

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ

ТЕМА ЛЕКЦИИ:

«СТРУКТУРИРОВАНИЕ ЗНАНИЙ»

**Массель Л.В., д.т.н., профессор
кафедры Автоматизированных систем
факультета Кибернетики ИрГТУ**

СТРАТИФИКАЦИОННЫЙ ПОДХОД К СТРУКТУРИРОВАНИЮ ЗНАНИЙ

При структурировании знаний наиболее часто используются свойства иерархичности и статификации, т.е. расслоения. В системологии рассматривается стратификационный подход к построению формальных структур моделей (от «страта», т.е. слой) .

Стратификационные модели представляют собой математические описания иерархии детерминирующих связей, механизмов и структур сложных многокомпонентных систем. Стратификационные модели представляют собой машинно-ориентированные понятия.

Т.е., говоря о стратификации систем, мы, как правило, имеем ввиду конструирование БД или БЗ, над которыми определены вычислительные процессы локально исследованных задач. Недостатком стратификационной модели является то, что при общем рассмотрении понятие стратификационной модели имеет громоздкую абстрактную структуру.

В теории систем БД в последние годы сложилось новое направление исследований, которое получило название *семантического и концептуального моделирования*. При практическом применении получил распространение термин «инфологическое моделирование».

Недостаток инфологических моделей состоит в том, что они, отражая основные объекты предметной области, являются статистическими, т.е. не отражают основных процессов предметной области.

Основными понятиями инфологического моделирования является экземпляр и тип. *Понятие «тип»* нашло отражение в теории программирования. В объектно-ориентированном программировании ему соответствует понятие «класс».

При структурировании знаний основной акцент делается на анализе процедур принятия решений, причем большое внимание уделяется эвристическим знаниям.

ФРАКТАЛЬНЫЙ ПОДХОД К СТРУКТУРИРОВАНИЮ ЗНАНИЙ

Рассмотрим концептуальную модель структурирования знаний, основанную на представлении разных форм или видов знаний как объектов расслоенного (стратифицированного) информационного пространства.

Определим *информационное пространство* как пространство, в которое отображается вся доступная нам информация о жизнедеятельности живых и функционировании технических объектов, т.е. известная нам часть знаний о картине мира.

Под *объектом информационного пространства* будем понимать некоторую совокупность свойств, проявляющуюся в виде реального объекта или процесса.

Сущность объекта – это инвариант, т.е. совокупность существенных и не изменяющихся свойств объекта.

Тогда *информация* – это совокупность описаний сущности и отношений между ними (последние порождают изменяющиеся свойства объектов).

ФРАКТАЛЬНЫЙ ПОДХОД К СТРУКТУРИРОВАНИЮ ЗНАНИЙ

Предлагаемая концептуальная модель названа фрактальной стратифицированной моделью (ФС-моделью). Она отражает такие свойства, присущие организации знаний, как стратификация, иерархичность и фрактальность.

В математике *фракталами* называют геометрические формы, нерегулярная структура которых повторяется в разных масштабах; или, иначе, множество с нецелой размерностью для описания негладких кривых поверхностей. Примеры фракталов: равносторонний треугольник; еловая ветвь; снежинка.

Понятие фрактальности широко используется в топологии, географии, физике, медицине, компьютерной графике.

Один из основоположников теории фракталов Мандельброт определяет фрактал как структуру, состоящую из частей, которая подобна целому (самоподобие).

ФРАКТАЛЬНОСТЬ

Под фрактальностью понимают свойство природы повторяться в разных масштабах и времени (например, модель солнечной системы и планетарная модель атома ядра и электронов – фрактальность в пространстве; дерево вырастает из семени – фрактальность во времени).

В нашем случае фрактальность используется не в математическом, а в философском смысле, т.е. как методологическое свойство, позволяющее одновременное рассмотрение разномасштабных объектов и процессов с сохранением инварианта при изменении масштаба.

При рассмотрении свойств фрактальности выделяют два базовых свойства:

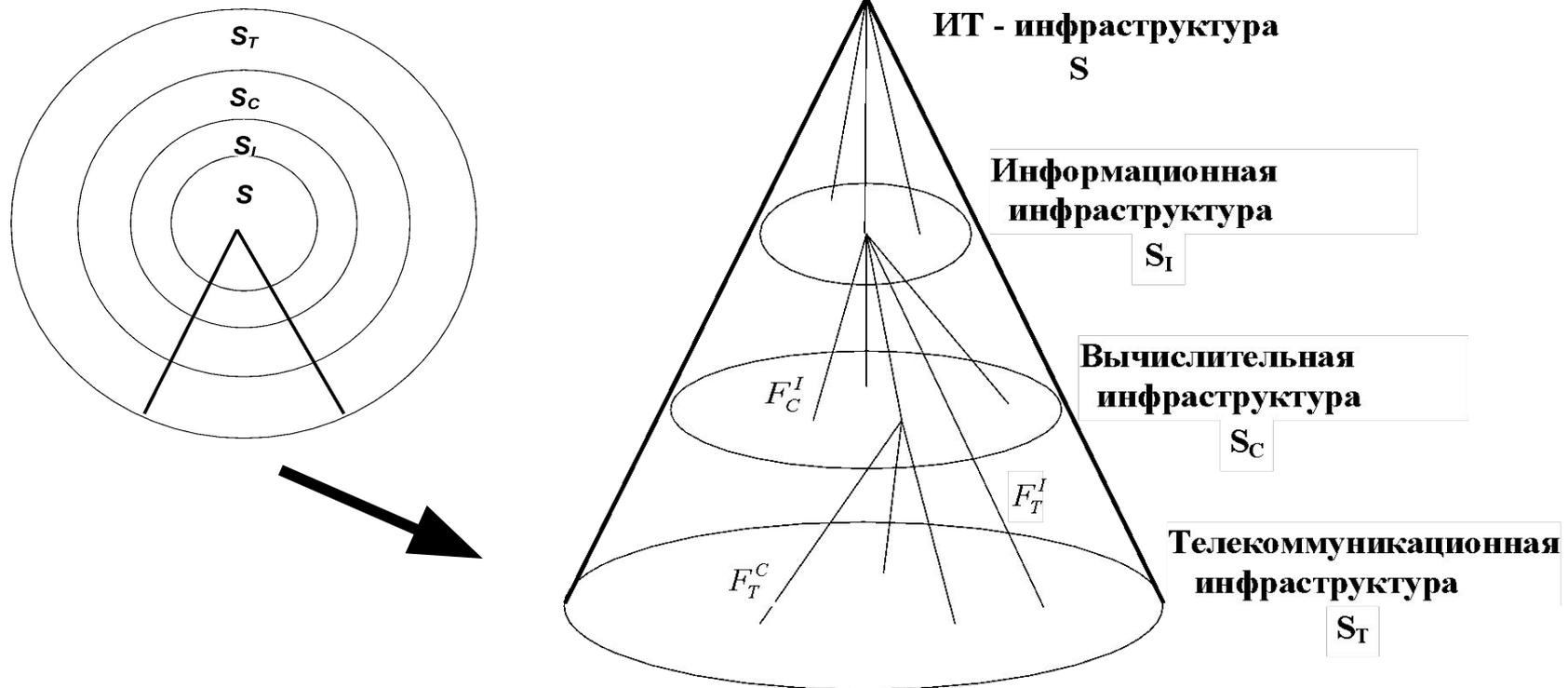
- ❑ самоподобие
- ❑ сохранение инварианта

ФРАКТАЛЬНАЯ СТРАТИФИЦИРОВАННАЯ МОДЕЛЬ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА

- Вся существующая информация о предметной области может быть отображена в информационное пространство, которое можно представить в виде совокупности информационных слоев (миров), состоящих из однотипных информационных объектов.
- Информационный объект представляется в виде совокупности свойств, отражающих реальный объект или процесс
- Знания некоторой предметной области могут быть структурированы с помощью ФС-модели, заданной тройкой (S, F, G) , где
 - S – информационный мир (множество информационных объектов),
 - F – множество отображений,
 - G – множество инвариантов

ФРАКТАЛЬНАЯ СТРАТИФИЦИРОВАННАЯ МОДЕЛЬ ИТ-ИНФРАСТРУКТУРЫ НАУЧНЫХ ИССЛЕДОВАНИЙ

Концепция создания ИТ-инфраструктуры методологически обосновывается с помощью фрактальной стратифицированной модели (ФС-модели) информационного пространства. Графически ФС-модель удобно изображать в виде вложенных сферических оболочек. Каждому уровню знаний соответствует свой слой (или страта) этого пространства и соответственно свой информационный мир. Последовательность отображений может отражать процесс познания.



МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ФС-МОДЕЛИ

Вся существующая информация о предметной области, т.е. знания о реальных процессах и объектах, может быть отображена в информационное пространство $L = \{l_i\}$, состоящих из информационных объектов l_i , $i = 1 \dots n$.

