


МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное
учреждение высшего образования
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Факультет инновационных технологий

УТВЕРЖДАЮ

Проректор по ОД

 Е.В. Луков



« 30 » сентября 2022г.

ПРОГРАММА

кандидатского экзамена по научной специальности
2.3.8. Информатика и информационные процессы

Программа кандидатского экзамена по научной специальности **2.3.8. Информатика и информационные процессы** рассмотрена и рекомендована к утверждению ученым советом *факультета инновационных технологий (ФИТ)*

протокол №06/22 от 24 сентября 2022 г.

Авторы-разработчики:

1. Шидловский Станислав Викторович, док. тех. наук, декан ФИТ
2. Шашев Дмитрий Вадимович, канд. тех. наук, доцент кафедры Управления качеством ФИТ

Руководитель ОП



С.В. Шидловский

1. Общие положения

На основании постановления Правительства Российской Федерации от 23.09.2013 № 842 «О порядке присуждения ученых степеней» кандидатские экзамены сдаются в соответствии с научной специальностью (научными специальностями) и отраслью науки, предусмотренными номенклатурой научных специальностей, по которым присуждаются ученые степени, утверждаемой Министерством науки и высшего образования Российской Федерации (далее – Минобрнауки России), по которым осуществляется подготовка (подготовлена) диссертации.

Кандидатский экзамен по специальной дисциплине в соответствии с темой диссертации на соискание ученой степени кандидата наук представляет собой форму оценки степени подготовленности соискателя ученой степени к проведению научных исследований по научной специальности **2.3.8. Информатика и информационные процессы** по техническим наукам (далее – кандидатский экзамен).

Программа кандидатского экзамена разработана на основе Паспорта научной специальности **2.3.8. Информатика и информационные процессы** (далее – Программа), утвержденного ВАК при Минобрнауки России <https://drive.google.com/drive/folders/1RNYkXhvAzaEF85GqxOH8HhbenJIoUMR7>.

Организация и проведение приема кандидатского экзамена осуществляется в соответствии с установленным в НИ ТГУ порядком.

Подготовка по Программе может осуществляться как самостоятельно, так и в рамках освоения соответствующей программы подготовки научных и научно-педагогических кадров в аспирантуре НИ ТГУ. Сдача аспирантом кандидатского экзамена является обязательным условием обучения и относится к оценке результатов освоения базовой дисциплины (модуля) образовательного компонента программы, осуществляемой в рамках промежуточной аттестации.

2. Структура кандидатского экзамена и шкала оценивания уровня знаний

Кандидатский экзамен проводится в форме устного экзамена по билетам продолжительностью один академический час и состоит из следующих частей:

1. Основные вопросы (не более трёх вопросов по содержанию курса «*Информатика и информационные процессы*»).

2. Дополнительные вопросы (не более трёх вопросов из 2-го раздела содержания Программы).

Оценка уровня знаний по каждому вопросу осуществляется по пятибалльной шкале со следующим принципом перерасчета:

«отлично» – 5 баллов;

«хорошо» – 4 балла;

«удовлетворительно» – 3 балла;

«неудовлетворительно» – 1-2 балла.

При оценивании ответов на каждый из вопросов экзаменационного билета учитываются следующие критерии:

| | |
|--|-----------|
| Ответ на вопрос исчерпывающий, продемонстрировано понимание и знание сути вопроса в полном объеме. Замечаний нет. | 5 баллов |
| Ответ на вопрос неполный, но раскрывающий основную суть вопроса, продемонстрировано понимание и знание вопроса в достаточном объеме. Замечания незначительные. | 4 балла |
| Ответ неполный с существенными замечаниями, знания по вопросу фрагментарные и частичные, в том числе и по тематике диссертационного исследования. | 3 балла |
| Ответ на вопрос отсутствует или дан неправильный | 1-2 балла |

Итоговая оценка за кандидатский экзамен выставляется решением экзаменационной комиссии:

«отлично» – при наличии не менее 80% 5-балльных ответов и отсутствии 3-2-1-балльных ответов;

«хорошо» – при наличии не менее 80% 4-балльных ответов и отсутствии 2-1-балльных ответов;

«удовлетворительно» – при наличии более 20% 3-балльных ответов и отсутствии 2-1-балльных ответов;

«неудовлетворительно» – при наличии 1-2 балльного ответа (или отказа отвечать на вопрос).

