

**Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Нижегородский государственный лингвистический
университет им. Н. А. Добролюбова»**



ПРОГРАММА КАНДИДАТСКОГО ЭКЗАМЕНА

по специальности

05.13.17 - «Теоретические основы информатики»

(технические науки)

КЭ.А.03

**Нижегород
2011**

Введение

В основу настоящей программы положены следующие учебные дисциплины: основы информатики; прикладная информатика; теория вероятностей и математическая статистика; актуальные элементы теории случайных процессов; вычислительная техника и языки программирования; технология и методы программирования; аппаратные средства ВТ; управление информационной безопасностью; сети и системы передачи информации; теория информации и её приложения; теория информационных систем и технологий; математическое моделирование речевых сигналов средствами ВТ; математическое моделирование социально-экономических процессов; защита информационных процессов в компьютерных системах; новые речевые технологии (в системах защиты информации). Программа разработана коллективом кафедры математики и информатики НГЛУ под руководством д.т.н. профессора В.В. Савченко.

Программа отвечает требованиям паспорта специальности *05.13.17 – Теоретические основы информатики* и, вместе с тем, учитывает профиль научной школы, сложившейся в НГЛУ (рук. проф. В.В. Савченко). Тематика диссертационных исследований аспирантов соответствует следующим областям паспорта специальности:

5. Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечения разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений. 6. Разработка методов, языков и моделей человеко-машинного общения; разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения данных из текстов на естественном языке. 7. Разработка методов распознавания образов, фильтрации, распознавания и синтеза изображений, решающих правил. Моделирование формирования эмпирического знания. 8. Исследование и когнитивное моделирование интеллекта, включая моделирование поведения, моделирование рассуждений различных типов, моделирование образного мышления. 9. Разработка новых интернет- технологий, включая средства поиска, анализа и фильтрации информации, средства приобретения знаний и создания онтологии, средства интеллектуализации бизнес-процессов. 10. Разработка основ математической теории языков и грамматик, теории конечных автоматов и теории графов. 11. Разработка методов обеспечения высоконадежной обработки информации и обеспечения помехоустойчивости информационных коммуникаций для целей передачи, хранения и защиты информации; разработка основ теории надежности и безопасности использования информационных технологий. 12. Разработка математических, логических, семиотических и лингвистических моделей и методов взаимодействия информационных процессов, в том числе на базе специализированных вычислительных систем. 13. Применение бионических принципов, методов и моделей в информационных технологиях.

Структура экзамена

Кандидатский экзамен по специальности 05.13.17 - «Теоретические основы информатики» (технические науки) проводится в устной форме и предполагает ответ аспиранта на три вопроса экзаменационного билета, а также на дополнительные вопросы по цели, задачам и основным результатам кандидатской диссертации.

Этапы проведения экзамена:

1. Подготовка аспиранта к ответу на вопросы билета (30 минут).
2. Ответ аспиранта на вопросы экзаменационного билета (20 минут).
3. Ответ аспиранта на дополнительные вопросы экзаменаторов (10 минут).

Оценка ответа аспиранта определяется как средняя из оценок ответов на каждый из вопросов экзаменационного билета и на дополнительные вопросы при условии, что все оценки положительные.

Содержание экзамена

1. Информатика как наука, отрасль промышленности и инфраструктурная область

Информатика — наука, отрасль индустрии и инфраструктура. 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.

Предметная область информатики. Информационные проблемы современного этапа научно-технической революции. Информационные потребности индивидуальных и коллективных пользователей. Информационные коммуникативные процессы. Современная информационная технология на базе широкого применения вычислительной техники и связи. Социальные аспекты информатизации и компьютеризации общества.

Информационные технологии и системы, их определение, назначение и классификация. Информационные системы и их основные характеристики. Общая модель (топология) информационных систем. Их разновидности и классификация. Информационные технологии.

Системы связи и их характеристики эффективности. Виды применяемых каналов связи.

Новые речевые технологии. Области применения речевых технологий в интеллектуальных и экспертных системах. Системы определения индивидуальности диктора по речевым характеристикам. Проверка прав доступа к различным информационным и физическим системам. Криминалистическая экспертиза: анализ записей переговоров при различных аварийных ситуациях; анализ записей телефонных переговоров при санкционированном их прослушивании; доказательство в суде. Ключевая роль фонетических баз данных (ФБД) национального языка, как отдельных его носителей с их индивидуальными особенностями речи, так и произвольных групп разных дикторов.

