

# Интеллектуальные информационные системы

## Лекция 2

# *Анализ предметной области на предмет применимости ИИС*

Самым первым вопросом, который необходимо решать всякий раз перед началом разработки конкретной ИИС, является вопрос: а следует ли вообще разрабатывать ИИС для данного приложения? Насколько оправдано будет использование методов и средств ИИС для данной задачи и даст ли это ощутимый эффект?

- **Предметная область** - объектно-ориентированным образом выделенная и формально описанная область человеческой деятельности (множество сущностей, описывающих область исследования или экспертизы).
- **Проблемная область** - предметная область плюс совокупность решаемых в ней задач.
- **Инженер по знаниям** (когнитолог, инженер-интерпретатор) - специалист по ИИ, выступающий в роли промежуточного буфера между экспертом и базой знаний.
- **Эксперт** - высококвалифицированный специалист, согласившийся поделиться опытом в рассматриваемой предметной области.

Системный анализ проблемной области на предмет применимости / неприменимости технологии ИИС заключается в общем случае в проведении трех видов исследований, которые могут быть оформлены в виде последовательных этапов:

- Проведение на основе системы выбранных критериев анализа на уместность (обоснованность) разработки ИИС для данной ПО.
- Проведение на основе системы выбранных критериев анализа на оправданность разработки ИИС.
- Проведение на основе системы выбранных критериев анализа на возможность разработки ИИС для данной ПО.