Информационный объект l_i представляется в виде совокупности свойств p_j , отражающих реальный объект или процесс

$$l_i = \{p_j\}, \quad j = 1 \dots n.$$

Знания некоторой предметной области могут быть структурированы с помощью ФС-модели, определяемой тройкой (S, F, G) и представляющей пространство L в виде слоев (или информационных миров) однотипных объектов l_i , если:

- а) информационный мир S может быть расслоен
- б) информационный объект s_i – совокупность элементарных или сложных свойств
- в) заданы разбиения S на слои однотипных объектов (причем эти слои не пересекаются), такие, что s_i , в свою очередь, может быть расслоен.

$$S = \{s_i\}, \quad i = 1 \dots n$$

$$\emptyset, \quad i \neq j$$

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ФС-МОДЕЛИ

- г) задано множество отображений F (из любого слоя в каждый), таких, что сохраняются инварианты, т.е. существенные свойства любого l_i или всех l_i , которые относятся к s_i
- д) задано множество инвариантов G , определяемое тройкой (G_0, G_i, g_j) , $i = 1 \dots n$, $j = 1 \dots m$, где
 G_0 – это инвариант разбиения S ,
 G_i – это инвариант слоя s_i ,
 g_j – это инвариант объекта l_j .
- $G_0(S)$ – «истина», если $[G_1(S_1) \& G_2(S_2) \& \dots \& G_n(S_n)]$ – «истина».
- $G_i(S_i)$ – «истина», если $[g_1(l_1) \& g_2(l_2) \& \dots \& g_m(l_m)]$ – «истина».
- Примечание. Инвариант задается предикатом (или логической функцией), принимающей значение «истина» или «ложь».

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ КАК СОВОКУПНОСТЬ ИНФОРМАЦИОННЫХ ОБЪЕКТОВ И ИХ ОТОБРАЖЕНИЙ

В основе построения любой информационной технологии лежит анализ процессов. Традиционная цепочка решения задачи выглядит следующим образом:

$Z \rightarrow M \rightarrow A \rightarrow P \rightarrow PP$

Z – задача

M – математическая модель

A – алгоритм

P – программа

PP – программный продукт

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ КАК СОВОКУПНОСТЬ ИНФОРМАЦИОННЫХ ОБЪЕКТОВ И ИХ ОТОБРАЖЕНИЙ

Для более сложных задач цепочка усложняется, в ней появляются

- DI – исходные данные
- DR – результирующие данные
- PS – программные системы (в том числе интеллектуальные)
- QE – выводы на основе эвристических знаний
- IP – информационный продукт

В соответствии с предложенным подходом будем рассматривать все элементы цепочки как типы информационных объектов, а соединяющие их стрелки – как типы отображений этих объектов. Например,

$F1: Z \rightarrow M$, $F2: M \rightarrow DI$, $F3: M \rightarrow A$, и т.д.

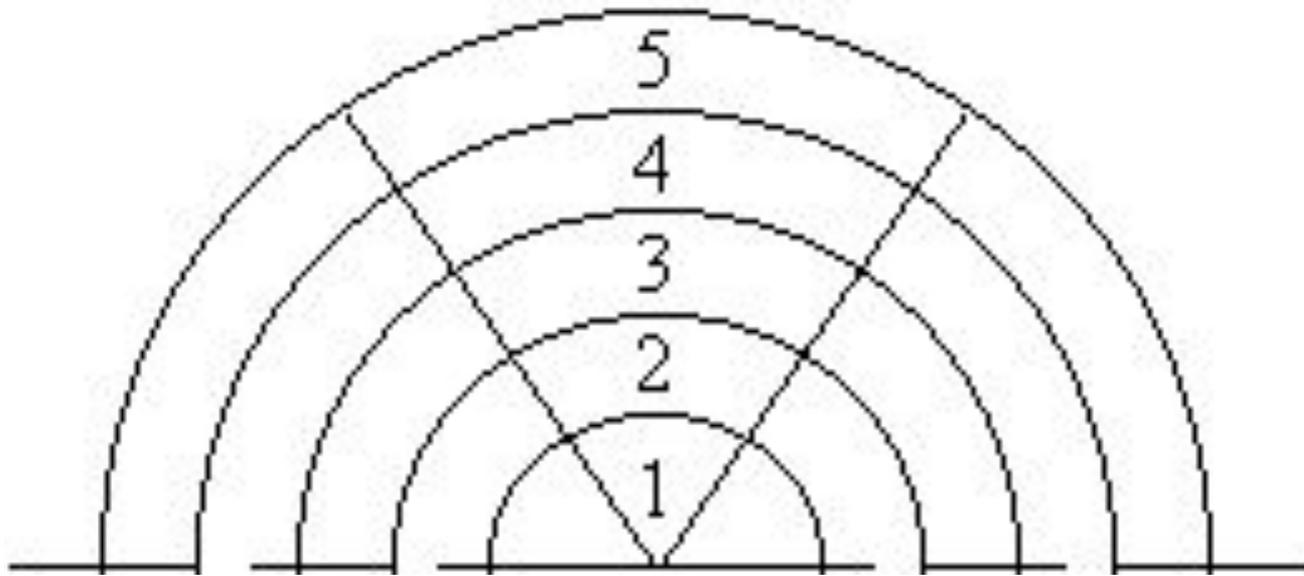
ПОСТРОЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ С ТОЧКИ ЗРЕНИЯ ФРАКТАЛЬНОГО ПОДХОДА

Основываясь на фрактальном подходе, информационную технологию можно определить как ***совокупность информационных объектов и их отображений.***

Процесс разработки ИТ, в таком случае, заключается в разработке способов описания информационных объектов и способов отображения объектов.