2. Концептуальные модели информатики

Информационная теория восприятия речи как основа новых речевых технологий. Вероятностный или теоретико-информационный подход в задачах автоматической обработки устной речи. Проблемы вариативности устной речи (под действием целого ряда факторов: времени суток, эмоционального состояния диктора и т.п.). Объединение близких друг другу по своему звучанию, однородных в теоретико-информационном смысле элементарных речевых единиц в соответствующие фонемы-кластеры. Главный постулат информационной теории восприятия речи (ИТВР): условно говоря, человеческий мозг объединяет и запоминает как нечто целое (в виде абстрактного образа) разные образцы (произношения) каждого отдельного слова в соответствующей «сфере» своей памяти вокруг абстрактного «центра» с заданным «радиусом»

Автоматические и экспертные системы. Решения задачи идентификации на основе автоматических систем и субъективных (экспертных) подходах. Системы поиска ключевых слов в слитной речи. Системы ограничения прав доступа с верификацией пользователя.

3. Математические основы информатики

Математические модели информационных технологий и систем: описание, оценка, оптимизация. Модели описания информационных процессов и технологий.

Энтропия как численная мера неопределенности случайных явлений: Дискретная случайная величина и ее энтропия по Шеннону. Свойства энтропии. Условная энтропия и ее свойства. Пример дискретной случайной величины: равновероятный закон распределения.

Дифференциальная энтропия непрерывной случайной величины. Непрерывная случайная величина и ее энтропия. Определение дифференциальной энтропии. Примеры вычислений. Свойства дифференциальной энтропии. Понятие относительной энтропии.

Дифференциальная энтропия системы случайных величин. Примеры: равномерный и гауссовский законы распределений.

Модель случайного процесса в задачах обработки информации
Энтропия случайных процессов. Случайный процесс и его основные статистические характеристики. Энтропия случайного процесса и ее свойства. Определение удельной энтропии. Случайный гауссовский процесс и его энтропия. Вычисление энтропии по спектральной плотности мощности.

Принцип максимума энтропии и экстремальные распределения. Постановка вариационной задачи. Информационный критерий. Метод множителей Лагранжа. Решение задачи для случая ограниченной области определения случайной величины X . Равномерный закон распределения. Решение вариационных задач с ограничениями на значения математического ожидания и дисперсии случайной величины. Особая роль гауссовского распределения.

Шенноновское количество информации. Определение количества информации для случайных дискретных и непрерывных объектов наблюдений. Основные свойства количества информации. Единицы измерения энтропии и количества информации.

4. Технические средства информатики и информационных технологий

Информационные характеристики сигналов и каналов связи. Сигнал и его виды в информационных системах. Информационная емкость сигнала. Избыточность сигналов. Скорость передачи информации и пропускная способность канала связи. Помехоустойчивость, эффективность и надежность.

Дискретные информационные системы без шумов. Задача оптимального кодирования информации. Фундаментальная теорема К.Шеннона. Примеры оптимальных кодов Шеннона-Фэно. Дискретные системы связи с шумами. Идеи помехоустойчивого кодирования сообщений. Коды с корректирующими свойствами.

Первая и вторая теоремы Шеннона. Обратная теорема Шеннона о кодировании сообщений при действии шумов. Пропускная способность каналов связи с шумами. Скорость передачи информации и пропускная способность непрерывного канала связи.

Постановка и решение оптимизационной задачи. Вывод формулы Шеннона для пропускной способности.

5. Программные средства информатики и информационных технологий

Теоретико-информационный подход в задаче оптимального прогнозирования. Математическая постановка оптимизационной задачи. Вывод формулы условного математического ожидания Проблема априорной

неопределенности. Линейная оценка прогнозирования. Ее связь с авторегрессионной моделью наблюдений. Проблема ее оптимальности (адекватности). Теоретико-информационный подход к задаче оптимизации линейной оценки прогнозирования. Критерий минимума взаимной энтропии. Синтез адаптивного алгоритма. Его информационный показатель эффективности в метрике Кульбака-Лейблера.

Оптимальное оценивание плотностей случайных временных рядов на основе принципа максимума энтропии. Задача восстановления (оценивания) неизвестной спектральной плотности мощности. Критерий максимума удельной энтропии. Синтез адаптивного алгоритма. Оценивание многомерной плотности вероятности. Принцип максимума энтропии. Синтез оптимального алгоритма.

