Демонстрационный вариант и методические рекомендации по направлению 080500 «Бизнес-информатика»
Профиль “Big Data Systems” (Системы больших данных)

Время выполнения задания – 180 мин.
Ответы могут быть даны на русском или английском языке

1. Дайте развернутые ответы

1. Advantages and disadvantages of using models of cloud computing in the enterprise
2. Виды информации и формы ее представления, использующиеся при принятии управленческих решений на предприятии

2. Решите задачи.

1. В реляционной базе данных предприятия хранятся данные о сотрудниках и выполненных ими проектах. В базе хранится следующая информация о сотрудниках: ФИО; дата рождения; пол; адрес проживания; телефон; наименование подразделения, в котором работает сотрудник; должность. Каждый сотрудник может работать только в одном подразделении, однако, может участвовать в одном или нескольких проектах. Информация о проектах включает: наименование проекта, дату начала, дату окончания, руководителя проекта, исполнителей проекта.
   Требуется:
   1) используя любую общепринятую нотацию, нарисовать схему базы данных, удовлетворяющую третьей нормальной форме, с выделением первичных и внешних ключей, типа и направления связи;
   2) сделать подробное описание таблиц с расшифровкой имен полей, типов и свойств данных;
   3) используя операторы языка SQL, написать один запрос для получения следующей информации:
      наименование подразделения, список сотрудников подразделения с указанием выполняемых ими проектов и роли в проекте (руководитель/исполнитель) за определенный период времени

2. Компания использует четыре источника поставки деталей одного вида для производства устройств N и M – фирмы A, B, C и D. Фирма A поставляет 40% деталей, B – 30%; C – 20%; D – 10%. Статистика использования поставок показывает, что среди деталей компании A - 3% бракованных деталей; у компании B – 5%; у компании C – 7%; у компании D – 8%. В каждом устройстве N используется 5 деталей. В каждом устройстве M – 3 детали. При использовании хотя бы одной бракованной детали устройства оказываются неработоспособными. В месяц на предприятии изготавливается 100 устройств N и 200 устройств M.

Какова вероятность того, что изготовленное устройство окажется неработоспособным? Каков процент выпуска годных устройств на предприятии?

3. В 1963 году Б. Мандельброт, исследуя цены на хлопок, предположил, что распределение доходности согласуется с устойчивым распределением Леви. Оказалось, что усечённые блуждания Леви очевидным образом наблюдаются в экономике для флуктуаций различных величин на фондовом рынке: ценных бумаг, биржевых фондовых индексов, которые очень хорошо интерпретируются как случайные процессы с независимыми приращениями. Наиболее простой вид плотности двустороннего распределения Леви:

$$p(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\cos(tx)\exp\left(-|t|^{\alpha}\right)dt\text{, где } 0<\alpha\le 2.$$

При каком целочисленном значении α данное распределение соответствует распределению Гаусса?

4. Обозначим: d – некий символ, например, цифра или буква латинского алфавита; d в степени 1(обозначим d ^ 1) - это другое обозначение символа d.
Если n - натуральное число и n больше 1, то d в степени n (обозначим d ^ n) - это последовательность, состоящая из n вхождений символа d.
Таким образом, b^3 g^2 c^4 – это выражение (или, другими словами, цепочка) *bbbggcccc.*
Выражения вида x → y, где x - это символ, y – символ или цепочка символов, будем понимать как подстановки, когда левая часть x заменяется на правую часть y.

Пусть задана следующая система подстановок:
S → A C, (1)
A → p B, (2)
C → w D, (3)
B → p A, (4)
E → w, (5)
D → w E, (6)
B → p , (7)
E → w C (8).

Требуется, используя введенные выше обозначения и теоретико-множественные обозначения, описать множество L, состоящее из ВСЕХ цепочек в алфавите {p, w}, которые можно получить подстановками из указанной системы, если всегда первой применяется подстановка (1), т.е. подстановка с левой частью S.

5. В 1885 году немецкий ученый Герман Эббингауз опубликовал результаты экспериментов, которые легли в основу теории обучения. Испытуемым предъявлялось на некоторое время множество бессмысленных триграмм (трехбуквенных сочетаний). Спустя определенное количество времени проверялось, насколько точно запоминалось это множество. Оказалось, что относительное количество сохраненной в памяти информации (I) зависело от времени (t) как

$$I\approx\exp(-t/s)\text{, где } s \text{ – сила памяти.}$$

Получите зависимость вероятности ошибки запоминания от числа предъявлений множества триграмм.