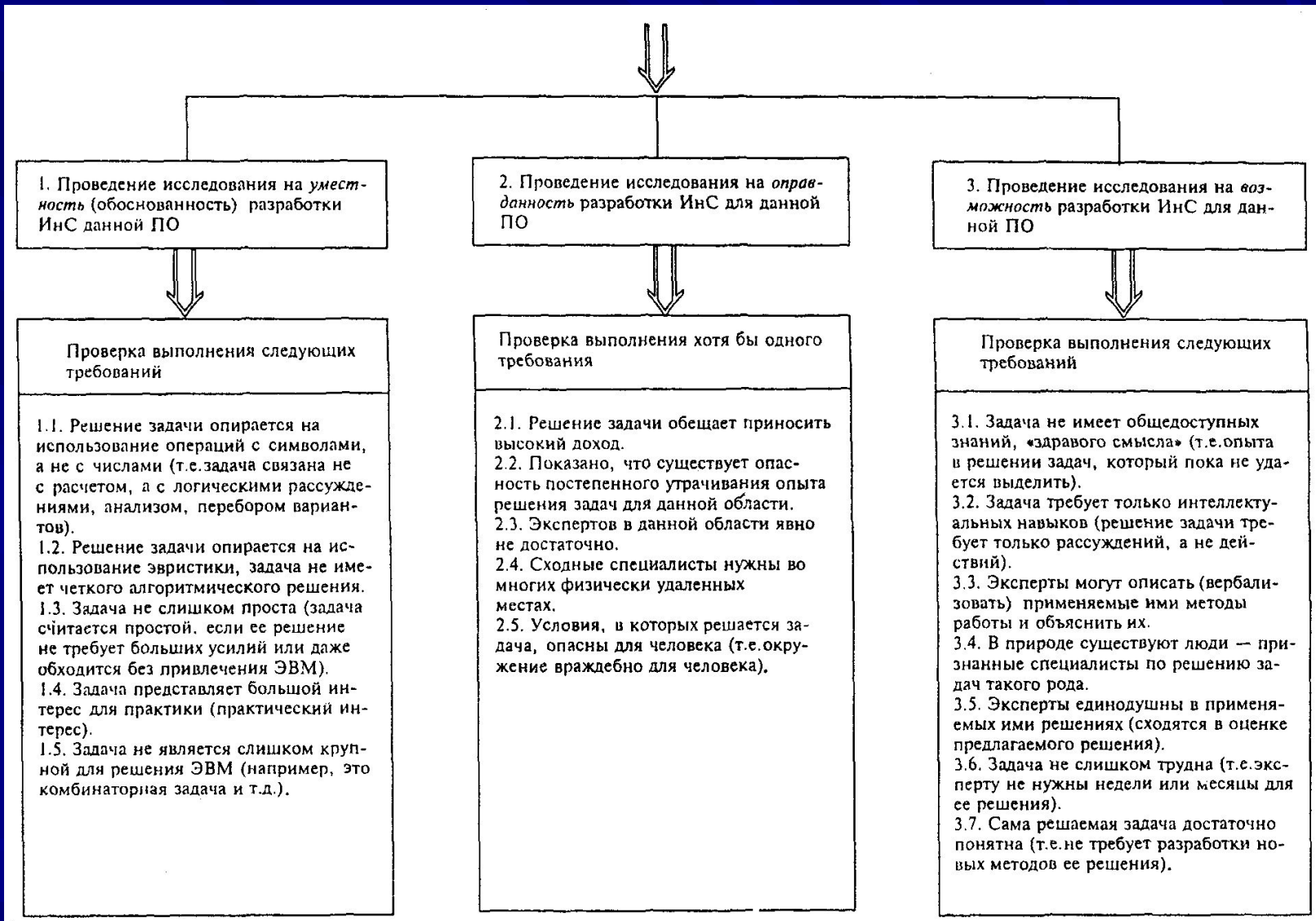


Схема системного анализа ПО на применимость технологии ИИС

# *Классификация ИИС*

# Все существующие ИИС можно разбить на два класса:

- общего назначения и
- специализированные.

К ИИС общего назначения относятся те, которые не только исполняют заданные процедуры, но и на основе метапроцедур генерируют и исполняют процедуры решения новых конкретных задач.



Технология использования ИИС общего назначения (инженерия знаний)



# К специализированным ИИС

отнесены те, которые выполняют решение фиксированного набора задач, predetermined при проектировании системы. Для использования таких систем требуется наполнить их данными, соответствующими выбранному приложению (прикладным задачам, предметной области).

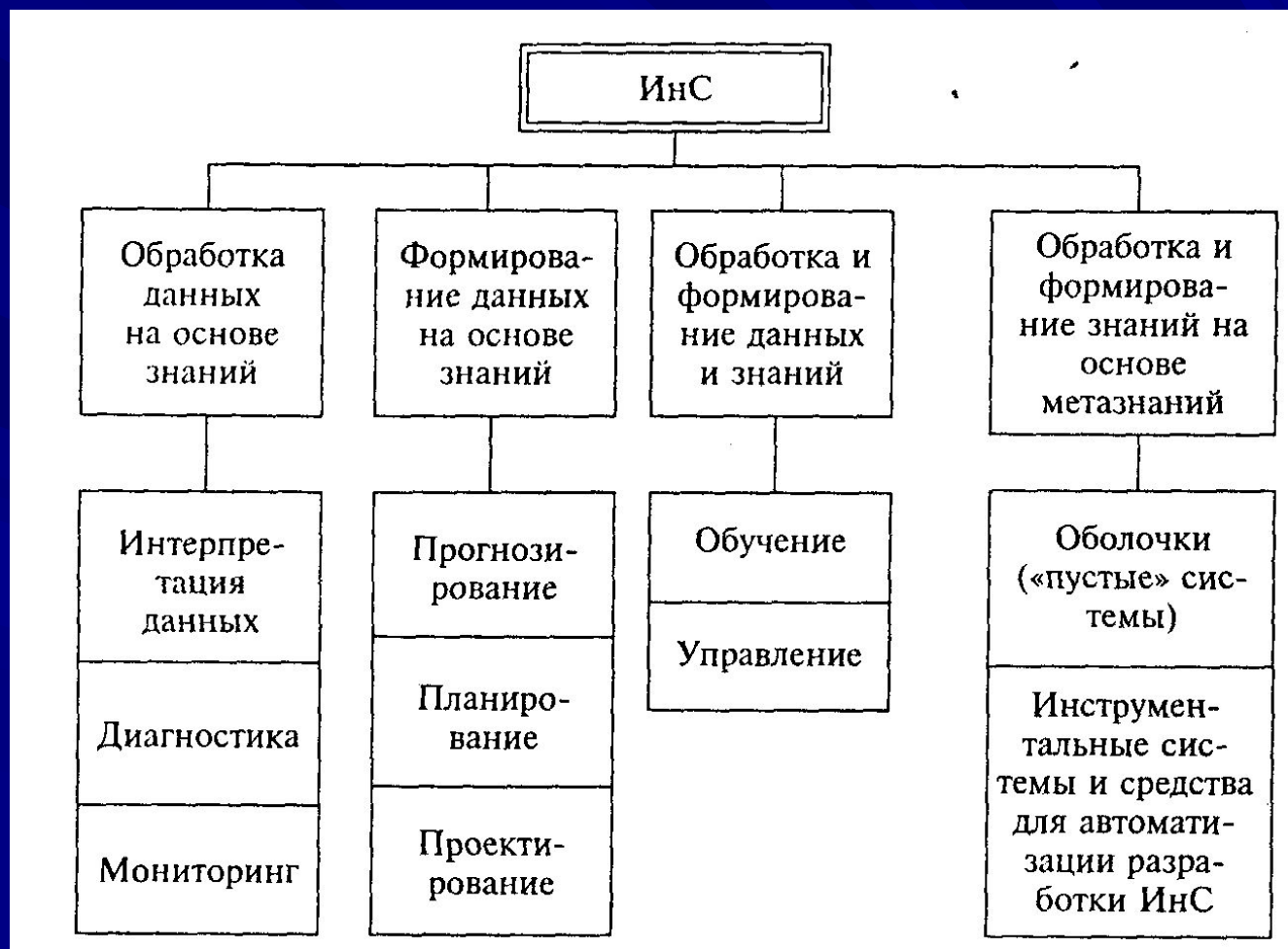
# При разработке специализированных ИС использовалась технология традиционного (процедурного) программирования