3. Перечень тем и вопросов для подготовки к сдаче экзамена

Раздел 1. Основные вопросы (по содержанию курса «Информатика и информационные процессы»).

Тема 1. Общие положения

1. Информатика как наука, изучающая информацию и ее свойства в естественных, искусственных и гибридных системах. Место информатики в системе наук. Информатика как обрабатывающая информацию отрасль индустрии и инфраструктурная область, ее роль и значение в ускорении научно-технического прогресса.

2. Понятие информационного продукта и информационной услуги. Классификация информационных продуктов и услуг. Жизненный цикл информационного продукта. Экономика информационных сетей. Методы управления производством и распределением информационных продуктов.

3. Информационные ресурсы. Принципы оценки информации как ресурса общества и объекта интеллектуальной собственности. Проблемы правового

регулирования научной интеллектуальной собственности. Государственная политика в области защиты информационных ресурсов общества.

Тема 2. Концептуальные модели

1. Машинное обучение. Задача машинного обучения. Объекты и признаки. Основные понятия: метод обучения, функционал качества, обобщающая способность, скользящий контроль. Алгоритмы классификации: анализ формальных понятий, метод опорных векторов, k ближайших соседей, Байесовские классификаторы, AdaBoost, скрытые модели Маркова, метод условных случайных полей. Факторный анализ. Алгоритмы кластеризации: k-средних, самоорганизующиеся карты Кохонена, графовые алгоритмы, иерархическая кластеризация.

2. Методы анализа текстовых данных. Приложения задач анализа текстовых данных: кластеризация, извлечение данных, выявление трендов. Алгоритмы выделения именованных сущностей, гиперонимов, описаний объектов. Разрешение неоднозначности. Алгоритмы выявления ассоциативных правил: a-priori, FP-growth.

3. Приближенные алгоритмы дискретной оптимизации. Примеры оптимизационных задач: минимальное трансверсальное множество, минимальное покрывающее множество. Жадные алгоритмы. Сведение к задаче линейного программирования. Алгоритмы со случайным выбором.

4. Алгоритмы анализа социальных сетей. Формализация понятия сообщества в социальной сети. Алгоритмы нахождения сообществ. Оценка авторитетности ресурсов. Алгоритм PageRank.

5. Методы представления знаний. Базы знаний. Общие принципы моделирования окружающей среды и мышления человека. Методы представления знаний: классификационные, тезаурусные, основанные на отношениях, семантические сети и фреймы, продукционные и непродукционные.

6. Онтологии. Введение в дескриптивную логику. Семейства логик. Характерные задачи. Логический вывод. Онтологии. Языки описания онтологий. Языки запросов к онтологиям.

Тема 3. Распознавание образов

1. Задача распознавания образов. Основные подходы: геометрический, вероятностный и комбинаторно-логический. Примеры задач распознавания. Геометрический подход. Линейные процедуры распознавания. Перцептроны. Теорема Новикова. Метод потенциальных функций.

2. Вероятностный подход. Процедура Байеса. Метод обобщенного портрета. Условия кластеризации. Основные процедуры построения кластеров. Метод скрытых марковских процессов.

3. Комбинаторно-логический подход. Линейные процедуры и информационные веса. Условия эффективности распознавания. Тесовые

процедуры распознавания. Алгоритм голосования. Оценки длины минимальных тестов.

4. Оценки числа тупиковых тестов для почти всех таблиц. Асимптотически оптимальный алгоритм построения множества коротких тестов. Полиномиальный характер решающих правил распознавания.

Тема 4. Управление данными и информационный поиск

1. Модели данных. Понятие модели данных. Иерархическая, сетевая модели данных, сравнительный анализ, противоречия и парадоксы. Реляционная модель данных. Экземпляры отношений, домены, атрибуты. Операции над отношениями: селекция, проекция, естественное соединение. Понятие реляционной полноты языка манипулирования данными. Модель данных «сущность-связь».

