - 1) Последовательность действий при решении лингвистической задачи
 - 2) Совокупность законов, методов и средств получения, хранения, передачи и преобразования информации о языке с помощью компьютера
 - 3) Программа, расширяющая возможности операционной системы (например, антивирусные программы, архиваторы и т.д.)
 - 4) Совокупность систематизированных сведений, фактов, событий в электронном варианте
2. Компьютерная лингвистика - направление в прикладной лингвистике, ориентированное на:
 - 1) использование программ, компьютерных технологий организации и обработки данных - для моделирования функционирования языка в тех или иных условиях, ситуациях, проблемных сферах
 - 2) применение компьютерных моделей языка в лингвистике и смежных дисциплинах при решении практических задач, связанных с изучением языка
3. Объяснение фактов языка является задачей
 - 1) теоретической лингвистики
 - 2) прикладной лингвистики
 - 3) математической лингвистики
 - 4) компьютерной лингвистики
4. Данные - это ...
 - 1) факты, отражающие объекты, процессы и явления предметной области
 - 2) закономерности, установленные в результате практической деятельности и накопления профессионального опыта в некоторой проблемной области и позволяющие специалистам ставить и решать задачи в этой области
 - 3) сведения, рассматриваемые в каком-либо контексте и позволяющие которого пользователю составить собственное мнение
5. Данные соответствуют ... аспекту отражения действительности
 - 1) прагматическому
 - 2) синтаксическому
 - 3) семантическому
6. Информация - это ...
 - 1) факты, характеризующие объекты, процессы и явления предметной области, а также их свойства
 - 2) результаты размышлений
 - 3) закономерности, установленные в результате практической деятельности и накопления профессионального опыта в некоторой проблемной области и позволяющие специалистам ставить и решать задачи в этой области
 - 4) сведения, рассматриваемые в каком-либо контексте, который имеет значение для пользователя
7. Информация соответствует ... аспекту отражения действительности
 - 1) синтаксическому
 - 2) семантическому
 - 3) прагматическому
8. При оценке информации различают такие её аспекты, как.
 - 1) логический

- 2) синтаксический
 - 3) математический
 - 4) физический
 - 5) прагматический
 - б) семантический
9. Объем данных - это ... мера оценки информации
- 1) логическая
 - 2) семантическая
 - 3) алгоритмическая
 - 4) физическая
 - 5) прагматическая
 - 6) синтаксическая
 - 7) абсолютная
10. Информация, рассматриваемая только относительно синтаксического аспекта, называется:
- 1) данными
 - 2) сигналами
 - 3) сообщением
 - 4) символами
 - 5) знаками
11. Семантический аспект передает ... содержание информации:
- 1) структурное
 - 2) полезное
 - 3) смысловое
 - 4) качественное
 - 5) формальное
12. Прагматический аспект отражает ... свойства информации:
- 1) потребительские
 - 2) структурные
 - 3) количественные
 - 4) качественные
 - 5) смысловые
 - 6) динамические
13. Коэффициент содержательности используется для ... оценки информации
- 1) логической
 - 2) семантической
 - 3) алгоритмической
 - 4) физической
 - 5) прагматической
 - 6) синтаксической
 - 7) абсолютной

Примерные тесты для текущего контроля №2

1. Знания - это ...
 - 1) факты, характеризующие объекты, процессы и явления предметной области, а также их свойства
 - 2) закономерности, установленные в результате практической деятельности и

накапливания профессионального опыта в некоторой проблемной области и позволяющие специалистам ставить и решать задачи в этой области

3) сведения, рассматриваемые в каком-либо контексте и позволяющие которого пользователю составить собственное мнение

2. Знания соответствуют ... аспекту отражения действительности

- 1) прагматическому
- 2) синтаксическому
- 3) семантическому

3. Знаниями являются ... факты

- 1) осмысленные
- 2) новые
- 3) зафиксированные

3. Факты, отображающие объекты, процессы и явления предметной области, а также их свойства, — это ...

- 1) данные
- 2) знания
- 3) информация
- 4) описания

4. Сведения, рассматриваемые в каком-либо контексте, которое имеют значение для пользователя, — это

- 1) данные
- 2) знания
- 3) описания
- 4) результаты
- 5) информация

5. Закономерности, установленные в результате практической деятельности и накапливания профессионального опыта в некоторой проблемной области и позволяющие специалистам ставить и решать задачи в этой области, - это ...