# Классификация ИИС на основе решаемых задач

В общем случае все ИИС можно подразделить на решающие задачи анализа и на решающие задачи синтеза.

- Примерами задач анализа являются задачи интерпретации данных и диагностики;
- Примерами задач синтеза являются задачи проектирования и планирования.
- Комбинированные задачи - задачи обучения, мониторинга, управления.



# Классификация ИИС по типам решаемых задач

# Типы задач, решаемых ИИС

Тип задачи	Определение (адресуемые задачи)
Интерпретация	Процесс определения смысла данных (построение описаний по наблюдаемым данным).
Диагностика	Процесс обнаружения неисправностей (в технике и в живых организмах).
Слежение (мониторинг)	Непрерывная интерпретация данных в реальном масштабе времени и сигнализация о выходах параметров за допустимые пределы.
Прогнозирование	Предсказание будущих событий на базе моделей прошлого и настоящего (вывод вероятных следствий из заданных ситуаций).
Планирование	Конструирование плана, т.е. программы действий
Проектирование	Построение спецификаций на создание объектов с заранее определенными свойствами.
Отладка, ремонт	Выработка рекомендаций по устранению неисправностей
Обучение	Диагностика, интерпретация, планирование, проектирование.
Управление	Интерпретация, прогноз, планирование, моделирование, оптимизация выработанных решений, мониторинг.

# Классификация ИИС на основе признаков интеллектуальности

В различных ИИС признаки интеллектуальности развиты в неодинаковой степени и редко, когда все признаки реализуются одновременно. Каждому из признаков интеллектуальности соответствует свой класс ИИС:

- Системы с интеллектуальным интерфейсом;
- Экспертные системы;
- Самообучающиеся системы;
- Адаптивные системы

**Интеллектуальные информационные системы**

**Коммуникативные способности (интеллектуальный интерфейс)**

Интеллектуальные базы данных

Естественно-языковой интерфейс

Гипертекстовые системы

Контекстные системы помощи

Когнитивная графика

**Решение сложных задач (экспертные системы)**

Классифицирующие системы

Доопределяющие системы

Трансформирующие системы

Многоагентные системы

**СПОСОБНОСТЬ К САМООБУЧЕНИЮ (САМООБУЧАЮЩИЕСЯ СИСТЕМЫ)**

Индуктивные системы

Нейронные сети

Системы на прецедентах

Информационные хранилища

**Адаптивность (адаптивные ИС)**

CASE-технологии

Компонентная технология

# *Системы с интеллектуальным интерфейсом*



# Интеллектуальные базы данных (БД)

- отличаются от обычных БД  
возможностью выборки по запросу  
необходимой информации, которая  
может явно не храниться, а выводиться  
из имеющейся в БД.

