


КЛАССИФИКАЦИЯ И КОДИРОВАНИЕ ИНФОРМАЦИИ

1. Система классификации информации.
2. Система кодирования информации.
3. Классификация информации по разным признакам.

1. Система классификации информации (общие сведения)

Классификация — система распределения объектов по классам в соответствии с определенным признаком.

Под **объектом** понимается любой предмет, процесс, явление материального или нематериального свойства.



Классификация объектов — это процедура группировки на качественном уровне, направленная на выделение однородных свойств.

Применительно к информации как к объекту классификации выделенные классы называют **информационными объектами.**

Свойства информационного объекта определяются информационными параметрами, называемыми **реквизитами**. Реквизиты представляются либо числовыми данными, например вес, стоимость, год, либо признаками, например цвет, марка машины, фамилия.

Реквизит — логически неделимый информационный элемент, описывающий определенное свойство объекта, процесса, явления и т.п.

.

-
- Классификация нужна для разработки правил (алгоритмов) и процедур обработки информации, представленной совокупностью реквизитов.
 - В любой стране разработаны и применяются государственные, отраслевые, региональные классификаторы. Например:



Классификаторы

По отраслям промышленности

По оборудованию

По профессиям

По единицам измерения

По статьям затрат

...

○ **Классификатор** — систематизированный свод наименований и кодов классификационных группировок.

При классификации широко используются понятия **классификационный признак** и **значение классификационного признака**, которые позволяют установить сходство или различие объектов.



Системы классификации объектов:

1. Иерархическая

2. Фасетная

3. Дескрипторная

Иерархическая система классификации информации

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.

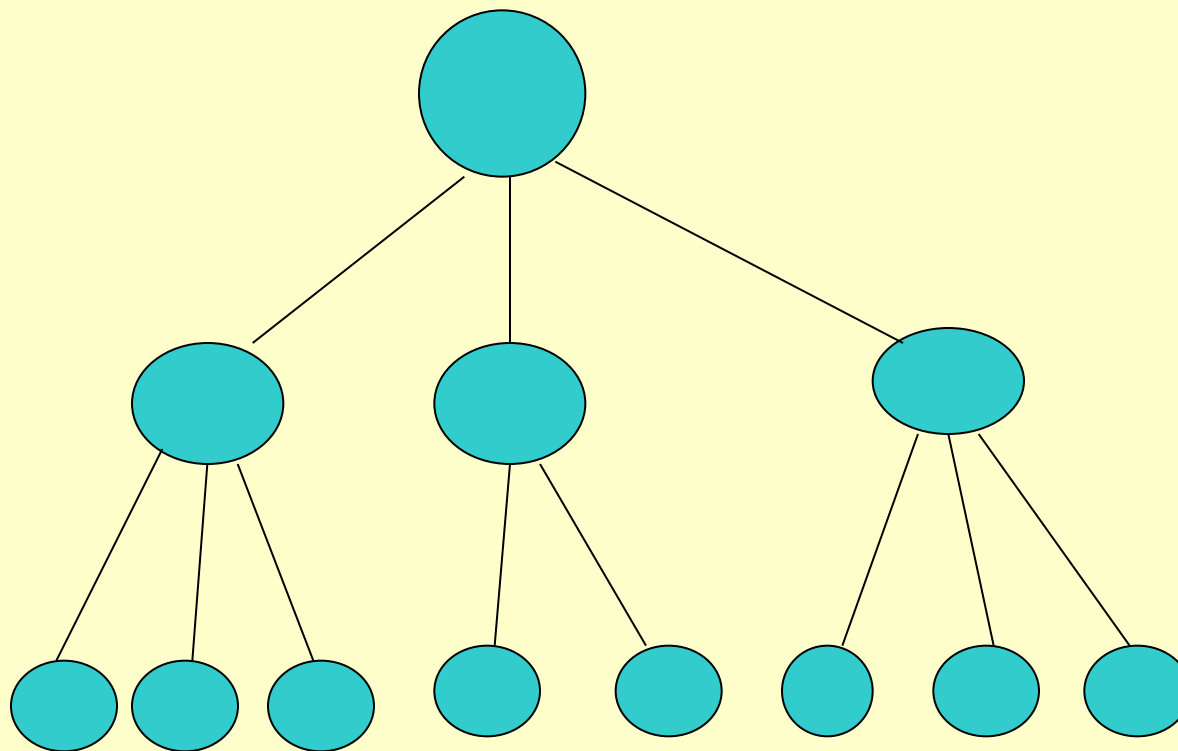
Иерархическая система классификации


0-й уровень

0-й уровень

1-й
уровень

2-й
уровень





В иерархической системе классификации каждый объект на любом уровне должен быть отнесен к одному классу, который характеризуется конкретным значением выбранного классификационного признака. Для последующей группировки в каждом новом классе необходимо задать свои классификационные признаки и их значения.

Фасетная система классификации информации

Фасетная система классификации позволяет выбирать признаки классификации независимо друг от друга, и от семантического содержания классифицируемого объекта.