Программа и методические материалы для подготовки к олимпиаде (магистерская программа “Big Data Systems” (Системы больших данных)

Теория информационных технологий и систем

Понятие информации, ее основные свойства и особенности. Понятие сообщения и его формы, знаки, алфавиты, понятие формального языка. Информация и данные. Конечный вероятностный источник сообщений. Энтропия источника.

Представление информации. Кодирование сообщений источника и текстов. Равномерное и неравномерное кодирование. Дерево кода. Однозначное декодирование, префиксные коды. Условия существования префиксного кода с заданными длинами слов, теорема Крафта. Методы построения префиксных кодов. Код Фано. Средняя длина кодового слова. Нижняя граница средней длины кодового слова. Оптимальное кодирование, свойства оптимальных кодов, построение оптимального кода методом Хафмана. Сжатие данных. Надежность передачи сообщений, способы повышения надежности. Принципы использования кодов, обнаруживающих и исправляющих

ошибки. Расстояние Хемминга. Связь минимального расстояния кода с его характеристиками. Корректирующие возможности кодов, границы Хэмминга и Варшамова-Гилберта. Понятие линейного группового кода.

Обработка информации. Понятие алгоритма и его свойства. Способы формальной записи алгоритмов. Распределенная обработка информации и проблемы взаимодействия параллельно выполняемых процессов обработки. Методы описания и анализа процессов распределенной обработки. Сети Петри.

Защита информации при передаче, основные угрозы и методы защиты от них. Симметричная, асимметричная и комбинированная криптосистемы. Электронная цифровая подпись и принципы ее использования

Анализ структурированной и неструктурированной информации

Основная литература

1 Новиков Ф.А. Дискретная математика для программистов – СПб.: Питер 2006. – 304 с.: ил.

2 Леонтьев В.К. Теория кодирования. М.: Знание, 1977.

3 Аршинов М.Н., Садовский Л.Е. Коды и математика М.: Наука, 1983.

Дополнительная литература

1 Гаврилов Г.П., Сапоженко А.А. Задачи и упражнения по курсу дискретной математики. М.: ФИЗМАТЛИТ, 2005.

2 Морелос-Сарагоса М. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение. М.: Техносфера, 2005.

3 Вернер М. Основы кодирования. М.: Техносфера, 2004.

4 Хэмминг Р.В. Теория кодирования и теория информации /Пер. с англ. – М.: Радио и связь, 1983.

5 Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А. Теория кодов, исправляющих ошибки: Пер. с англ. –М.: Связь, 1979.

6 Брандт З. Анализ данных. Статистические и вычислительные методы для научных работников и инженеров. М.: Мир, 2003.

7 Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989.

8 Айвазян С. А., Бухштабер В. М., Юнюков И. С., Мешалкин Л. Д. Прикладная статистика: Классификация и снижение размерности. М.: Финансы и статистика, 1989.

9 Хайкин С. Нейронные сети: полный курс, 2-е издание.: Пер. с англ. М.: Издательский дом «Вильямс», 2006.

10 Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование. М.: Финансы и статистика, 2010.

11 Малинецкий Г.Г. Математические основы синергетики: хаос, структуры, вычислительный эксперимент. М.: Издательство ЛКИ, 2007.

Базы данных

Виды моделей данных. Структурированные модели данных: сетевые, иерархические, реляционные. Особенности, преимущества и недостатки. Влияние типа модели на особенности проектирования. Понятие предметной области. Состав инфологической модели (ИЛМ). Требования, предъявляемые к ИЛМ. Способы описания предметной области. ER-модели. Объекты и классы объектов. Атрибуты объектов. Типы объектов. Виды связей. Классы членства.

Проектирование баз данных на основе использование ER-моделей. Алгоритм проектирования (алгоритм перехода от базовой ER-модели). Проектирование с использованием CASE-систем. Целостность баз данных. Классификация ограничений целостности. Причины, вызывающие нарушение ограничений целостности. Процедурный и декларативный способы задания ограничений целостности.

Основная литература

1. Диго С.М. Базы данных. Проектирование и использование. Учебник - М.:

2. «Финансы и статистика», 2005.- 592 с.
3. Мишенин А.И. Теория экономических информационных систем. - М.:
4. «Финансы и статистика», 2008.

