



Национальный исследовательский университет «Высшая школа экономики»
Программа кандидатского экзамена по научной специальности «1.2.3. Теоретическая информатика, кибернетика»

Федеральное государственное автономное образовательное
учреждение
высшего образования «Национальный исследовательский университет
«Высшая школа экономики»

УТВЕРЖДАЮ

Проректор

_____ С.Ю. Рощин

Одобрено на заседании
Академического совета
Аспирантской школы по
компьютерным наукам
Протокол № ___ от ____

Согласовано

Академический директор
Аспирантской школы по
компьютерным наукам

Программа кандидатского экзамена по научным специальности

1.2.3. Теоретическая информатика, кибернетика



1. Область применения и нормативные ссылки

Настоящая программа разработана в соответствии с Программой-минимум кандидатского экзамена по специальности 1.2.1 «Искусственный интеллект и машинное обучение» и Паспорта научной специальности 1.2.1 «Искусственный интеллект и машинное обучение».

2. Структура кандидатского экзамена

Форма проведения испытания:

Экзамен проводится в устной форме.

Продолжительность испытания:

не более одного часа

Структура кандидатского экзамена:

Экзамен проводится в виде беседы в свободной форме по выбранным экзаменуемым двум пунктам программы из части 3 настоящей Программы и по теме диссертационного исследования экзаменуемого.

Оценка уровня знаний (баллы):

Ответ оценивается по десятибалльной шкале (0 – 10). Итоговая оценка выставляется по пятибалльной шкале («неудовлетворительно» – «отлично») по следующему принципу пересчета:

«отлично» — 8 – 10 баллов (по 10-балльной шкале);

«хорошо» — 6 – 7 баллов (по 10-балльной шкале);

«удовлетворительно» — 4 – 5 баллов (по 10-балльной шкале);

«неудовлетворительно» — 0 – 3 балла (по 10-балльной шкале).

3. Содержание

1. Теория информации

Измерение информации дискретных источников. Энтропия. Неравномерное кодирование. Неравенство Крафта. Прямая и обратная теоремы побуквенного кодирования. Коды Хаффмена, Шеннона, Гилберта-Мура. Задача универсального кодирования. Двухпроходное побуквенное кодирование, нумерационное, адаптивное кодирование. Сравнение алгоритмов. Монотонные коды. Интервальное кодирование и метод «стопка книг». Метод скользящего словаря. Алгоритм LZW. Предсказание по частичному совпадению. Алгоритм Барроуза-Уиллера. Сравнение способа кодирования. Характеристики архиваторов.

2. Формальные языки и грамматики

Основные понятия и определения формальных языков и грамматик. Классификация грамматик и языков по Хомскому. Порождающие и аналитические (распознающие) грамматики. Регулярные грамматики, конечные автоматы и регулярные выражения. Минимизация детерминированных конечных автоматов. Построение детерминированного автомата, эквивалентного данному недетерминированному автомату. Лемма о разрастании для конечных автоматов и ее применение. Контекстно-свободные грамматики и деревья вывода. Нормальные формы контекстно свободных



грамматик. Автоматы с магазинной памятью.

3. Теория сложности алгоритмов и вычислений.

Машины Тьюринга, частично рекурсивные функции, машины с произвольным доступом к памяти (РАМ-машины). Тезис Черча. Неразрешимость проблемы останова машины Тьюринга. Анализ сложности алгоритмов. Классы задач P и NP, EXP, PSPACE, EXPSPACE, примеры. Сводимость задач по Карпу и Тьюрингу. NP-полнота. Теорема Кука – Левина. Классы задач по памяти: L, NL, coNL, PSPACE. Их соотношение с классами задач по времени. Вероятностные алгоритмы. Классы задач BPP, ZPP, RP. Подходы к проектированию алгоритмов: «разделяй и властвуй», динамическое программирование, жадная стратегия. Алгоритмы сортировки, двоичного поиска, быстрое возведение в степень, вычисление расстояния Левенштейна. Двоичные деревья поиска, кучи, хеш- таблицы.

4. Математическая теория языков и грамматик.

Основные понятия и определения формальных языков и грамматик. Классификация грамматик и языков по Хомскому. Порождающие и аналитические (распознающие) грамматики. Регулярные грамматики, конечные автоматы и регулярные выражения. Минимизация детерминированных конечных автоматов. Построение детерминированного автомата, эквивалентного данному недетерминированному автомату. Лемма о разрастании для конечных автоматов и ее применение. Контекстно-свободные грамматики и деревья вывода. Нормальные формы контекстно свободных грамматик. Автоматы с магазинной памятью.