Реализация этих способов дает инструментальные средства поддержки конкретной ИТ.

ИНФОРМАЦИОННЫЕ МИРЫ



1 – мир математических моделей

2 – мир алгоритмов

3 – мир программ

4 – мир моделей данных

5 – мир логико-лингвистических моделей

данных

ПРЕИМУЩЕСТВА ФРАКТАЛЬНОГО ПОДХОДА

Использование ФС-моделей при построении ИТ позволяет:

- концептуально объединить разные способы структурирования знаний, применяемые в различных областях знаний, необходимых для решения поставленной задачи
- применять различные способы выделения и упорядочивания миров в зависимости от поставленной цели
- акцентировать внимание на мирах программ, данных, знаний, являющихся предметом данной работы, отдавая должное значимости других миров
- обеспечить последовательное осуществление стратегии построения ИТ (что достигается введением инварианта) независимо от глубины расслоения отдельных миров

ВОПРОСЫ К ЛЕКЦИИ

1. Стратификационный подход к структурированию знаний.
2. Фрактальный подход к структурированию знаний.
3. Фрактальность (определение, базовые свойства, примеры)
4. Понятие ФС-модели
5. Графическое представление ФС-модели
6. Математическое описание ФС-модели
7. Задание инвариантов в ФС-модели
8. Традиционная и усложненная цепочка решения задачи.
9. Информационная технология с точки зрения фрактального подхода.
10. Базовые информационные миры при построении ИТ
11. Преимущества фрактального подхода

МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ:

- **Логические**
- **Логико-лингвистические**
 - Продукционные
 - Фреймовые
 - Семантические сети

■ **Онтологии**

Онтологии – базы знаний специального вида, которые могут «читаться» и пониматься, отчуждаться от разработчика и/или физически разделяться их пользователями. Это формально представленные знания на базе концептуализации (описания множества объектов и понятий, знаний о них и связей между ними)

МОДЕЛИРОВАНИЕ ЗНАНИЙ: ПРИМЕР ОПИСАНИЯ ОНТОЛОГИИ НА ЯЗЫКЕ XML

The screenshot displays a software application window titled "VAZ_Расчет и оптими..." with a sidebar menu on the left. The main window shows an XML document titled "C:_VAZ\model\Онтология приложения ТЭЦ ВА3а.xml" in Microsoft Internet Explorer. The XML content is as follows:

```
<?xml version="1.0" encoding="windows-1251" ?>
<tree>
  <entity id="e1">
    <description>Иерархия схем</description>
    <oncontextmenu />
    <image>images/book.gif</image>
    <imageOpen>images/bookOpen.gif</imageOpen>
    <link>fwork.htm</link>
  </entity>
  <contents>
    <entity id="e2">
      <description>Общая схема ТЭЦ</description>
      <image>images/book.gif</image>
      <imageOpen>images/bookOpen.gif</imageOpen>
      <link>model/tec/scheme.htm</link>
    </entity>
    <entity id="e3">
      <description>Расчет модели</description>
      <image>images/book.gif</image>
      <imageOpen>images/bookOpen.gif</imageOpen>
      <link>model/tec/36/edit.htm</link>
    </entity>
    <entity id="e4">
      <description>Исходное состояние</description>
      <image>images/book.gif</image>
      <imageOpen>images/bookOpen.gif</imageOpen>
      <link>edit.htm</link>
    </entity>
    <entity id="e5">
      <description>Исходное состояние</description>
      <image>images/book.gif</image>
      <imageOpen>images/bookOpen.gif</imageOpen>
      <link>edit.htm</link>
    </entity>
    <entity id="e6">
      <description>Параметры турбин Т-100/120</description>
      <image>images/book.gif</image>
      <imageOpen>images/bookOpen.gif</imageOpen>
      <link>model/tec/120/xtable.htm</link>
    </entity>
  </contents>
</tree>
```

Below the XML editor, a technical diagram titled "На сальниковые подогреватели" is shown. It includes a schematic with components labeled "Д-1" and "КН", and a table of parameters:

Г _{ОХ.В.}	155.578 т/ч
ρ _{ОХ.В.} _{ВХ}	31.463 ^с кг
ρ _{ОХ.В.} _{ВЫХ}	2.7989 ^с кг
Т _{ОХ.В.} _{ВХ}	665.545 С
Т _{ОХ.В.} _{ВЫХ}	3.2805 С

The bottom of the screenshot shows the Windows taskbar with the system clock at 11:07.