# Примеры запросов:

- *Вывести список товаров, цена которых выше среднеотраслевой. Для выполнения запроса необходимо сначала проведение статистического расчета среднеотраслевой цены по всей базе данных, а уже после этого собственно отбор данных.*
- *Вывести список товаров-заменителей некоторой продукции. Для выполнения запроса необходимо вывести значения характерных признаков объекта, а затем поиск по ним аналогичных объектов.*
- *Вывести список потенциальных покупателей некоторого товара. Для выполнения запроса требуется сначала определить список посредников-продавцов, выполняющих продажу данного товара, а затем провести поиск связанных с ними покупателей.*

# *Естественно-языковой интерфейс*

- предполагает трансляцию естественно-языковых конструкций на внутримашинный уровень представления знаний.
- Для этого необходимо решать задачи морфологического, синтаксического и семантического анализа и синтеза высказываний на естественном языке.

- **Морфологический** анализ предполагает распознавание и проверку правильности написания слов по словарям.
- **Синтаксический** контроль - разложение входных сообщений на отдельные компоненты (определение структуры) с проверкой соответствия грамматическим правилам внутреннего представления знаний и выявления недостающих частей.
- **Семантический** анализ - установление смысловой правильности синтаксических конструкций.
- **Синтез** высказываний решает обратную задачу преобразования внутреннего представления информации в естественно-языковое.

# *Естественно-языковой интерфейс используется для:*

- доступа к интеллектуальным БД;
- контекстного поиска документальной текстовой информации;
- голосового ввода команд в системах управления;
- машинного перевода с иностранных языков.

# Системы когнитивной графики

позволяют осуществлять интерфейс пользователя с ИИС с помощью графических образов, которые генерируются в соответствии с происходящими событиями.

Такие системы используются в мониторинге и управлении оперативными процессами.

Графические образы в наглядном и интегрированном виде описывают множество параметров изучаемой ситуации.

• **Когнитивность**  
(лат. *cognitio*, «познание, изучение, осознание») — способность к умственному восприятию и переработке внешней информации.

# *Системы когнитивной графики*

используются также в обучающих и тренажерных системах на основе использования принципов виртуальной реальности, когда графические образы моделируют ситуации, в которых обучаемому необходимо принимать решения и выполнять определенные действия.

# *Гипертекстовые системы*

*Предназначены для реализации  
поиска по ключевым словам в  
базах текстовой информации*



# *Понятие гипертекста*

- Гипертекст (ГТ) — одна из фундаментальных моделей представления знаний, выраженных в текстовом виде.
- Обычный (одномерный) текст рассматривается как длинная строка символов, читаемая в одном направлении.
- Многомерный текст (ГТ) включает точки ветвления, в которых чтение можно продолжать в нескольких направлениях в зависимости от информационных потребностей читателя.

- Современные гипертекстовые системы позволяют пользователю самостоятельно формировать альтернативные траектории навигации по ГТ, максимально отвечающие его текущим интересам.

# *В основе ГТ лежат следующие основные идеи:*

- Текст разбивается на фрагменты, представляющие его семантические единицы (сеты). Между ними устанавливаются связи, которые могут наделяться именами.
- В отличие от обычного текста, который читается последовательно (в порядке, определенном его автором), ГТ можно читать, двигаясь по разным траекториям, образованным связанными сетями.
- Активируемые переходы выбираются читателем (пользователем). Имена (типы) связей облегчают решение задачи выбора перехода. Например, «раздел А», «аргументы за...», «определение термина...», «замечания», «детализация положения...» и др.

Под гипертекстом понимается форма организации семантической информации, предусматривающая ее разделение на фрагменты, для каждого из которых заданы переходы к родственным фрагментам. Исторически первым гипертекстовым документом можно считать Библию.