Теоретико-информационный подход в задаче автоматического распознавания речи. Модель случайного гауссовского процесса при анализе речи. Взаимная энтропия двух гауссовских процессов. Критерий минимального информационного рассогласования. Синтез оптимального алгоритма и его программная реализация. Модель АР-процесса и метод обескураживающего фильтра.

Информационная теория колебаний рыночной конъюнктуры. Понятие рыночной конъюнктуры, линейная авторегрессионная модель ее колебаний. Общесистемный принцип минимума информационного рассогласования. Понятия кластера однотипных колебаний рыночных цен в динамике и типологической базы данных. Задача автоматической диагностики текущего состояния рынка. Синтез алгоритма. Примеры его практического применения на фондовых рынках США и России.

6. Информационное и лингвистическое обеспечение информационных технологий

Системы определения индивидуальности диктора по речи. Разделение задач определения личности диктора на верификацию и идентификацию. Ошибки 1-го рода (захват ложной цели или, принятие злоумышленника за зарегистрированного пользователя) и ошибки 2-го рода (пропуск цели или отказ признать зарегистрированного пользователя). Идентификация, когда имеется ограниченная и строго контролируемая группа пользователей системы. Групповая идентификация. Текстозависимые и текстонезависимые системы идентификации.

Перспективы применения новых речевых технологий в информационных системах. Перспективы для решения задач идентификации и верификации диктора. Новые решающие правила на основе ИТВР. Создание полноценной речевой базы данных, на которой можно проводить исследования и

проверять эффективность предлагаемых решений. Поиск новых признаков описания речевого сигнала, связанных с нелинейными моделями и определением характеристик голосового источника и формы артикуляционного тракта. Анализ эргономических проблем, связанных с решением конкретных задач.

7. Телекоммуникационное обеспечение информационных технологий

Информационные технологии и их поддержка. Примеры информационных технологий: SAP/R3, операционный день банка, рабочее место брокера. Проектирование и разработка информационных технологий. Государственные стандарты на разработку и создание информационных систем. CASE-технологии создания информационных систем. Стандарт ITIL.

8. Правовое обеспечение информатики и информационных технологий

Технология защиты информации. Основные угрозы информации в компьютерных системах. Ценности, опасности, потери, риски, угрозы в компьютерных системах. Основные угрозы информации в компьютерных системах; специфика возникновения угроз в открытых сетях; особенности защиты информации на узлах компьютерной сети; системные вопросы защиты программ и данных. Анализ рисков. Модель противника, возможности противника; параллельный анализ целей и возможностей злоумышленника в компьютерной сети. Анализ критических технологий.

Государственная политика в области безопасности компьютерных систем. Основные категории требований к программной и программно-аппаратной реализации средств защиты информации. Система лицензирования и сертификации средств защиты. Аттестация защищенных систем. Структуры в РФ, обеспечивающие лицензирование и сертификацию. Нормативная база и ответственность за защиту информации в компьютерных системах. Руководящие документы Гостехкомиссии и ФСТЭК по оценке защищенности информационных технологий, систем. Политика безопасности для компьютерных систем. Дискреционная политика безопасности. Многоуровневая политика. Политика защиты целостности.

Американские и европейские стандарты по защите информации. Стандарт "Розовая книга". Построение гарантированно защищенных баз данных и их оценка по стандарту "Оранжевая книга". Система ITSEC - функциональные требования. Вопросы гарантий и эффективности.

Общие критерии оценки защищенности информационных технологий - Common Criteria (CC). Базовые концепции. Профиль защиты. Функции поддержки политики безопасности. Гарантии безопасности. Требования по безопасности информационных технологий. Оценки защищенности.

Компоненты подсистем поддержки политики безопасности. Классы в СС. Требования к подсистемам аудита. Подсистемы подтверждения подлинности отправки и получения сообщения. Подсистемы разграничения доступа. Подсистемы аутентификации. Подсистемы защиты функций защиты. Подсистемы защиты ресурсов системы. Подсистемы защиты связи. Каналы утечки и их анализ по СС. Виды каналов утечки. Методология анализа каналов утечки информации. Безопасное функционирование по СС. Управление конфигурацией. Безопасная установка систем защиты информационных технологий. Безопасная модернизация информационных технологий.