2. Базы данных. Независимость программ и данных. Интегрированное использование данных. Непротиворечивость данных. Целостность и защита данных. Структуры БД. Администрирование банков данных. Типы пользователей. Администратор БД. Понятие концептуальной, логической, физической структуры БД. Представления пользователей и подсхемы. Понятие о словарях данных, языках описания и манипулирования данными. БД и файловые системы.

3. Системы управления базами данных. Состав и структура. Типовые функции СУБД: хранение, поиск данных; обеспечение доступа из прикладных программ и с терминала конечного пользователя; преобразование данных; словарное обеспечение БД; импорт и экспорт данных из(в) файлов ОС ЭВМ. Языки описания и манипулирования данными SQL. Типовая структура СУБД: ядро, обрамление, утилиты, интерпретатор/компилятор пользовательского языка манипулирования данными.

4. Полнотекстовые БД. Физическая и логическая структура. Файл полного текста. Частотный словарь, инверсный файл. Положительный и отрицательный словари. Стандартные строки и словосочетания, включаемые в частотный словарь. Описание БД. Обработка текстов при загрузке БД. Понятие экспорта импорта документов-данных.

5. Построение распределенных баз данных. Основные способы построения распределенных приложений. Распределенные транзакции. Механизмы репликации.

6. Обеспечение безопасности в СУБД. Методы шифрования протоколов обмена данными. Построение виртуальных баз данных. Защита структур данных. Распределение прав доступа к данным. Политики распределения ресурсов в СУБД.

7. Обработка слабоструктурированных данных. Графовые и древовидные модели данных. Понятие путевого запроса. Языки запросов XPath и XQuery. Методы выполнения запросов.

8. Критерии оценки информационных технологий и систем. Оценки качества поиска (полнота, точность и др.). Скалярные и векторные оценки. Смешанные критерии (полезная работа, корреляционный критерий, свертки и пр.). Рабочие

характеристики информационно-поисковых систем (ИПС) в различных координатах. Вероятностная модель ИПС. Теоретико-множественная модель ИПС. Оптимизация режима ИПС. Линейное представление документов, запросов, индексирования, поиска.

9. Методы выполнения и оптимизации локальных и распределенных запросов. План выполнения запроса. Методы доступа к данным. Синтаксическая и стоимостная оптимизация. Адаптивные методы вычисления запросов.

10. Виртуальная интеграция данных. Задачи интеграции данных. Синтаксическая и семантическая неоднородность. Методы сопоставления схем данных. Перезапись запросов.

Тема 5. Элементы теории информации

1. Неопределенность и информация. Кодирование информации. Алфавитное кодирование. Теорема Маркова. Понятие энтропии стохастического источника. Измерение количества информации по Шеннону. Аксиоматика Шеннона-Хартли. Корректирующие свойства кодов. Коды Хэмминга, коды БЧХ. Теорема Шеннона о передаче при наличии помех.

2. Алгоритмический подход к понятию количества информации. Сложность конечного объекта по А.Н. Колмогорову. Существование оптимального способа описания. Количество информации. Сложность по Колмогорову и ее связь с шенноновской энтропией.

Тема 6. Элементы математической логики и теории алгоритмов

1. Логика первого порядка. Интерпретация формул. Нормальная форма Сколема. Проблема выполнимости. Теорема Черча. Метод резолюций.

2. Модели теории алгоритмов. Вычислимость. Интуитивное и формализованное понятие алгоритма. Машины Тьюринга, нормальные алгорифмы Маркова, частично-рекурсивные функции. Тезис Черча. Алгоритмически неразрешимые проблемы. Универсальный алгоритм.

3. Вычислимые и гёделевы (главные) нумерации вычислимых функций. Теорема Райса о неразрешимости проблемы распознавания любого нетривиального свойства вычислимой функции по ее описанию.