*Гипертекстовый документ может быть как электронным, так и бумажным*

В настоящее время под ГТ также понимают многоцелевой информационный фонд, характеризующийся полнотой изложения сведений по определенной тематике и наличием ссылок между статьями.

В гипертекстовом документе может быть представлено несколько уровней детализации материала. Такие документы моделируются деревьями или сетями. Если в обычном тексте автором или экспертом расставлены точки ветвления (ссылки), позволяющие читать его, двигаясь по разным траекториям, то текст превращается в ГТ. В графовой модели ГТ вершины соответствуют вычлененным фрагментам текста, а ребра — возможным переходам между ними. Каждый путь на графе представляет отдельную линию прочтения текста.

# ГТ как информационная модель интегрирует положительные стороны энциклопедий, монографий и тезаурусов

- От энциклопедий ГТ наследует возможности детального представления понятий, быстрого просмотра материала (без использования ссылок), алфавитного поиска;
- от монографий — возможности представления материала с разной степенью глубины и детальности, поиска по оглавлению;
- от тезаурусов — раскрытие объема и содержания понятий, а также связей между понятиями.

# *Гипертекстовая информационная технология (ГИТ)*

— технология обработки семантической информации, основанная на использовании ГТ. Она относится к проблематике ИИ, так как ее содержанием является представление, поиск и обработка семантической информации, выраженной в текстах.



# *Области применения ГИТ:*

- информационные ресурсы и технологии Internet;
- гипертекстовые информационно-поисковые системы;
- гипертекстовые информационные модели экономических систем;
- базы данных с гипертекстовой организацией;
- представление электронной документации (в том числе, контекстно-зависимой и ситуативно-зависимой справки по программным средствам);
- электронные записные книжки;
- электронные картотеки, словари, энциклопедии, справочники;
- обучающие системы;
- экспертные системы;
- организация пользовательского интерфейса и др.

# Формализованная модель гипертекста

- В основе моделей ГТ лежит понятие информационно-справочной статьи (ИСС), выступающей в качестве информационной единицы ГТ.
- В формализованной модели ГТ ИСС соответствует информационному объекту, содержание которого характеризуется смысловым единством и логической целостностью.

В конкретных технологиях ИСС называют по-разному:

- страница,
- статья,
- тема и др.

Она может включать информацию, представленную в разных формах:

- текст,
- таблицы,
- фрагменты программного кода (макросы, скрипты),
- внедренные цифровые объекты, а также ссылки на подобные объекты (графика, звук, видео,
- управляющие элементы пользовательского интерфейса и т. д.), включаемые в ИСС при ее загрузке.

Элементом ИСС могут быть присвоены метки, уникальные в рамках ИСС. Кроме того, элементы (слова, фразы, предложения, ячейки таблиц, пиктограммы, фрагменты изображений, кнопки и др.) могут наделяться интерактивным поведением. Такие элементы называются гиперссылками.

При воздействии на гиперссылку  
(например, щелчке на ней мышью)  
инициируется переход:

- к началу другой ИСС;
- фрагменту другой ИСС,  
начинающемуся с элемента, который  
имеет указанную метку;
- фрагменту данной ИСС,  
начинающемуся с элемента, который  
имеет указанную метку.

- Гиперссылки задают направления переходов между ИСС и фрагментами ИСС, что фактически соответствует точкам ветвления при чтении документа.
- Гиперссылка содержит указатель на ИСС и, возможно, ее фрагмент. В Internet подобные указатели представляются в виде URL, задающих адреса соответствующих ресурсов.
- Гиперссылки, указывающие на фрагменты текущей ИСС, называются локальными. Гиперссылки, указывающие на другие ИСС, называются глобальными.

Графические иллюстрации и мультимедийные представления, содержащие интерактивные элементы, называются гиперграфикой и гипермедиа соответственно. Эти же понятия часто используются по отношению к документам, включающим гиперграфику и гипермедиа.