Основная литература

1. *Кудряшов Б.Д.* Теория информации: Учебник для вузов. – Изд. Дом Питер. 2009.
2. *Хохлов Г.И.* Основы теории информации. М.: Изд. центр «Академия». 2008.
3. Информационные системы в экономике: / Под ред. Г.А. Титоренко. - 2-е изд., перераб. и доп. - М.: ЮНИТИ-ДАНА, 2007
4. *Бабешко Л.О.* Математическое моделирование финансовой деятельности: учеб. пособие / Л. О. Бабешко; Фин. акад. при Правительстве РФ. - М.: Кнорус, 2009.
5. *Потапова Р.К.* Речь: коммуникация, информация, кибернетика: Учебное пособие. – М.: Эдиториал УРСС, 2004.
6. *Ярочкин В.И.* Информационная безопасность: учебник для вузов. - 5-е изд. - М.: Академический проект, 2008.
7. *Мартин Дж.* Организация баз данных в вычислительных системах. М.: Мир, 2000.
8. *Макарова Н.В.* Информатика: учебник для вузов / Н. В. Макарова, В. Б. Волков. - СПб.: Питер, 2011. - 573 с.
9. *Попов И.И., Максимов Н.В., Храмов П.Б.* Введение в сетевые информационные ресурсы и технологии: Учеб. пособие для вузов. М.: Изд-во РГГУ, 2001.
10. *Савченко В.В.* Теория вероятностей: Конспект лекций – Н.Новгород: НГТУ, 1997.
11. *Савченко В.В.* Теория вероятностей и математическая статистика: Конспект лекций – Н.Новгород: НГЛУ, 2003.
12. *Савченко В.В., Грушин В.А.* Актуальные главы высшей математики: Конспект лекций и планы семинарских занятий. – Н.Новгород: НГЛУ, 2003.
13. *Савченко В.В.* Математика и информатика: Учеб. пособие для студ. лингвист. специальностей / В. В. Савченко, В. А. Никольская; ГОУ ВПО НГЛУ им. Н.А. Добролюбова. - Н.Новгород: НГЛУ, 2008
14. *Родькина О.Я.* Информационные системы в подготовке принятия решений: Учеб. пособие /; ГОУ ВПО НГЛУ Н.Новгород: НГЛУ, 2007

15. *Галеева И.С.* Интернет как инструмент библиографического поиска / И. С. Галеева. - СПб.: Профессия, 2007. - 245 с
16. *Потапова Р.К.* Язык, речь, личность / Р. К. Потапова, В. В. Потапов. - М.: Языки славянской культуры, 2006. - 491с.
17. *Степанов Е.А.* Информационная безопасность и защита информации: Учеб.пособие / Е. А. Степанов, И. К. Корнеев. - М.: ИНФРА-М, 2001. - 301с.

Дополнительная литература

1. Основы государства и права: Учеб. пособие для вузов / Под ред. О.Е. Кутафина. М.: Юрист, 1994.
2. *Попов И.И.* Автоматизированные информационные системы (по областям применения): Учеб. пособие для вузов. М.: Изд-во РЭА им. Г.В. Плеханова, 1999.
3. *Феллер В.* Введение в теорию вероятностей и ее приложения. В 2 т. М.: Мир, 1982.
4. *Савченко В.В.* Принцип минимакса энтропии в задаче многомерного спектрального анализа.// Радиотехника и электроника (издание РАН),1990, № 8.
5. *Савченко В.В.* Принцип минимакса энтропии в задачах статистических решений// Радиотехника и электроника (издание РАН),1990, № 9.
6. *Савченко В.В.* Рекуррентный метод восстановления многомерного закона распределения по принципу ММЭ //Известия вузов. Радиофизика,1991, № 3.
7. *Савченко В.В.* Прогнозирование социально-экономических процессов на основе спектрального оценивания.// АВТОМЕТРИЯ (издание СО РАН),1999, № 3.
8. *Савченко В.В.* Теоретико-информационное обоснование линейных оценок прогнозирования// АВТОМЕТРИЯ (издание СО РАН),2001, № 5.
9. *Савченко В.В.* Автоматическая обработка речи по критерию МИР// Радиотехника и электроника (издание РАН),2005, № 3.
10. *Орлов В.А., Филиппов Л.И.* Теория информации в упражнениях и задачах. М.: Высш. школа. 1976.