- 1) данные
- 2) знания
- 3) информация

6. Установите соответствие ...

Знания - это ...

Данные - это ...

Информация - это ...

Варианты:

- a. записанные на каком-либо носителе факты
- b. понятые субъектом факты и их зависимости, запоминаемые для последующего применения
- c. новые и полезные для решения задач факты

7. ... может быть элементарной единицей структурного знания

- 1) Объект
- 2) Значение
- 3) Факт
- 4) Коэффициент уверенности
- 5) Правило

8. Понятие это:

- 1) мысль, отражающая общие и существенные признаки определенной совокупности предметов, процессов, явлений действительности
- 2) соглашение (явное или неявное) о приписывании чему-либо (означающему) какого-либо определённого смысла (означаемого)
- 3) форма мышления, отражающая существенные свойства, связи и отношения предметов и явлений в их противоречии и развитии; мысль или система мыслей, обобщающая, выделяющая предметы некоторого класса по определённым общим и в совокупности специфическим для них признакам
- 4) термин предметной области
- 5) отображение сущности предмета в человеческом мышлении; предмет при этом может трактоваться расширительно: как объект, группа объектов, отношения между ними, как связи свойств, абстрагированных от объектов.

9. Определение понятия (дефиниция) - это:

- 1) операция, состоящая в раскрытии содержания понятия, т. е. нахождении существенно-общих признаков определенного класса предметов, мыслимого в данном понятии
- 2) логическая операция, направленная на выявление правильного значения термина или содержания понятия
- 3) операция, состоящая в раскрытии объема понятия

10. Обобщение понятия - это:

- 1) совершение перехода от понятия с меньшим объемом, но большим содержанием к понятию с большим объемом и меньшим содержанием
- 2) совершение перехода от понятия с большим объемом и меньшим содержанием к понятию с меньшим объемом, но большим содержанием

11. При обобщении осуществляется переход:

- 1) от видового понятия к родовому
- 2) от родового понятия к видовому
- 3) от определяющего понятия к определяемому
- 4) от определяемого понятия к определяющему

12. Суждение - это:

- 1) умственный акт, выражающий отношение говорящего к содержанию высказываемой мысли посредством утверждения модальности сказанного и сопряжённый обычно с психологическим состоянием убеждённости или веры.
- 2) форма мышления, в которой что-либо утверждается или отрицается о предмете, его свойствах или отношениях между предметами
- 3) форма мышления, в которой, сочетая понятия, что-либо утверждается или отрицается о самых реальных вещах и явлениях.
- 4) соглашение (явное или неявное) о приписывании чему-либо (означающему) какого-либо определённого смысла (означаемого)

13. Умозаключение - это:

- 1) умственное действие, связывающее в ряд "посылок" и "следствий" мысли различного содержания
- 2) форма мышления, в которой что-либо утверждается или отрицается о предмете, его свойствах или отношениях между предметами
- 3) мыслительный процесс, в ходе которого из одного или нескольких суждений, называемых посылками, выводится новое суждение, называемое заключением, или следствием.
- 4) умственный акт, выражающий отношение говорящего к содержанию высказываемой

мысли посредством утверждения модальности сказанного и сопряжённый обычно с психологическим состоянием убеждённости или веры.

Критерии оценки:

- Оценка «отлично» выставляется студенту, если он дал правильные ответы в диапазоне 85-100%, тем самым показав знание теоретических основ математической лингвистики, умение применять эти знания.
- Оценка «хорошо» выставляется студенту, если он дал правильные ответы на 76-84% вопросов теста, тем самым показав неплохое знание основных положений математической лингвистики, умение применять эти знания.
- Оценка «удовлетворительно» выставляется студенту, если он дал правильные ответы на 61-75% вопросов, показав посредственное знание основных положений математической лингвистики, несистемное умение применять эти знания.
- Оценка «неудовлетворительно» выставляется студенту, если он дал правильные ответы менее чем на 61% вопросов, показав знание только отдельных положений математической лингвистики, слабое умение применять эти фрагментарные знания, отсутствие навыков применения математических методов к решению задач лингвистики.