# Формализованная модель ГТ состоит из двух слоев.

- Первый слой представляет отображаемое на экране содержимое документа, в котором гиперссылки по умолчанию выделены цветом, подчеркиванием или изменением шрифта.
- Адреса переходов (идентификаторы ИСС и метки их фрагментов) хранятся во втором, скрытом слое модели.



Для работы с гипертекстовой системой, включающей множество связанных документов, не требуется «сборка» интегрального документа. Входящие в систему документы могут храниться на одном или множестве компьютеров (узлах сети). При этом физически распределенная система является логически единой.

$(x_0, x_1, \dots, x_{11})$

# описывает кортеж:

$$(x_0, x_1, \dots, x_{11}), \quad (1)$$

где  $x_0$  – имя ИСС;

$x_1$  – заголовок ИСС;

$x_2$  – аннотация ИСС;

$x_3$  – точка входа в ИСС;

$x_4$  – множество текстовых фрагментов, входящих в ИСС;

$(x_0, x_1, \dots, x_{11})$   
 $x_5$  – множество цифровых информационных объектов, входящих в ИСС;

$x_6$  – множество программных объектов, входящих в ИСС;

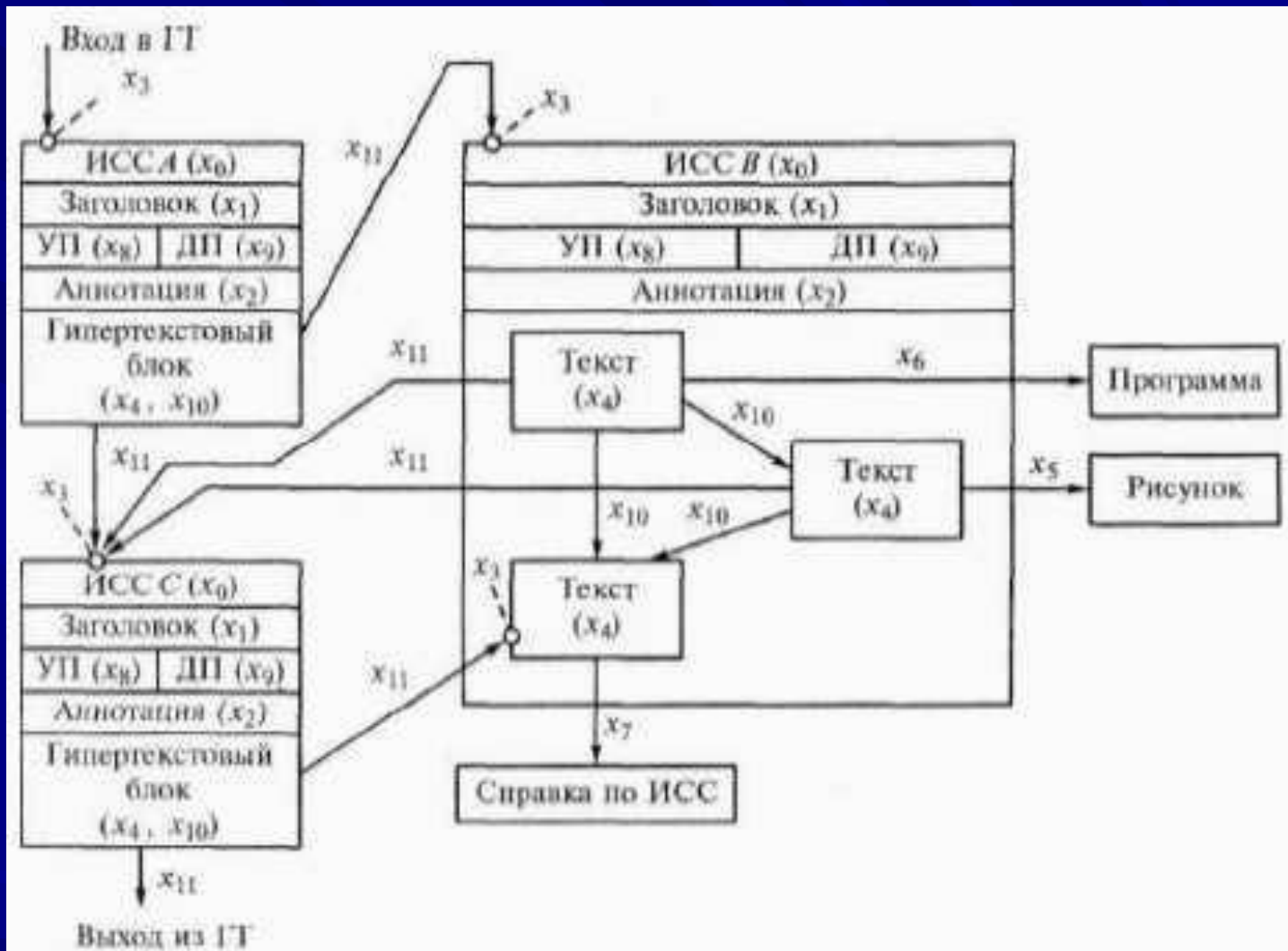
$x_7$  – справка по ИСС;