4. Сложность алгоритмов. Классы сложности P и NP, полиномиальная сводимость задач. Теория NP-полноты. Основные NP-полные задачи. Метод ветвей и границ и динамическое программирование в комбинаторных алгоритмах. Градиентные и другие приближенные алгоритмы для NP-полных задач. Теоремы о соотношении сложности и точности приближенных алгоритмов для NP-полных задач. Метод рекурсии и оценки сложности алгоритмов, использующих рекурсию. Алгоритм быстрого преобразования Фурье. Принцип жадности. Теорема Радона-Эдмонса об условиях оптимальности жадного алгоритма.

Тема 7. Формальные языки и конечные автоматы

1. Детерминированные и недетерминированные конечные автоматы. Регулярные языки и регулярные выражения. Свойства регулярных языков. Лемма о накачке для регулярных языков и ее применение. Теорема Клини. Эквивалентные состояния и минимизация автоматов. Структурные автоматы. Алгоритмическая неразрешимость проблемы полноты для автоматов.

2. Понятие формального языка. Примеры формальных языков. Задание языков конечными автоматами. Порождение языков формальными грамматиками. Операции над языками.

3. Понятие вывода в формальной грамматике. Язык, порождаемый грамматикой. Линейные и автоматные грамматики и их свойства. Эквивалентность автоматных грамматик и конечных автоматов.

4. Контекстно-свободные грамматики (КС-грамматики) и контекстно-свободные языки (КС-языки). Деревья разбора. Автоматы с магазинной памятью (МП-автоматы). Эквивалентность МП-автоматов и КС-грамматик. Свойства КС-языков.

5. Нормальная форма Бэкуса-Наура. Примеры для арифметических выражений. Понятие о LL и LR разборах. Процедура LR(1) разбора выражений, соответствующих формальной грамматике. Операции сдвига и свертки.

Тема 8. Программные и технические основы

1. Классификация языков программирования. Структурное программирование. Объектно-ориентированный подход. Императивное, функциональное и логическое программирование. Статическая и динамическая типизация. Компиляция и интерпретация программ.

2. Язык программирования Си. Организация программы. Аргументы командной строки. Схема трансляции программ на языке Си. Препроцессор и компоновщик. Типы данных. Переменные. Константы. Массивы. Размещение массивов в памяти. Строки. Арифметические выражения. Логические выражения. Операторы if, while, for, do while, switch. Функции. Передача параметров и возврат значений. Область видимости и время существования переменных. Глобальные и локальные переменные. Рекурсия. Ввод-вывод. Работа с файлами.

3. Представление чисел. Представление целых чисел. Представление вещественных чисел. Стандарт IEEE 754. Специальные значения. Порядок байт (big endian, little endian).

4. Операционные системы. Архитектура ЭВМ. Ядро операционной системы. Файловая система. Управление памятью. Процессы, группы процессов. Планирование процессов. Межпроцессное взаимодействие (сигналы, каналы, разделяемая память, сокеты). Поток, синхронизация потоков.

5. Архитектура TCP/IP. Модель OSI. Маршрутизация и топология сети. Маршрутизация в IP. Протокол UDP. Протокол TCP (установление и завершение соединения, трехэтапное рукопожатие, диаграмма состояний). Элементарные

сокеты TCP (пример приложения клиент/сервер). Архитектура WWW. Протокол HTTP.

6. Принципы построения распределенных приложений. Основные понятия: открытость, совместное использование ресурсов, конкуренция, масштабируемость, отказоустойчивость, прозрачность. Базовые распределенные алгоритмы: выбор ведущего процесса, взаимное исключение, достижение согласия, многоадресная передача.

7. Математические методы анализа программных систем. Формальные модели описания систем: конечные автоматы, сети Петри, линейные временные логики. Алгоритмические задачи проверки выполнимости свойств модели.

8. Принципы создания информационных систем в сети Интернет. Клиент-серверная и многоуровневая архитектура программных систем. Особенности Web-приложений. Шаблоны проектирования сервера приложений.

Рекомендуемая литература

1. Михайлов А.И., Черный А.И., Гиляревский Р.Э. Основы информатики. М.: Наука, 1978.