Вопросы к коллоквиуму:

1. Современные стратегии доступа к филологической (лингвистической) информации.
2. Типы лингвистических Интернет-ресурсов.
3. Сетевой поиск лингвистической информации.
4. Компьютерный инструментарий, применяемый в лингвистических исследованиях.
5. Новейшие компьютерные технологии в переводоведении.
6. Современное состояние и перспективы развития машинного перевода.
7. Поиск и публикация информации в Интернет. Основные понятия информационного поиска.
8. Новейшие компьютерные технологии в лексикографии.
9. Виртуальные библиотеки (на примере одной библиотеки).
10. Системы дистанционного обучения и компьютерные обучающие системы.
11. Система TextAnalyst как инструмент анализа содержания текстов, смыслового поиска информации, формирования электронных архивов.
12. Психолингвистические программы.
13. Лингвистическая проблематика обучения с помощью ЭВМ.
14. Перевод как аналитико-синтетический процесс.
15. Характеристика основных возможностей систем машинного перевода PROMT, SYSTRAN и Google.

Критерии оценки:

«Отлично», повышенный уровень: изложение полученных знаний в устной, письменной или графической форме, полное, в системе, в соответствии с требованиями учебной программы; допускаются единичные несущественные ошибки, самостоятельно исправляемые студентами;

«Хорошо», пороговый уровень: изложение полученных знаний в устной, письменной и графической форме, полное, в системе, в соответствии с требованиями учебной программы; допускаются отдельные несущественные ошибки, исправляемые студентами после указания преподавателя на них;

«Удовлетворительно», пороговый уровень: изложение полученных знаний неполное, однако это не препятствует усвоению последующего программного материала; допускаются отдельные существенные ошибки, исправленные с помощью преподавателя;

«Неудовлетворительно», уровень не сформирован: изложение учебного материала неполное, бессистемное, что препятствует усвоению последующей учебной информации;

существенные ошибки, неисправляемые даже с помощью преподавателя.

Кейс-задания №1. Итогом работы трёх научных коллективов явилось издание трёх толковых словарей. Известно, что все слова, которые можно найти одновременно в словарях I и II, являются также общими и для словарей II и III. Задание. Можно ли на основании этих сведений утверждать, что все слова словаря I входят в словарь III? Докажите свою точку зрения.

№2. Чтобы проводить сортировку писем автоматически, Министерство связи ввело специальные индексы, которые состоят из цифр, имеющих вполне определённое начертание⁸. В состав каждой цифры могут входить отрезки девяти известных типов. Представьте себе, что машина «опознаёт» цифры, совершая ряд проверок. При каждой проверке определяется, есть ли в начертании цифры какой-то один отрезок; результатом каждой проверки является ответ «да» или «нет». После k проверок каждой цифре будет приписана цепочка из k ответов, например: «да-нет-нет-...» (В определённой последовательности для всех цифр проверяются одни и те же k отрезков, и для каждой цифры должны быть проверены все k отрезков.). Будем говорить, что цифры «опознаются» за k проверок, если в результате k проверок цепочки ответов оказываются у всех цифр различными. Задание. Каково минимальное число проверок, необходимое для опознания 10 цифр в принятом для них начертании?

№3. В устной речи обычно нет такого чёткого указателя границы между словами, как пробел в письменном тексте. Поэтому не всегда можно однозначно разделить текст на слова, особенно если это текст на незнакомом языке. Ниже приводится часть фразы на японском языке, записанная в русской транскрипции без пробелов между словами, и фрагмент словаря, содержащий, в частности, и все слова, входящие в данную фразу: какикуэбаканэганарунари.

в ак	ган	ка на у	
ака	ганар	каки	нари ун
аканэ	ганару	кан	нару унари наруна
аки	ганаруна	канэ	уэба

ан	и	канэга	нарунари	э
ана	ику	ки	нэ	эба
анэ	икуэ	кику	ри	
ари		кикуэ		
бака		ку		
бакан		куэ		
баканэ		куэба		

частности, и все слова, входящие в данную фразу: какикуэбаканэганарунари.

баканэ куэба

Задание. Определите, сколькими способами, используя приводимый словарь, можно было бы разделить данную фразу на слова. Объясните своё решение.

№4. В арабском языке 28 согласных звуков и 6 гласных, из них 3 кратких и 3 долгих. Слова в этом языке составляются из звуков по следующим правилам:

а) первый звук слова — согласный, второй — гласный; б) не могут встречаться два гласных подряд; в) не могут встречаться три согласных подряд; г) после долгого гласного

не могут идти два разных согласных подряд. Задание. Сколько различных слов, состоящих из 5 звуков, можно составить в соответствии с правилами этого языка?

