

# **Лабораторная работа**

## **Задание к лабораторной работе**

1. Изучить теоретический материал.
2. Загрузить MS Word. Результаты работы студента должны быть сохранены.
3. Выбрать определенную предметную область, согласовать с преподавателем.
4. Создать иерархическую систему классификации для выбранных объектов.
5. Построить фасетную систему классификации для этих же объектов.
6. Составить словарь дескрипторов: 10 - 15 ключевых слов. Указать между выбранными из словаря словами любую связь: синонимическую, родо-видовую, ассоциативную. Если получится, то представьте все три типа отношений.
7. Результаты работы представить преподавателю.

## **Теоретическая часть Основные понятия**

**Информатика** – научная дисциплина, изучающая вопросы, связанные с поиском, сбором, хранением, преобразованием и использованием информации в самых различных сферах человеческой деятельности.

Генетически информатика связана с вычислительной техникой, компьютерными системами и сетями, так как именно компьютеры позволяют порождать, хранить и автоматически перерабатывать информацию в таких количествах, что научный подход к информационным процессам становится одновременно необходимым и возможным.

Главная функция информатики заключается в разработке методов и средств преобразования информации и их использовании в организации технологического процесса переработки информации.

Информатика служит для формирования мировоззрения в информационной сфере и определенного уровня информационной культуры, то есть умения целенаправленно работать с информацией, профессионально используя для ее получения, обработки и передачи компьютерную информационную технологию и соответствующие ей технические и программные средства.

Предмет информатики определяется многообразием ее приложений. Различные информационные технологии, функционирующие в разных видах человеческой деятельности (управление производственным процессом, системы проектирования, финансовые операции, образование и т.п.), имея общие черты, в то же время существенно различаются между собой.

Тем самым образуются различные «предметные» информатики, базирующиеся на разных наборах операций и процедур, различных видах оборудования (во многих случаях наряду с

компьютером используются специализированные приборы и устройства), разных информационных носителях и т.п.

Можно выделить следующие аспекты информатики:

информатика как область интеграции знаний о роли информации в развитии общества, об управлении знаниями, о логических основах построения компьютера и обработки информации, о методах и средствах моделирования информационных объектов, об основах построения баз данных и информационных системах, об информационной безопасности;

техническая база информатики, где формируется представление об аппаратной части компьютера и о формах представления в нем данных, об основах построения компьютерных сетей, о состоянии и тенденции развития компьютерных систем;

алгоритмическое и программное обеспечение информатики, содержащее основы теории алгоритмов, технологии и инструменты программирования, классификацию программного обеспечения, характеристики и особенности разных классов программных продуктов.

Информатика рассматривает информацию как концептуально связанные между собой сведения, понятия, изменяющие наши представления о явлении или объекте окружающего мира. Наряду с информацией в информатике часто употребляется понятие данных

Данными называют информацию, когда делают акцент на том, что она поступает в качестве входных параметров некоторого алгоритма обработки. Данные - от слова "дано", так же как в записи условия задачи мы записываем "дано" и "найти".

Этот термин часто употребляется, когда речь идет о компьютерных системах. Одни и те же данные, записанные в компьютерной памяти, могут быть отображены в зависимости от выбранной программы обработки или как графическая информация, или как символьная, или как звуковая, или как числовая.

Данные - это любые зарегистрированные сигналы. Слово "информация" чаще используется в более узком смысле, чем просто все, что отражается в материальном объекте в результате воздействия на него другого материального объекта. Предполагается, что получение информации дает получившему ее возможность принимать решения, действовать, осуществлять выбор или пополнить свою систему знаний.

Если полученные данные не приводят ни к чему из перечисленного, то с субъективной точки зрения считается, что для получателя они информации не несут, хотя и занимают определенный объем его памяти.

Говоря об информации и ее свойствах, обычно имеют в виду один из трех аспектов:

Технический - точность, надежность, скорость передачи сигналов, объем, занимаемый в памяти зарегистрированными сигналами, способы регистрации сигналов. В этом аспекте информация = данные, и никак не учитывается ее полезность для получателя или ее смысловое содержание. Это информация в самом широком общем для всей материи смысле, информация в аспекте восприятия, хранения, передачи.

*Прагматический* - насколько эффективно информация влияет на поведение получателя. В этом аспекте говорят о полезности и ценности информации. В определенных случаях ценность информации становится отрицательной, а сама информация становится дезинформацией. Это информация в аспекте управления поведением.

*Семантический* - передача смысла с помощью кодов. Семантической называется информация, активизирующая образы, уже имеющиеся в тезаурусе получателя (узнаваемая) или вносящая изменения в его тезаурус (систему знаний). Это информация в аспекте знаний.