$x_8$  – признак ускоренного просмотра ИСС;

$x_9$  – признак детального просмотра ИСС;

$x_{10}$  – список гиперссылок внутри ИСС;

$x_{11}$  – список гиперссылок между ИСС.



Структура гипертекста, описываемого (1)

# *Условно-типовая модель гипертекста*

Недостатком формализованной модели ГТ связан с отсутствием в ней возможности явного определения типов гиперссылок.

В условно-типовой модели все гиперссылки имеют явно указанный тип. Данная модель ГТ включает:

- *тезаурус,*
- *список главных тем,*
- *совокупность указателей.*

Обязательным компонентом является тезаурус Предметной области, к которой относится информационная система (ИС).

Тезаурус — упорядоченный перечень терминов, в котором отражены семантические отношения между ними.

Каждый термин в тезаурусе снабжается его текстовой характеристикой (статьей).

Тезаурус позволяет пользователю ГТ уточнять как содержание (смысл), так и объем интересующего его термина.

- Для упрощения работы с ГТ, а также повышения эффективности поиска по нему как в полуавтоматическом режиме (с участием человека), так и в автоматическом режиме (в ГИПС) в условно-типовую модель ГТ включаются список главных тем и указатели.
- Список главных тем делит ГТ на сегменты, соответствующие более или менее независимым частям (срезам или аспектам) Предметной области. Таким образом, он отражает самое общее представление о тематике ГТ.

# Указатель

- - упорядоченная установленным образом последовательность информационных объектов (понятий, выражений, обозначений и т. п.), ссылающихся на ИСС, в которых эти объекты упоминаются.

В зависимости от характера объектов указатели подразделяются на

- предметные,
- именные,
- событийные,
- библиографические и др.

По принципу упорядочения различают

- алфавитные,
- хронологические,
- систематические и прочие виды указателей.

Гипертекст может включать один или несколько указателей.



В лингвистике выделено около 200 семантических типов отношений. Наиболее часто употребляются 10 типов, используемых в условно-типовой модели.

Тип связи	Обозначение
Синоним	СН
Род-вид	РВ
Вид-род	ВР
Часть-целое (укрупнение)	ЧЦ
Целое-часть (декомпозиция)	ЦЧ
Процесс-надпроцесс	ПН
Процесс-подпроцесс	ПП
Причина-следствие	ПС
Следствие-причина	СП
Ассоциация	АС

Графовой интерпретацией условно-типовой модели является семантическая сеть.

В рамках условно-типовой модели ИСС включает

- имя,
- заголовков,
- собственно текст (содержимое)
- список ссылок на ИСС, связанные с данной ИСС различными типами отношений.

Ссылки относятся только к «ближайшим» родственникам. Такой список ссылок образует локальный справочный аппарат ИСС. Он может быть организован тремя способами.

- Первый способ — в виде списка.
- При втором способе ссылки внедряются в текст (как в энциклопедиях).
- Третий способ является комбинированным. Часть ссылок помещаются после заголовка статьи в виде списка, оставшаяся часть — в самом тексте. При отображении ГТ на экране имена ИСС заменяются их заголовками.