

Олимпиада для студентов и выпускников – 2016 г.
Демонстрационный вариант и методические рекомендации
по направлению «Бизнес-информатика»

Профиль: «Системы больших данных / Big Data Systems»

Время выполнения задания – 180 мин.,
язык – русский и английский.

1. Дайте развернутые ответы (на английском языке):

- 1.1. Каковы возможности и преимущества использования контекстно-зависимых систем?
1.2. Опишите характеристики современной системы поддержки принятия решений в области управления предприятием

2. Решите задачи:

2.1. Предприятие закупает для своей деятельности товары, сведения о которых хранятся в реляционной базе данных. Для каждого товара в базе данных предприятия хранится информация, содержащая код товара, наименование товара, необходимое максимальное количество единиц товара на складе предприятия, текущее количество единиц товара на складе предприятия. По каждому товару хранится история закупок: дата закупки, количество закупленных единиц товара, код поставщика товара.

В базе данных хранятся также данные о поставщиках товаров: код поставщика; его наименование; контактная информация; перечень поставляемых товаров; цена за единицу каждого товара; максимальное количество единиц товара по каждому товару, которое может быть закуплено одновременно.

Товар может закупаться у любого из поставщиков, который способен обеспечить необходимое количество товара.

Закупка товара производится, когда текущее количество единиц товара на складе предприятия составляет 20% от максимального количества единиц данного товара на складе предприятия. Товар закупается в количестве, определяемом максимальным количеством единиц данного товара на складе предприятия (предполагается, что за время оформления и доставки закупленного товара имеющиеся на складе остатки будут израсходованы).

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. написать запросы для получения следующей информации с использованием языка MySQL:
 - а) коды и наименования товаров, которые должны быть закуплены в текущий момент, количество закупаемых единиц товара по каждому товару, перечень предполагаемых поставщиков, включая цену закупки по каждому поставщику.

Олимпиада для студентов и выпускников – 2016 г.

б) по товару с кодом N вывести перечень поставщиков за определенный период времени $[t_1, t_2]$, расположив поставщиков по убыванию количества закупленных единиц товара.

2.2. Для производства определенного изделия компания использует два вида деталей – деталь А и деталь В. Работнику компании требуется распределить по двум ящикам 4 детали, 2 вида А и 2 вида В.

Каким образом работнику следует разместить детали в ящиках, чтобы обеспечить максимальную вероятность извлечения из ящика детали В?

Какова эта вероятность?

Считать, что работник выбирает наугад один из ящиков и вытаскивает из него одну деталь.

2.3. В 1996 году Р. Мантенья и Х. Стенли установили, что распределение плотности вероятностей изменений индекса S&P500 ($Z_{\Delta t} = SP500(t + \Delta t) - SP500(t)$) на временном интервале Δt подчиняется усеченному распределению Леви:

$$p_L(\beta) = \frac{2}{\pi} \int_0^{\infty} \cos(\beta x) \exp(-\gamma \Delta t x^\alpha) dx$$

где $\beta = \frac{Z_{\Delta t}}{0,0508}$, $\gamma = 0,00375$, $\alpha = 1,4$.

Для $\Delta t = 32$ найти плотности распределения вероятностей при $\alpha = 1$ и $\alpha = 2$.

Какие это распределения?

2.4. Обозначим: d – произвольный 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, $b^3 d^2 c^4$ – это выражение (или, другими словами, цепочка) $bbbddcccc$. Выражения вида $x \rightarrow y$, 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:

$$E \rightarrow ccN \quad (1)$$

$$S \rightarrow cYdd, \quad (2)$$

$$H \rightarrow \text{пусто}, \quad (3)$$

$$E \rightarrow bbEddd, \quad (4)$$

$$Y \rightarrow \text{пусто}, \quad (5)$$

$$H \rightarrow ccN, \quad (6)$$

$$S \rightarrow bbEddd, \quad (7)$$

$$Y \rightarrow cYdd \quad (8).$$

Требуется, используя введенные выше обозначения и теоретико-множественные обозначения, описать множество L , состоящее из ВСЕХ цепочек в алфавите $\{b, c, d\}$, которые можно получить подстановками (1)-(8) из указанной системы, если всегда первой применяется какая-либо подстановка с левой частью S .

Национальный исследовательский университет «Высшая школа экономики»

Олимпиада для студентов и выпускников – 2016 г.

2.5. 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 в подавляющем большинстве случаев вероятность ошибки зависела от числа предъявлений множества (N) как

$$P = C \exp(-\lambda N)$$

где $C, \lambda = const$.