## **Классификация информации**

Множество объектов легко поддается изучению, если оно систематизировано.

**Классификация** - упорядочение объектов предметной области по выбранным признакам в соответствии с существующей системой классификации.

Под системой классификации понимается совокупность методов и правил, в результате которых заданное множество разбивается на подмножества, то есть классификационные группировки.

Свойства или характеристики объектов, по которым осуществляется классификация, называются признаками или основаниями классификации.

При классификации объектов предметной области необходимо соблюдать ряд правил:

- классификация должна осуществляться только по одному основанию, а не по нескольким;
- полученные в результате деления подклассы должны исключать друг друга;
- деление на подклассы должно быть непрерывным (то есть с последующим уточнением характеристик).

Разработаны три метода классификации объектов:

- иерархический,
- фасетный,
- дескрипторный.

Эти методы различаются разной стратегией применения классификационных признаков.

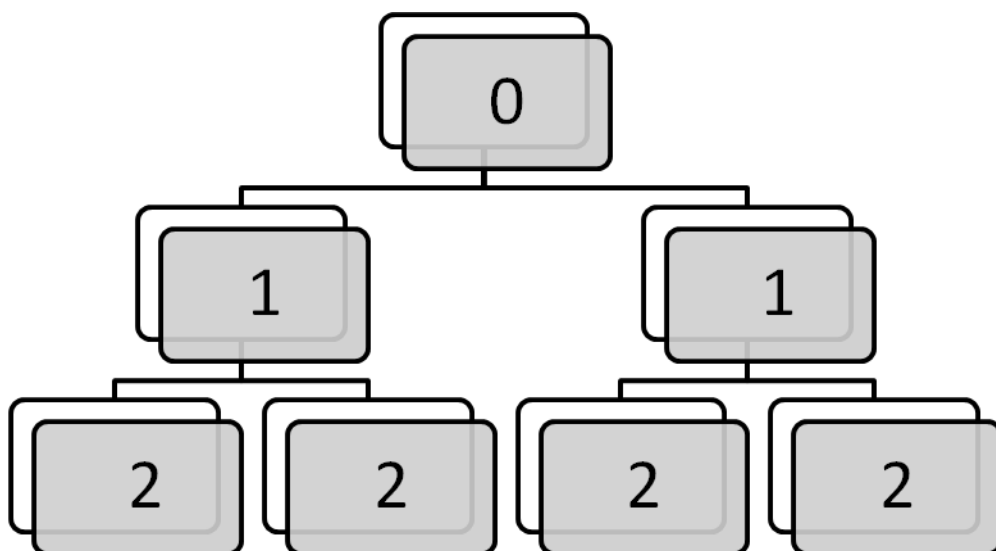
**Иерархическая система классификации** обеспечивает последовательное разделение множества объектов на соподчиненные группировки. Исходное множество объектов классификации делится сначала по какому-либо признаку на крупные группировки, каждая из которых, в свою очередь дробится на ряд последующих группировок, конкретизируя свойства объектов в виде иерархического логического дерева

Преимуществом иерархической классификации является то, что она не требует высокой квалификации исполнителей при ее организации. Детализация может быть любой глубины, что зависит от конкретных целей.

К недостаткам иерархической классификации относится то, что введение дополнительного признака требует существенной переработки всей структуры классификационного множества.

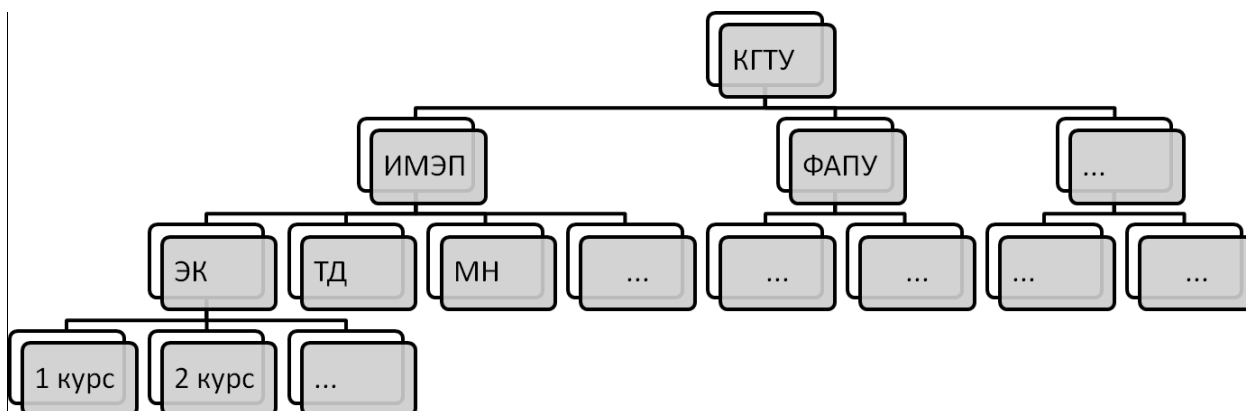
Иерархическая система классификации строится следующим образом:

- исходное множество элементов составляет 0-й уровень и делится в зависимости от выбранного классификационного признака на классы (группировки), которые образуют 1-й уровень;
- каждый класс 1-го уровня в соответствии со своим, характерным для него классификационным признаком делится на подклассы, которые образуют 2-й уровень;
- каждый класс 2-го уровня аналогично делится на группы, которые образуют 3-й уровень и т.д.



Учитывая достаточно жесткую процедуру построения структуры классификации, необходимо перед началом работы определить ее цель, т.е. какими свойствами должны обладать объединяемые в классы объекты. Эти свойства принимаются в дальнейшем за признаки классификации.

В качестве примера рассмотрим классификацию информации о студентах.



Классификация информации о студентах (иерархический метод). Квалификационные признаки: факультет, направление, курс, ...

**Система фасетной классификации** основана на параллельном разделении множества объектов на независимые классификационные группировки, которые затем формируются в параллельные, независимые фасеты. Таким образом, фасет - это признак, принятый для изучения

При фасетном методе классификации информация об объектах предметной области может детализоваться, то есть не иметь жесткой структуры. Это обеспечивает раскрытие содержания выбранных характеристик до необходимой степени детализации. Важным преимуществом фасетной системы классификации перед иерархической является то, что она имеет более гибкую структуру: изменение содержания фасетов или введение дополнительных не требует пересмотра всей структуры.

При построении фасетной системы классификации необходимо, чтобы значения, используемые в различных фасетах, не повторялись. Фасетную систему легко можно модифицировать, внося изменения в конкретные значения любого фасета.

В качестве примера опять рассмотрим классификацию информации о студентах.

Факультет	Направление	Курс	Форма обучения	Пол	...
ИМЭП	Менеджмент	1	очная	муж	
ФАПУ	Прикладная информатика	2	заочная	жен	
ФСЭ	Торговое дело	3	очно-заочная		
....	Экономика	4			
	Кораблестроение				
	...				

Фасетная система классификации информации о студентах. Классификационные признаки по фасетам: Факультет, Направление, Курс, Форма обучения, Пол...

Достоинства фасетной системы классификации:

возможность создания большой емкости классификации, т.е. использования большого числа признаков классификации и их значений для создания группировок;

возможность простой модификации всей системы классификации без изменения структуры существующих группировок.

Недостатком фасетной системы классификации является сложность ее построения, так как необходимо учитывать все многообразие классификационных признаков.

Для организации поиска информации, для ведения тезаурусов (словарей) эффективно используется дескрипторная (описательная) система классификации, язык которой приближается к естественному языку описания информационных объектов. Особенно широко она используется в библиотечной системе поиска.

**Дескрипторный метод классификации** заключается в следующем:

отбирается совокупность ключевых слов или словосочетаний, описывающих определенную предметную область или совокупность однородных объектов.

выбранные ключевые слова и словосочетания подвергаются нормализации, т.е. из совокупности синонимов выбирается один или несколько наиболее употребимых;

создается словарь дескрипторов, т.е. словарь ключевых слов и словосочетаний, отобранных в результате процедуры нормализации.

Классифицируем учебную деятельность в высшем учебном заведении. Ключевыми словами могут быть выбраны: студент, обучаемый, учащийся, преподаватель, учитель, педагог, лектор, ассистент, доцент, профессор, коллега, факультет, подразделение университета, аудитория, комната, лекция, практическое занятие, занятие и т.д.

Среди указанных ключевых слов встречаются синонимы, например: студент, обучаемый, учащийся, преподаватель, учитель, педагог, факультет, подразделение университета и т.д.

После нормализации словарь дескрипторов будет состоять из следующих слов: студент, преподаватель, лектор, ассистент, доцент, профессор, факультет, аудитория, лекция, практическое занятие и т.д.

Между дескрипторами устанавливаются связи, которые позволяют расширить область поиска информации. Связи могут быть трех видов:

синонимические указывающие некоторую совокупность ключевых слов как синонимы;

родо-видовые, отражающие включение некоторого класса объектов в более представительный класс;

ассоциативные, соединяющие дескрипторы, обладающие общими свойствами.

Синонимическая связь: студент – учащийся - обучаемый.

Родо-видовая связь: университет – факультет - кафедра.

Ассоциативная связь: студент – экзамен – профессор - аудитория.