Дополнительная литература

1. Диго С. М. Access: учебно-практич пособие. - М.: ТК Велби, изд-во Проспект, 2006
2. Диго С. М. Базы данных. Проектирование и создание.Учебно-методический комплекс. М.:Изд. Центр УАОИ, 2008
3. Когаловский М. Р. Энциклопедия технологий баз данных. – М.: «Финансы и статистика», 2008.
4. Кошелев В.Е.. Access 2007- М.:Бином, 2008
5. Каленик А. И. Использование новых возможностей Microsoft SQL Server 2005, Русская редакция; Питер, 2006.
6. Маклаков С. В. Создание информационных систем с ALLFusion Modeling Suite. М.:ДИАЛОГ-МИФИ, 2005 – 432 с.
7. Харрис Энди. PHP/MySQL для начинающих. Пер. с англ., Кудиц-образ, 2005.
8. Грабер М. SQL. – М.: ЛОРИ, 2001.
9. Дейт К.Дж. Введение в системы баз данных, 6-е изд.: Пер. с англ. – СПб.: Издательский дом «Вильямс», 2000.
10. Грабер М. SQL. Справочное руководство. – М.: ЛОРИ, 1997.
11. 11Грабер М. Введение в SQL.: Пер. с англ. – М.: ЛОРИ, 1996. .

Моделирование и оптимизация бизнес-процессов

Методологии описания деятельности. Понятие о моделировании деятельности. Моделирование деятельности и моделирование процессов. Предметные области в деятельности организации. Уровни описания. Общие принципы моделирования деятельности. Эволюция развития методологий описания. Моделирование организации, модели организационной структуры, моделирование и анализ бизнес-процессов, процессные модели проекта. Методологии и стандарты SADT,

IDEF, DFD. Методология ARIS. Методология UML. Сравнительный анализ методологий моделирования. Инструментальные системы для моделирования бизнеса. Требования к инструментальным системам для моделирования бизнеса. Инструментальная система ARIS. Сравнительный анализ инструментальных средств.

Основная литература

1 Хаммер М., Чампи Дж. Реинжиниринг корпорации: Манифест революции в бизнесе. – М.: Изд-во «Манн, Иванов и Фербер», 2011.
2 Андерсен Б. Бизнес-процессы. Инструменты совершенствования. – М.: РИА Стандарты и качество, 2007.
3 Каменнова М.С., Громов А.И., Ферапонтов М.М., Шматалюк А.Е. Моделирование бизнеса. Методология ARIS. – М.: Весть-МетаТехнология, 2001.
4 Шеер А.-Б. Бизнес-процессы. Основные понятия. Теория. Методы. – М.: Весть-МетаТехнология, 1999.

Дополнительная литература

1 А.И.Громов, В.Г.Чеботарев. Применение системного подхода к идентификации процессов организации // Информационные технологии в проектировании и производстве. №3. –М. Изд-во ФГУП «ВИМИ», 2008. – С. 18-22.
2 В.Г. Чеботарев, А.И. Громов. Эволюция подходов к управлению бизнес-процессами// Бизнес-информатика, №1. -М.: Изд-во ГУ-ВШЭ, 2010
3 Марка Д.А., Мак Гоуэн К. Методология структурного системного анализа и проектирования SADT. – М.: МетаТехнология, 1993.
4 Оценка и аттестация зрелости процессов создания и сопровождения программных средств и информационных систем (ИСО/МЭК ТО 15504) – М.: Книга и бизнес,

2001.
5 Process Classification Framework. Developed By APQC's International Benchmarking Clearinghouse In Partnership With Arthur Andersen & Co., SC. 1996
6 Process Classification Framework. APQC, 2004

Проектирование информационных систем

Основные классы информационных систем. Архитектура информационной системы. Основные этапы создания информационных систем. Регламентация процессов проектирования, состава и содержания проектной документации в отечественных (ГОСТ 34.601-603) и международных (ISO/IEC 12207) стандартах.

Основы теории систем. Система, элемент, связь. Информационная система как средство реализации информационных технологий. Функции и ресурсы информационных систем. Структура и принципы функционирования информационных систем. Основные типы информационных систем. ERP, CRM, OLAP и OLTP, базы данных, архивы.