2. Попов И.И. Информационные ресурсы и системы: реализация, моделирование, управление. М.: ТПК «Альянс», 1996.

3. Шемякин Ю.И. Введение в информатику. М.: Финансы и статистика, 1985.

4. С.В.Алешин, Распознавание динамических образов, Изд-во МГУ, М., 1996

5. Шеннон К. Работы по теории информации и кибернетике. М., 1963.

6. А.Н. Колмогоров, Теория информации и теория алгоритмов, М., 1987

7. Х. Роджерс, Теория рекурсивных функций и эффективная вычислимость, М., 1972

8. Дж. Ульман, Основа систем баз данных, М. Финансы и статистика, 1983

9. Гери М., Джонсон Д., Вычислительные машины и труднорешаемые задачи. Москва, Мир, 1982 г.

10. Ахо А., Хопкрофт Дж., Ульман Дж., Построение и анализ вычислительных алгоритмов. Москва, Мир, 1970 г.

11. Хопкрофт Дж., Мотвани Р., Ульман Дж. Введение в теорию автоматов, языков и вычислений. М.: Издательский дом "Вильямс", 2008.

12. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. Том 1: Синтаксический анализ. М.: Мир, 1978.

13. Б.У. Керниган, Д.М. Ритчи. Язык программирования С. Вильямс, 2013.

14. М. Липовач. Изучай Haskell во имя добра. ДМК Пресс, 2012.

15. Р. Лав. Linux. Системное программирование. Питер, 2008.

16. У. Р. Стивенс, Б. Феннер, Э. М. Рудофф. UNIX. Разработка сетевых приложений, 3-е изд. Питер, 2007.

17. С. Фейт. TCP/IP. Архитектура, протоколы, реализация (включая IPv6 и IP Security). Лори, 2009.

18. Мартин Дж. Организация баз данных в вычислительных системах. М.: Мир, 2000.
19. Л.А. Калиниченко, Методы и средства интеграции неоднородных баз данных, М. Наука, 1983.
20. С.Д. Кузнецов. Базы данных. Академия, Серия: Университетский учебник, 2012.
21. Маннинг К., Рагхаван П., Шютце Х. Введение в информационный поиск. Вильямс, 2011.
22. Allemang D., Hendler J. Semantic web for the working ontologist: effective modeling in RDFS and OWL. – Elsevier, 2011.
23. Antoniou G., Van Harmelen F. A semantic web primer. – MIT press, 2004.
24. Baader F., Nutt W. Basic description logics //Description logic handbook. – 2003. – С. 43-95.
25. Т Кормен, Ч Лейзерзон, Р.Ривест. Алгоритмы, Построение и анализ. М 1999.
26. Х Пападимитриу, К. Стайглиц. Комбинаторная оптимизация. Алгоритмы и анализ. М 1985.
27. В.Б. Кудрявцев, Э.Э. Гасанов, А.С. Подколзин. Введение в теорию интеллектуальных систем. М. Изд-во МГУ, 2006.
28. В.Н. Сачков. Введение в комбинаторные методы дискретной математики. М.2004.
29. В.Б. Кудрявцев, С.В. Алешин, А.С. Подколзин. Введение в теорию автоматов. М.Наука. 1985.
30. В.Б.Кудрявцев, А.Е.Андреев, Э.Э.Гасанов. Теория тестового распознавания. М. ФизМатЛит, 2007.

Раздел 2. Дополнительные вопросы.