Список вопросов к экзамену

1. Понятие энтропии случайного процесса. Пример вычислений для гауссовского распределения.
2. Многомерный гауссовский закон распределения. Автокорреляционная матрица. Энтропия гауссовского случайного процесса.
3. Расчет энтропии через спектральную плотность мощности. Статистические оценки гауссовского процесса. Особая роль гауссовского распределения.
4. Математическая постановка задачи прогнозирования. Вывод формулы условного математического ожидания Проблема априорной

- неопределенности.
5. Линейная оценка прогнозирования. Ее связь с авторегрессионной моделью наблюдений. Проблема ее оптимальности (адекватности).
 6. Теоретико-информационный подход к задаче оптимизации линейной оценки прогнозирования. Критерий минимума взаимной энтропии.
 7. Синтез адаптивного алгоритма прогнозирования. Его информационный показатель эффективности в метрике Кульбака-Лейблера.
 8. Модель случайного гауссовского процесса при анализе речи.
 9. Взаимная энтропия двух гауссовских процессов.
 10. Критерий минимального информационного рассогласования. Синтез оптимального алгоритма и его программная реализация.
 11. Модель АР-процесса и метод обеляющего фильтра.
 12. Задача восстановления (оценивания) неизвестной спектральной плотности мощности. Критерий максимума удельной энтропии. Синтез адаптивного алгоритма.
 13. Оценивание многомерной плотности вероятности. Принцип максимума энтропии. Синтез оптимального алгоритма.
 14. Модель случайного гауссовского процесса при анализе речи.
 15. Взаимная энтропия двух гауссовских процессов.
 16. Критерий минимального информационного рассогласования. Синтез оптимального алгоритма и его программная реализация.
 17. Оценивание многомерной плотности вероятности. Принцип максимума энтропии. Синтез оптимального алгоритма.
 18. Задача типологического анализа динамики фондового рынка.
 19. Области применения речевых технологий при защите информации.
 20. Системы определения индивидуальности диктора по речевым характеристикам.
 21. Проверка прав доступа к различным информационным и физическим системам.
 22. Криминалистическая экспертиза: анализ записей переговоров при различных аварийных ситуациях; анализ записей телефонных переговоров при санкционированном их прослушивании; доказательство в суде.
 23. Ключевая роль фонетических баз данных (ФБД) национального языка.
 24. Классификация систем определения индивидуальности диктора по речи.
 25. Разделение задач определения личности диктора на верификацию и идентификацию.
 26. Ошибки 1-го и 2-го рода.
 27. Идентификация, когда имеется ограниченная и строго контролируемая группа пользователей системы.
 28. Групповая идентификация.
 29. Текстозависимые и текстонезависимые системы идентификации.
 30. Автоматические и экспертные системы защиты информации.

31. Решения задачи идентификации на основе автоматических систем и субъективных (экспертных) подходах.
32. Системы поиска ключевых слов в слитной речи.
33. Системы ограничения прав доступа с верификацией пользователя.
34. Информационная теория восприятия речи как основа новых речевых технологий.
35. Вероятностный или теоретико-информационный подход в задачах автоматической обработки устной речи.
36. Проблемы вариативности устной речи.
37. Объединение близких друг другу по своему звучанию, однородных в теоретико-информационном смысле элементарных речевых единиц в соответствующие фонемы-кластеры.
38. Главный постулат информационной теории восприятия речи.
39. Перспективы применения новых речевых технологий в системах защиты информации.
40. Перспективы для решения задач идентификации и верификации диктора.
41. Новые решающие правила на основе ИТВР.
42. Создание полноценной речевой базы данных, на которой можно проводить исследования и проверять эффективность предлагаемых решений.
43. Поиск новых признаков описания речевого сигнала, связанных с нелинейными моделями и определением характеристик голосового источника и формы артикуляционного тракта.

Программа кандидатского экзамена по специальности 05.13.17 – «Теоретические основы информатики» разработана на основе Федеральных государственных требований к Основной профессиональной образовательной программе послевузовского профессионального образования для обучающихся в аспирантуре, утвержденных приказом Министерства образования и науки Российской Федерации от 16 марта 2011 г. № 1365 (зарегистрирован Министерством юстиции Российской Федерации 10 мая 2011 г., регистрационный № 20700).

Программу составил:  к. т. н., проф. Д.Ю. Акатъев

Программа обсуждена и одобрена на заседании кафедры математики и информатики «29» августа 2011 г., протокол № 1.

Зав. кафедрой
математики и информатики  д.т.н., проф. В.В. Савченко

Программа утверждена _____
(дата)

Проректор по научной работе



Е.С. Гриценко