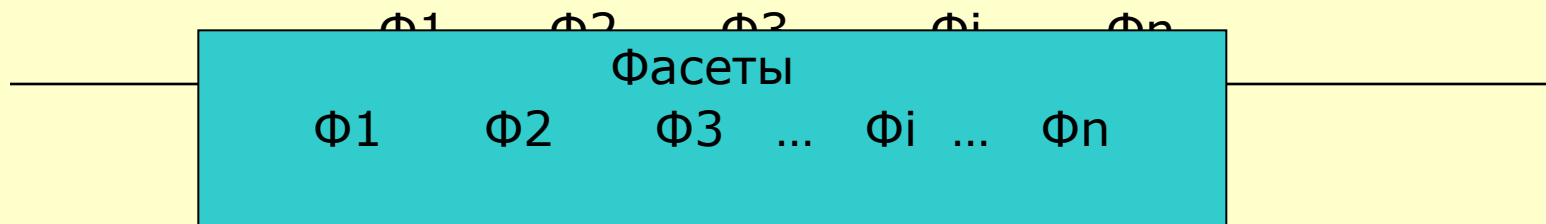
Признаки классификации называются **фасетами** (facet — рамка). Каждый фасет (Ф) содержит совокупность однородных значений данного классификационного признака. Причем значения в фасете могут располагаться в произвольном порядке.

- **Пример.** Фасет *цвет* содержит значения: красный, белый, зеленый, черный, желтый.

Фасет *специальность* содержит названия специальностей.

Фасет *образование* содержит значения: среднее, среднее специальное, высшее.

-
- Схема построения фасетной системы классификации представляется в виде таблицы. Названия столбцов соответствуют выделенным классификационным признакам (фасетам), обозначенным $\Phi 1, \Phi 2, \dots, \Phi i, \dots, \Phi n$. Например, цвет, размер одежды, вес и т.д. Произведена нумерация строк таблицы. В каждой клетке таблицы хранится конкретное значение фасета. Например, фасет *цвет*, обозначенный $\Phi 2$, содержит значения: красный, белый, зеленый, черный, желтый.



Значения
фасетов

1	•	•	•	•	•
2	•	•	•	•	•
...	•		•	•	•
K	•			•	

Процедура классификации состоит в присвоении каждому объекту соответствующих значений из фасетов. При этом могут быть использованы не все фасеты. Для каждого объекта задается конкретная группировка фасетов структурной формулой, в которой отражается их порядок следования:

$$K_s = (\Phi_1, \Phi_2, \dots, \Phi_i, \dots, \Phi_n)$$

где Φ_i - i -й фасет;

n — количество фасетов.


Значения, используемые в различных фасетах, не должны повторяться. Фасетную систему легко можно модифицировать, внося изменения в конкретные значения любого фасета.

Дескрипторная система классификации

- Для организации поиска информации, для ведения тезаурусов (словарей) эффективно используется дескрипторная (описательная) система классификации. Особенно широко она используется в библиотечной системе поиска.

Суть дескрипторного метода классификации :

- отбирается совокупность ключевых слов или словосочетаний, описывающих определенную предметную область или совокупность однородных объектов. Причем среди ключевых слов могут находиться синонимы;
- 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.



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.


2. СИСТЕМА КОДИРОВАНИЯ ИНФОРМАЦИИ

Применяется для замены названия объекта на условное обозначение (код) в целях обеспечения удобной и более эффективной обработки информации.

Система кодирования — совокупность правил кодового обозначения объектов.

Код строится на базе алфавита, состоящего из букв, цифр и других символов. Код характеризуется:

- *длиной* — число позиций в коде;
- *структурой* — порядок расположения в коде символов, используемых для обозначения классификационного признака




Процедура присвоения объекту кодового обозначения называется *кодированием*. Можно выделить две группы:

- *классификационную систему кодирования*, ориентированную на проведение предварительной классификации объектов либо на основе иерархической системы, либо на основе фасетной системы;
- *регистрационную систему кодирования*, не требующую предварительной классификации объектов.

Классификационное кодирование

Классификационное кодирование применяется после проведения классификации объектов. Различают последовательное и параллельное кодирование.

- *Последовательное* кодирование используется для иерархической классификационной структуры. Суть метода заключается в следующем: сначала записывается код старшей группировки 1-го уровня, затем код группировки 2-го уровня, затем код группировки 3-го уровня и т.д. В результате получается кодовая комбинация, каждый разряд которой содержит информацию о специфике выделенной группы на каждом уровне иерархической структуры.



○ *Параллельное* кодирование используется для фасетной системы классификации. Суть метода заключается в следующем:

все фасеты кодируются независимо друг от друга;
для значений каждого фасета выделяется определенное количество разрядов кода.

Регистрационное кодирование

- Регистрационное кодирование используется для однозначной идентификации объектов и не требует предварительной классификации объектов. Различают порядковую и серийно-порядковую систему.
- *Порядковая* система кодирования предполагает последовательную нумерацию объектов числами натурального ряда или по алфавиту.