Тема 9. Информационная безопасность

1. Комплексный подход к обеспечению информационной безопасности (ИБ). Выделение уровней комплексного подхода к обеспечению ИБ: законодательный, административно-организационный, операционный и программно-технический. Обзор мер и методов на уровнях комплексного подхода к обеспечению ИБ.
2. Модели угроз и нарушителей. Основные принципы построения систем защиты информации в компьютерных системах. Модель угроз и модель нарушителя. Распространенные угрозы и атаки. Проектирование и выбор архитектуры системы защиты. Автоматизированные системы в защищенном исполнении.
3. Аутентификация. Цифровая подпись. Идентификация и аутентификация в компьютерных системах. Общие положения создания систем и механизмов идентификации и аутентификации. Примеры реализации методов идентификации и аутентификации в операционных системах, в веб-приложениях.
4. Криптография и криптоанализ. Обзор решаемых задач и основных методов. Понятия криптографической системы, криптографического алгоритма,

криптографического протокола. Обзор нормативной базы в области криптографии. Обзор типовых областей применения криптографических систем, алгоритмов и протоколов. Криптографическая стойкость. Обзор распространенных методов атак.

Тема 10. Системы принятия решений

1. Технология и процедуры разработки и принятия управленческих решений
2. Описание неопределенностей в теории принятия решений
3. Методы принятия решений
4. Моделирование в теории принятия решений.

Тема 11. Нейро-сетевые и нечеткие технологии

1. Нечеткие множества. Функции принадлежности. Лингвистические переменные. Термы. Операции над нечеткими множествами. Нечеткие высказывания и операции над ними. Свойства минимаксных операций над нечеткими множествами. Свойства алгебраических операций над нечеткими множествами.

2. Алгоритм Мамдани: активизация, импликация, агрегирование. Методы дефаззификации. Алгоритм Сугено. Нечеткий регулятор.

3. Математическая модель перцептрона

4. Архитектурные принципы построения полносвязной нейронной сети.

Пример сети.

5. Архитектурные принципы построения сверточной нейронной сети.

Пример сети.

6. Архитектурные принципы построения рекуррентной нейронной сети.

Пример сети.

Тема 12. Обработка цифровых изображений.

1. Математическая модель изображения и цветовые пространства.

2. Операторы низкоуровневой обработки изображений.

Рекомендуемая литература.

1. Галатенко В.А. Основы информационной безопасности: курс лекций: учебное пособие. ИНТУИТ. РУ "Интернет-университет Информационных Технологий", 2006.

2. Орлов А.И. Теория принятия решений. Учебное пособие / А.И.Орлов.- М.: Издательство «Экзамен», 2005. - 656 с.

3. Лотов А.В., Поспелова И.И. Многокритериальные задачи принятия решений: Учебное пособие. – М.: МАКС Пресс, 2008. – 197 с.

4. Шидловский С.В. Автоматическое управление. Перестраиваемые структуры. – Томск: Томский государственный университет, 2006. – 288 с.

5. Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А. Поспелова. – М.: Наука, 1986. – 312 с.

6. Компьютерное зрение. Современные методы и перспективы развития /ред. Р. Дэвис, М. Терк. – М.: ДМК Пресс, 2022. -690 с.

4. Пример экзаменационного билета

Билет №1

1. Машинное обучение. Алгоритмы классификации: анализ формальных понятий, метод опорных векторов, k ближайших соседей, Байесовские классификаторы, AdaBoost, скрытые модели Маркова, метод условных случайных полей.

2. Системы управления базами данных. Типовые функции СУБД: хранение, поиск данных; обеспечение доступа из прикладных программ и с терминала конечного пользователя; преобразование данных; словарное обеспечение БД; импорт и экспорт данных из(в) файлов ОС ЭВМ. Языки описания и манипулирования данными SQL.

3. Детерминированные и недетерминированные конечные автоматы. Регулярные языки и регулярные выражения. Свойства регулярных языков. Лемма о накачке для регулярных языков и ее применение. Теорема Клини. Эквивалентные состояния и минимизация автоматов. Структурные автоматы. Алгоритмическая неразрешимость проблемы полноты для автоматов.

4. Аутентификация. Цифровая подпись. Идентификация и аутентификация в компьютерных системах. Общие положения создания систем и механизмов идентификации и аутентификации. Примеры реализации методов идентификации и аутентификации в операционных системах, в веб-приложениях.

5. Алгоритм Мамдани: активизация, импликация, агрегирование. Методы дефаззификации. Алгоритм Сугено. Нечеткий регулятор.

6. Операторы низкоуровневой обработки изображений.