Получите зависимость относительного количества сохраненной в памяти информации от числа предъявлений множества триграмм.

Методические рекомендации

1. Предварительные критерии оценивания

Результаты олимпиадного задания оцениваются по 100-балльной системе. Проверяется только чистовик.

Максимальные баллы по пунктам задания:

1. Дайте развернутые ответы

1.1 – 15 баллов

1.2 – 15 баллов

2. Решите задачи

2.1.– 15 баллов

1.2.– 10 баллов

1.3.– 15 баллов

1.4.– 15 баллов

1.5.– 15 баллов

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 (задача 2.1).

4. Ответы на вопросы 1.1 и 1.2 должны быть даны на английском языке.

2. Перечень и содержание тем олимпиадных состязаний. Список рекомендуемой литературы

Основы 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. Обработка информации. Понятие алгоритма и его свойства. Способы формальной записи алгоритмов. Распределенная обработка информации и проблемы взаимодействия параллельно выполняемых процессов обработки.

Современные методы анализа структурированной и неструктурированной информации. Обработка текстов на естественном языке.

Современные методы хранения и обработки информации.

Основные классы информационных систем. Архитектура информационной системы. Основные элементы информационных систем. Структура и принципы функционирования информационных систем.

Основная литература

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

Олимпиада для студентов и выпускников – 2016 г.

2. Джордж Риз. Облачные вычисления. Изд-во: БХВ-Петербург, 2011.
3. Брандт З. Анализ данных. Статистические и вычислительные методы для научных работников и инженеров. М.: Мир, 2003.
4. В.Грекул, Г. Денищенко, Н.Коровкина. Проектирование информационных систем М.: Изд-во: Бинوم. Лаборатория знаний. Сер.: Основы информационных технологий. 2008.

Дополнительная литература

1. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989.
2. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование. М.: Финансы и статистика, 2010.
3. Е. Горяинова, А.Панков, Е.Платонов. Прикладные методы анализа статистических данных. М.: Изд-во: Высшая Школа Экономики, 2012.
4. В.А.Фомичев. Математические основы представления содержания посланий компьютерных интеллектуальных агентов. М.: Изд-во: Высшая Школа Экономики, 2007.
5. А. Е. Пентус, М. Р. Пентус. Математическая теория формальных языков. Изд-во: Бинوم. Лаборатория знаний, Интернет-университет информационных технологий. Серия: Основы информатики и математики. 2006

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,

Инструментальные системы для моделирования бизнеса.

Основная литература

1. Хаммер М., Чампи Дж. Реинжиниринг корпорации: Манифест революции в бизнесе. – М.: Изд-во «Манн, Иванов и Фербер», 2011.
2. Шеер А.-Б. Бизнес-процессы. Основные понятия. Теория. Методы. – М.: Весть-МетаТехнология, 1999.

Дополнительная литература

1. Марка Д.А., Мак Гоуэн К. Методология структурного системного анализа и проектирования SADT. – М.: МетаТехнология, 1993.
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.

Базы данных

Виды моделей данных. Структурированные модели данных: сетевые, иерархические, реляционные. Особенности, преимущества и недостатки. Влияние типа модели на особенности проектирования. Понятие предметной области. 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

Олимпиада для студентов и выпускников – 2016 г.

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

Дополнительная литература

1. Диго С. М. Access: учебно-практич пособие. - М.: ТК Велби, изд-во Проспект, 2006
2. Коголовский М. Р. Энциклопедия технологий баз данных. – М.: «Финансы и статистика», 2008.
3. Харрис Энди. PHP/MySQL для начинающих. Пер. с англ., Кудиц-образ, 2005.
4. Грабер М. SQL. – М.: ЛОРИ, 2001.
5. Дейт К.Дж. Введение в системы баз данных, 6-е изд.: Пер. с англ. – СПб.: Издательский дом «Вильямс», 2000.
6. Виктор Гольцман. MySQL 5.0. Изд-во: Питер, 2009

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. Устойчивые распределения. 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. Мантенья Р., Стенли Х. Введение в эконофизику. Корреляции и сложность в финансах. Пер. с англ., М.: УРСС, 2007.
4. Пугачев В.С. Теория вероятностей и математическая статистика. М.: Физматлит, 2002.
5. Borak S., Hardle W., Weron R. Statistical Tools for Finance and Insurance. Stable Distributions. Springer, 2005.

Дополнительная литература

1. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической
2. Романовский М.Ю., Романовский Ю.М. Введение в эконофизику. Статистические и динамические модели. М.: 2007.

Национальный исследовательский университет «Высшая школа экономики»