№5. В гавайском языке имеются гласные звуки i, e, a, o, и, причём гласный звук может быть долгим или кратким, и согласные p, k, ʻ, m, n, w, l, h. Каждый слог состоит либо из одного гласного звука, либо из одного согласного и одного следующего за ним гласного. В слове не могут стоять рядом два гласных одинакового качества, например, «a» краткое и «a» долгое, два кратких «a» или два долгих «a». Задание 1. Определите, сколько можно составить двусложных гавайских слов, произносящихся по-разному. Задание 2. Гавайское письмо несовершенно: на письме не обозначается долгота и краткость гласных; кроме того, никак не отражается на письме согласный ʻ. Определите, сколько можно составить двусложных гавайских слов, которые пишутся по-разному. **Задание 3.** Приведите пример четырёхсложного слова, которое можно прочесть по-гавайски наибольшим числом способов.

Критерии оценки:

«Отлично», повышенный уровень: кейс-задание выполнено полностью, в рамках регламента, установленного на публичную презентацию, студент(ы) приводит (подготовили) полную четкую аргументацию выбранного решения на основе качественно сделанного анализа. Демонстрируются хорошие теоретические знания, имеется собственная обоснованная точка зрения на проблему(ы) и причины ее (их) возникновения. В случае ряда выявленных проблем четко определяет их иерархию. При устной презентации уверенно и быстро отвечает на заданные вопросы, выступление сопровождается приемами визуализации. В случае письменного отчета-презентации по выполнению кейс-задания сделан структурированный и детализированный анализ кейса, представлены возможные варианты решения (3-5), четко и аргументировано обоснован окончательный выбор одного из альтернативных решений;

«Хорошо», пороговый уровень: кейс-задание выполнено полностью, но в рамках установленного на выступление регламента, студент(ы) не приводит (не подготовили) полную четкую аргументацию выбранного решения. Имеет место излишнее теоретизирование, или наоборот, теоретическое обоснование ограничено, имеется собственная точка зрения на проблемы, но не все причины ее возникновения установлены. При устной презентации на дополнительные вопросы выступающий отвечает с некоторым затруднением, подготовленная устная презентация выполненного кейс-задания не очень структурирована. При письменном отчете-презентации по выполнению кейс-задания сделан не полный анализ кейса, без учета ряда фактов, выявлены не все возможные проблемы, для решения могла быть выбрана второстепенная, а не главная проблема, количество представленных возможных вариантов решения — 2-3, затруднена четкая аргументация окончательного выбора одного из альтернативных решений;

«Удовлетворительно», пороговый уровень: кейс-задание выполнено более чем на 2/3, но в рамках установленного на выступление регламента, студент(ы) расплывчато раскрывает решение, не может четко аргументировать сделанный выбор, показывает явный недостаток теоретических знаний. Выводы слабые, свидетельствуют о недостаточном анализе фактов, в основе решения может иметь место интерпретация фактов или предположения. Собственная точка зрения на причины возникновения проблемы не обоснована или отсутствует. При устной презентации на вопросы отвечает с трудом или не отвечает совсем. Подготовленная презентация выполненного кейс-задания не структурирована. В случае письменной презентации по выполнению кейс-задания не сделан детальный анализ кейса, далеко не все факты учтены, для решения выбрана второстепенная, а не главная проблема, количество представленных возможных вариантов решения — 1-2, отсутствует четкая аргументация окончательного выбора решения;

«Неудовлетворительно», уровень не сформирован: кейс-задание не выполнено, или выполнено менее чем на треть. Отсутствует детализация при анализе кейса, изложение

устное или письменное не структурировано. Если решение и обозначено в выступлении или отчете-презентации, то оно не является решением проблемы, которая заложена в кейсе.

Темы письменных работ (эссе, рефераты, курсовые работы и др.)

Перечень примерных тем рефератов/докладов

1. Проблемы автоматизации синтаксического анализа предложений.
2. Проблемы обнаружения кореференций и анафор в текстах на ЕЯ.
3. Применение алгоритмов и методов обработки текстовой информации в технике.
4. Применение алгоритмов и методов обработки текстовой информации в медицине.
5. Применение алгоритмов и методов обработки текстовой информации в системах безопасности.
6. Возможности программных приложений для анализа социальных сетей.
7. Проблемы автоматической идентификации авторов текстов.
8. Методы автоматического построения онтологий.
9. Сравнение алгоритмов обнаружения и исправления ошибок и опечаток.
10. Сравнение алгоритмов морфологического анализа.