5. Представление знаний

Классификационные системы: иерархические классификации, фасетные классификации, алфавитно-предметные классификации. Тезаурусные методы представления знаний. Системы, основанные на отношениях. Объектно-характеристические таблицы. Предикатно-октантные структуры. Семантические сети. Понятие «сущности». Семантические отношения и их виды. Лингвистические, логические, теоретико-множественные, квантификационные отношения. Абстрактные и конкретные семантические сети. Фреймы-системно-структурное описание предметной области. Принципы фрейм-представлений. Понятие «СЛОТА». Продукционные системы представления знаний. Канонические системы Поста. Представление неформальных знаний. Редукционные системы. Синтез плана решения задач с автоматическим построением редукционной модели.

6. Представление данных

Обработка данных. Структуры данных. Уровни представления данных. Языка описания и манипулирования данными. Система управления базами данных. Архитектура СУБД. Основные конструкции структур данных. Функции СУБД. Категории пользователей. Классы структур данных. Иерархическая структура. Сетевые структуры. Реляционные структуры. Системы управления базами данных (СУБД), 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 файлы ОС ЭВМ. Типовая структура СУБД: ядро, обрамление, утилиты, интерпретатор / компилятор пользовательского языка манипулирования данными. Среда конечного пользователя.

7. Базы данных.

Основные понятия. 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. Документальные и фактографические базы данных, базы знаний. 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. 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".

8. Программные средства

Операционные системы и их функции. Управление задачей и использование памяти. Понятие тома и файла данных. Сообщения операционной системы. Команды и директивы оператора.

Системы программирования. Понятие разработки приложений. Состав системы программирования: язык программирования (ЯП); обработчик программ; библиотека программ и функций. История развития и сравнительный анализ ЯП. Типы данных. Понятие блока и процедуры. Операторы ЯП. Стандартные арифметические, логические, строчные функции.

Программирование в средах современных информационных систем: создание модульных программ, элементы теории модульного программирования, объектно-ориентированное проектирование и программирование. Объектно-ориентированный подход к проектированию и разработке программ: сущность объектно-ориентированного подхода; объектный тип данных; переменные объектного типа; инкапсуляция; наследование; полиморфизм; классы и объекты. Логическое программирование. Компонентное программирование.



Литература

1. Кудряшов Б. Д. Теория информации, СПбГУ НИУ ИТМО
2. Гладкий А.В. Формальные грамматики и языки. — М.: Наука, 1973.
3. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. — М.: Мир, 1982.
4. Кузнецов О.П. Дискретная математика для инженера. — М.: Лань, 2004.
5. Ларичев О.И. Теория и методы принятия решений. — М.: Логос, 2002.
6. Басакер Р., Саати Т. Конечные графы и сети. — М.: Наука, 1974.
7. Мартыненко Б.К. Языки и трансляции: учеб. пос. СПб.ГУ, 2002 г.
8. Шень А. Х. Программирование: теоремы и задачи. М.: МЦНМО, 2004.
9. Хопкрофт Дж., Мотвани Р., Ульман Дж. Введение в теорию автоматов, языков и вычислений. — 2-е изд.: Пер. с англ. — М.: Вильямс, 2008.
10. Sipser M. Introduction to the Theory of Computation. — Boston, Mass.: Thomson Course Technology, 2006
11. Гарсиа-Молина Г., Ульман Дж., Уидом Дж. Системы баз данных. Полный курс. — Вильямс, 2003.
12. Дасгупта С., Пападимитриу Х., Вазирани У. Алгоритмы. — М.: МНЦМО, 2014. 13. Дейт К. Дж. Введение в системы баз данных. — 8-е изд. — М.: Вильямс, 2005.
13. Кормен Т.Х., Лейзерсон Ч.И., Ривест Р.Л., Штайн К. Алгоритмы: построение и анализ. — 2-е издание: Пер. с англ. — М.: Вильямс, 2007.
14. Крупский В.Н. Введение в сложность вычислений. — М.: Факториал
15. Hermann Helbig: Knowledge Representation and the Semantics of Natural Language, Springer, Berlin, Heidelberg, New York 2006
16. Arthur B. Маркман: Knowledge Representation Lawrence Erlbaum Associates, 1998
17. John F. Sowa: Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks/Cole: New York, 2000
18. Adrian Walker, Michael McCord, John F. Sowa, and Walter G. Wilson: Knowledge Systems and Prolog, Second Edition, Addison-Wesley, 1990