-
- *Серийно-порядковая* система кодирования предусматривает предварительное выделение групп объектов, которые составляют серию, а затем в каждой серии производится порядковая нумерация объектов. Каждая серия также будет иметь порядковую нумерацию. Серийно-порядковая система является смешанной: классифицирующей и идентифицирующей. Применяется тогда, когда количество групп невелико.

3. КЛАССИФИКАЦИЯ ИНФОРМАЦИИ ПО РАЗНЫМ ПРИЗНАКАМ

Один и тот же объект может быть классифицирован по разным признакам или критериям. Часто встречаются ситуации, когда в зависимости от условий внешней среды объект может быть отнесен к разным классификационным группировкам. Это актуально при классификации видов информации без учета ее предметной ориентации, так как она часто может быть использована для разных целей.

Пример классификации информации:

На рис. будет приведена одна из схем классификации циркулирующей в организации (фирме) информации. В основу классификации положено пять наиболее общих признаков:

1. место возникновения,
2. стадия обработки,
3. способ отображения,
4. стабильность,
5. функция управления

Информация



Место возникновения

- По этому признаку информацию можно разделить на входную, выходную, внутреннюю, внешнюю.
 - *Входная* информация — это информация, поступающая в фирму или ее подразделения
 - *Выходная* информация — это информация, поступающая из фирмы в другую фирму, организацию (подразделение).
 - Одна и та же информация может являться входной для одной фирмы, а для другой, ее вырабатывающей, выходной. По отношению к объекту управления (фирма или ее подразделение: цех, отдел, лаборатория) информация может быть определена как внутренняя, так и внешняя
 - *Внутренняя* информация возникает внутри объекта, *внешняя* информация — за пределами объекта

Стадия обработки.

По стадии обработки информация может быть первичной, вторичной, промежуточной, результатной.

- *Первичная* информация — это информация, которая возникает непосредственно в процессе деятельности объекта и регистрируется на начальной стадии.
- *Вторичная* информация — это информация, которая получается в результате обработки первичной информации и может быть промежуточной и результатной.
- *Промежуточная* информация используется в качестве исходных данных для последующих расчетов.
- *Результатная* информация получается в процессе обработки первичной и промежуточной информации и используется для выработки управленческих решений.

Способы отображения информации.

- По способу отображения информация подразделяется на текстовую и графическую.
- *Текстовая* информация — это совокупность алфавитных, цифровых и специальных символов, с помощью которых представляется информация на физическом носителе (бумага, изображение на экране дисплея).
- *Графическая* информация — это различного рода графики, диаграммы, схемы, рисунки и т.д.

Стабильность.

По стабильности информация может быть переменной (текущей) и постоянной (условно-постоянной).

- *Переменная* информация отражает фактические количественные и качественные характеристики производственно-хозяйственной деятельности фирмы. Она может меняться как по назначению, так и по количеству. Например, количество произведенной продукции за смену, еженедельные затраты на доставку сырья, количество исправных станков и т.п.

-
- *Постоянная* (условно-постоянная) информация — это неизменная и многократно используемая в течение длительного периода времени информация. Постоянная информация может быть справочной, нормативной, плановой:
 - • постоянная справочная информация включает описание постоянных свойств объекта в виде устойчивых длительное время признаков. Например, табельный номер служащего, профессия работника, номер цеха и т.п.;
 - • постоянная нормативная информация содержит местные, отраслевые и общегосударственные нормативы. Например, размер налога на прибыль, стандарт на качество продуктов определенного вида, размер минимальной оплаты труда, тарифная сетка оплаты государственным служащим;
 - • постоянная плановая информация содержит многократно используемые в фирме плановые показатели. Например, план выпуска телевизоров, план подготовки специалистов определенной квалификации.

Функция управления.

- По функциям управления обычно классифицируют *экономическую информацию*. При этом выделяют следующие группы: плановую, нормативно-справочную, учетную и оперативную (текущую).
- *Плановая* информация — информация о параметрах объекта управления на будущий период. На эту информацию идет ориентация всей деятельности фирмы.

Примером плановой информацией фирмы могут быть такие показатели, как план выпуска продукции, планируемая прибыль от реализации, ожидаемый спрос на продукцию и т. д.

Нормативно-справочная информация содержит различные нормативные и справочные данные. Ее обновление происходит достаточно редко.

Примером нормативно-справочной информации на предприятии являются:

- время, предназначенное для изготовления типовой детали (нормы трудоемкости);
- среднедневная оплата рабочего по разряду;
- оклад служащего;
- адрес поставщика или покупателя и т.д.

Учетная информация — это информация, которая характеризует деятельность фирмы за определенный прошлый период времени. На основании этой информации могут быть проведены следующие действия: скорректирована плановая информация, сделан анализ хозяйственной деятельности фирмы, приняты решения по более эффективному управлению работами и пр. На практике в качестве учетной информации может выступать информация бухгалтерского учета, статистическая информация и информация оперативного учета.

- **Оперативная (текущая)** информация — это информация, используемая в оперативном управлении и характеризующая производственные процессы в текущий (данный) период времени. К оперативной информации предъявляются серьезные требования по скорости поступления и обработки, а также по степени ее достоверности.