Основная литература

1 Грекул В.И., Коровкина Н.Л., Куприянов Ю.В. Проектированиеинформационных систем. Практикум Москва: Национальный открытый университет «ИНТУИТ», 2012. 186 с.
2 Грекул В.И., Денищенко Г.Н., Коровкина Н.Л. Проектирование ИС. Учебное пособие. Интернет-университет, М., 2008
3 Г.Н. Смирнова, А.А.Сорокин, Ю.Ф. Тельнов Проектирование экономических информационных систем. Учебник. М., «Финансы и статистика»,2002
4 Вендров А.М. Проектирование программного обеспечения экономических информационных систем. М., «Финансы и статистика», 2000

Дополнительная литература

1 Маклаков С.В. Создание ИС с AllFusion Modelling Suite. М., «Диалог- МИФИ», 2003

Теория вероятностей и математическая статистика

Основные понятия. Независимость. Функции распределения. Числовые характеристики случайных величин. Условные вероятности. Биномиальное и нормальное распределения вероятностей. Распределение Пуассона и другие распространенные распределения вероятностей. Применение вероятностного аппарата к статистической информации. Исследование выборками. Статистические оценки. Их свойства и методы построения

Доверительные интервалы. Проверка гипотез. Дисперсионный анализ. Примеры многомерных распределений, многомерное нормальное распределение. Моменты, ковариационная матрица. Информация Фишера. Свойства оценок максимального правдоподобия. Порядковые статистики.

Основная литература

1. Шведов А.С. Теория вероятностей и математическая статистика. 2-е изд. М.: Изд-во ГУ-ВШЭ, 2005.
2. Шведов А.С. Теория вероятностей и математическая статистика – 2 (промежуточный уровень). М.: Изд-во ГУ-ВШЭ, 2007.
3. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики. – С.-П.: Лань, 1999.
4. Пугачев В.С. Теория вероятностей и математическая статистика. М.: Физматлит, 2002.

5. Hogg R.V., Craig A.T. (1995) Introduction to Mathematical Statistics. 5th ed., Upper Saddle River, Prentice Hall.
6. Newbold P. (1995) Statistics for Business and Economics. 4th ed. London: Prentice-Hall.

Дополнительная литература

1. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической
2. Сборник задач по теории вероятностей и математической статистике (сост. Ратникова Т.А., Шведов А.С.). М.: ВШЭ, 2004.

Системный анализ

Место и роль системного подхода и анализа в бизнесе. Виды систем. Свойства систем. Принципы системного анализа. Сложные системы. Признаки сложной системы. Классификация систем. Самоорганизующиеся системы. Принципы синергетического подхода. Закон простоты сложных систем. Закон конечности скорости распространения взаимодействия. Теорема Геделя о неполноте. Закон эквивалентности вариантов построения сложных систем. Закон Онсагера максимизации убывания энтропии.

Математические технологии системного анализа: математическое программирование (линейное программирование; целочисленное программирование; нелинейное программирование и др.); теория игр; теория оптимального управления; имитационное моделирование систем и процессов; теория автоматов; системы искусственного интеллекта; многокритериальный анализ.

Основная литература

1. Системный анализ и принятие решений. М.: Высшая школа, 2004. – 616 с.

Дополнительная литература

1. Афанасьев М.Ю. Исследование операций в экономике: модели, задачи, решения / М.Ю. Афанасьев, Б.П. Суворов. – М.: Инфра-М, 2003.
2. Карманов В.Г. Моделирование в исследовании операций / В.Г.Карманов, В.В.Федоров. – М. : Твема, 1996.

Методика проверки и оценки заданий олимпиады для студентов и выпускников вузов

профиль “Big Data Systems” (Системы больших данных)

1. Проверяется только чистовик.

2. Установлена 100-балльная шкала оценки работы.

3. Максимальные баллы по пунктам задания:

3.1. Дайте развернутые ответы
A_1 – 15
A_2 – 15
3.2. Решите задачи
B_1 – 15
B_2 – 10
B_3 – 15
B_4 – 15

B_5 – 15

3. Дифференцирование оценок. Дифференцированные оценки возможны только в заданиях типа 1 (развернутые ответы). Для заданий типа 2 (решение задач) дифференцированные оценки в основном не предусмотрены: выставляется либо максимальная оценка при правильном решении и ответе, либо ноль баллов. Дифференцирование возможно только в задачах, состоящих из нескольких частей, при которых каждая часть может быть оценена независимо, либо в заданиях, предполагающих наличие вариантов правильных решений.