Критерии оценки:

«Отлично», повышенный уровень: системность, обстоятельность и глубина излагаемого материала; знакомство с научной и научно-популярной литературой, рекомендованной к докладу преподавателем; письменная форма доклада (от руки); способность воспроизвести основные тезисы доклада без помощи конспекта; способность быстро и развернуто отвечать на вопросы преподавателя и аудитории; способность докладчика привлечь внимание аудитории. «Хорошо», пороговый уровень: развернутость и глубина излагаемого в докладе материала; знакомство с основной научной литературой к докладу; письменная форма доклада; при выступлении частое обращение к тексту доклада; некоторые затруднения при ответе на вопросы (неспособность ответить на ряд вопросов из аудитории). «Удовлетворительно», пороговый уровень: правильность основных положений доклада; наличие недостатка информации в докладе по целому ряду проблем; использование для подготовки доклада исключительно учебной литературы; неспособность ответить на несложные вопросы из аудитории и преподавателя; неумение воспроизвести основные положения доклада без письменного конспекта.

«Неудовлетворительно», уровень не сформирован: подготовка доклада в печатном виде с привлечением неизвестного информационного источника; поверхностный, неупорядоченный, бессистемный характер информации в докладе; при чтении доклада постоянное использование текста; выступление сбивчивое, с долгими паузами, монотонное; полное отсутствие внимания к докладу аудитории.

3.Оценочные средства для промежуточной аттестации

Вопросы к теоретическому зачёту

1. Множество лингвистических объектов
2. Действительные числа
3. Лингвистическое явление как математическая величина
4. Понятие функции
5. Числовые функции в лингвистике
6. Элементарные функции
7. Диахронический скачок и его моделирование с помощью элементарных функций
8. Моделирование информационного построения речи
9. Моделирование периодичности речи
10. Понятия бесконечно малой величины и предела в квантитативной лингвистике

11. Число Эйлера и модель роста словаря
12. Глоттохронология
13. Информационные модели слова и текста
14. Диахроническая скорость и понятие производной
15. Дифференциал
16. Исследование функций, аппроксимирующих лингвистические процессы
17. Основные понятия теории рядов
18. Каков максимальный объем информации в слове?
19. Лингвистические задачи, приводящие к понятию интеграла
20. Основные понятия интегрирования и применение их к лингвистическим задачам
21. Комбинаторные схемы
22. Лингвистическое событие
23. Вероятность элементарного лингвистического события
24. Вероятности сложных лингвистических событий
25. Информационные измерения в тексте
26. Повторение независимых испытаний в тексте
27. Случайная лингвистическая величина, ее характеристики и функция распределения
28. Законы распределения, моделирующие образование языковых единиц текста
29. Понятие о законе больших чисел
30. Статистическая совокупность лингвистических объектов и ее организация
31. Вариационные ряды лингвистических признаков
32. Статистические характеристики лингвистических вариационных рядов
33. Исследование лингвистических вариационных рядов с помощью эмпирических моментов
34. Точечная оценка параметров генеральной лингвистической совокупности
35. Оценка математического ожидания с помощью доверительного интервала и статистическая параметризация стилей
36. Доверительный интервал для дисперсии и среднего квадратического отклонения
37. Доверительные интервалы для вероятности качественного лингвистического признака
38. Оценка функции генерального распределения по данным лингвостатистического наблюдения
39. Элементы теории статистических гипотез
40. Гипотеза о лексической нормативности текста и ее проверка с помощью порядковых критериев
41. Проверка гипотез о характере расхождений статистических характеристик языков, функциональных стилей и подязыков с помощью параметрических критериев
42. Проверка статистических гипотез о тождестве двух лингвистических распределений
43. Распределение средних длин словоформ в языках мира
44. Доминантные смысловые единицы и элементы заполнения текста

Критерии итоговой оценки по дисциплине (зачет)

«Зачтено», повышенный уровень: теоретическое содержание дисциплины освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные рабочей программой дисциплины учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному;

«Зачтено», пороговый уровень: теоретическое содержание дисциплины освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных рабочей программой дисциплины учебных заданий выполнено, некоторые из выполненных заданий, возможно, содержат ошибки;

«Не зачтено», уровень не сформирован: теоретическое содержание дисциплины не освоено. Необходимые практические навыки работы не сформированы, все предусмотренные рабочей программой дисциплины учебные задания выполнены с грубыми ошибками. Дополнительная самостоятельная работа над материалом дисциплины не приведет к какому-либо значимому повышению качества выполнения учебных заданий.