

Тема 3.

Методы представления знаний

- Тузовский А.Ф., Чириков С.В., Ямпольский В.З. **Системы управления знаниями (методы и технологии)**. - Томск: Изд-во НТЛ, 2005
- Балдин К.В., Уткин В.Б. Уткин В.Б. **Информационные системы в экономике: Учебник — 5-е изд.** - М.: Издательско-торговая корпорация «Дашков и К^о», 2008

Методы представления знаний

Стоит ещё раз вспомнить, что

Данными называют формализованную информацию пригодную для последующей обработки, хранения и передачи средствами автоматизации профессиональной деятельности.

Примеры: показатели температуры (воздуха, воды и т.д.), давление, скорость, геометрические параметры (длина, ширина, высота), громкость звука, сила тока, напряжение и многие другие

Тогда что представляют собой знания, **что превращает данные в знания?**

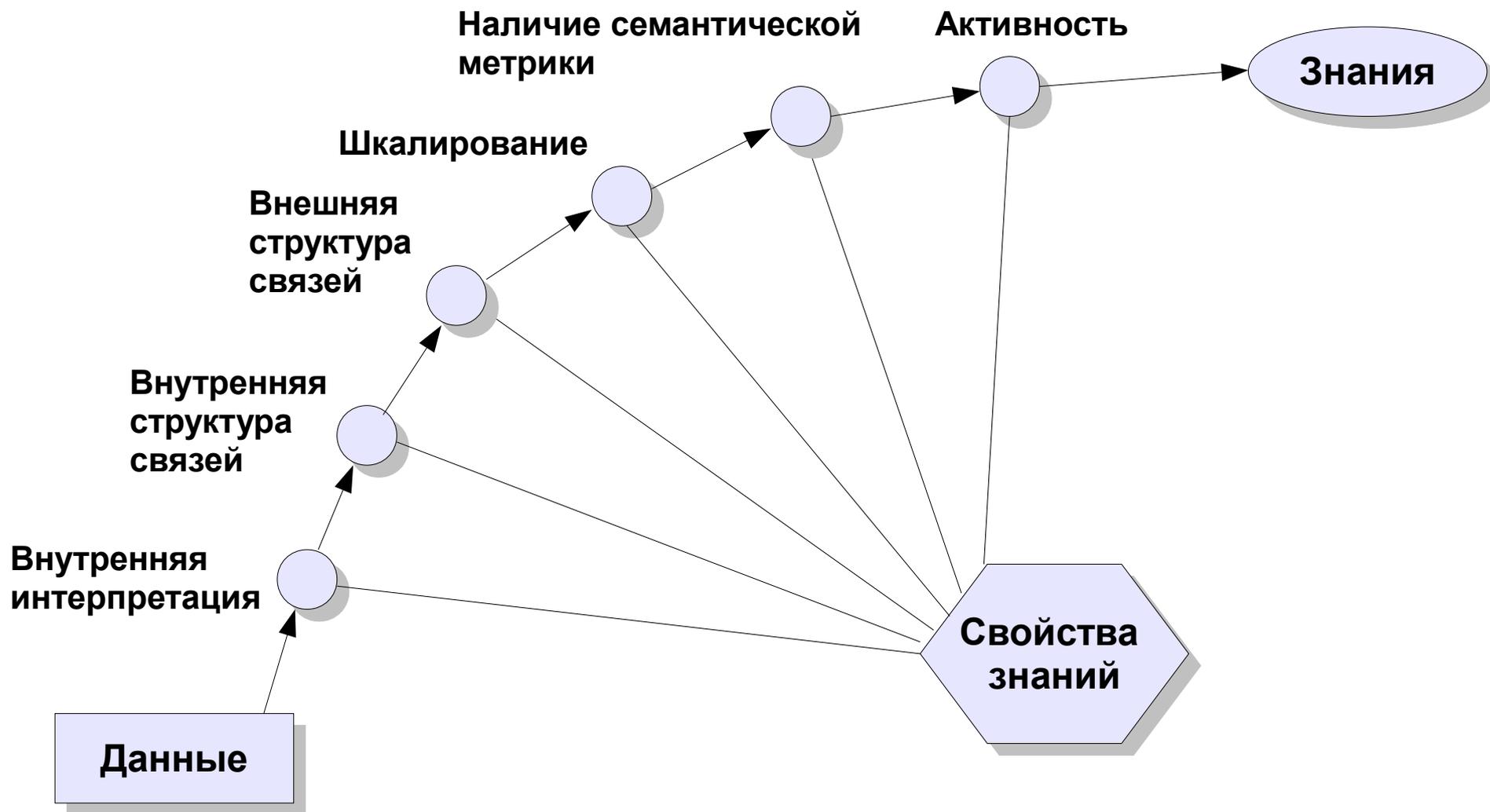


Рис. 3.1. Свойства знаний

Пояснения к рис. 3.1

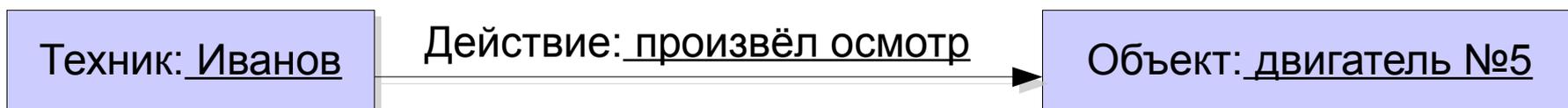
Внутренняя интерпретация. Это свойство предполагает, что в ЭВМ хранятся не только сами данные, но и «**данные о данных**» (метаданные)

Внутренняя структура связей предполагает, что в качестве информационных единиц используются не отдельные данные, а их упорядоченные определёнными отношениями (родовидовыми, причинно-следственными и др.) **структуры**.

*Например: «**факультет — курс — группа — студент**»*

Внешняя структура связей позволяет описывать отдельный объект (понятие). Объекты способны находиться в других отношениях.

*Например: объекты «**техник Иванов**» и «**двигатель №5**»
могут находиться в связи «**произвёл осмотр**»*



Шкалирование подразумевает введение соотношений между различными информационными единицами и их упорядочение путём измерения **интенсивности отношений и свойств**.

*Пример: объект «учебная группа У-05-01»
и понятие «успеваемость»
связаны отношением «занимает первое место»,
которое имеет численную характеристику (шкалу).*



Наличие **семантической метрики** - установление **смысловой связи** между понятиями и объектами, не имеющие «количественного» толкования.

Часто на практике встречаются понятия, к которым не применимы количественные шкалы, однако имеется потребность в установлении их связи и меры близости.

Семантики классифицируются:

- **значение** (объективное содержание)
- **контекстуальный смысл** (определяется связями данного понятия с другими соседствующими в данной ситуации или предметной области)
- **личный смысл** (объективное значение, отраженное через систему взглядов эксперта)
- **прагматический смысл** (определяется текущим знанием о конкретной ситуации).

Например, «условие выполнено» может нести положительный или отрицательный смысл в различных ситуациях.

Более подробно типы семантических связей будут рассмотрены далее

Последнее свойство знаний, которое в наибольшей степени отличает их от данных, это:

Активность: свойство принципиально отличает понятие «знания» от понятия «данные».

Знаниям человека свойственна **познавательная активность** (обнаружение противоречий, неполнота знаний).

В отличие от данных, **знания позволяют получать новые знания.**

Активность знаний позволяет человеку **решать** не только типовые, но и **принципиально новые, нестандартные задачи.**

Рассел Аккоф, один из классиков исследования операций, предложил следующую иерархию:

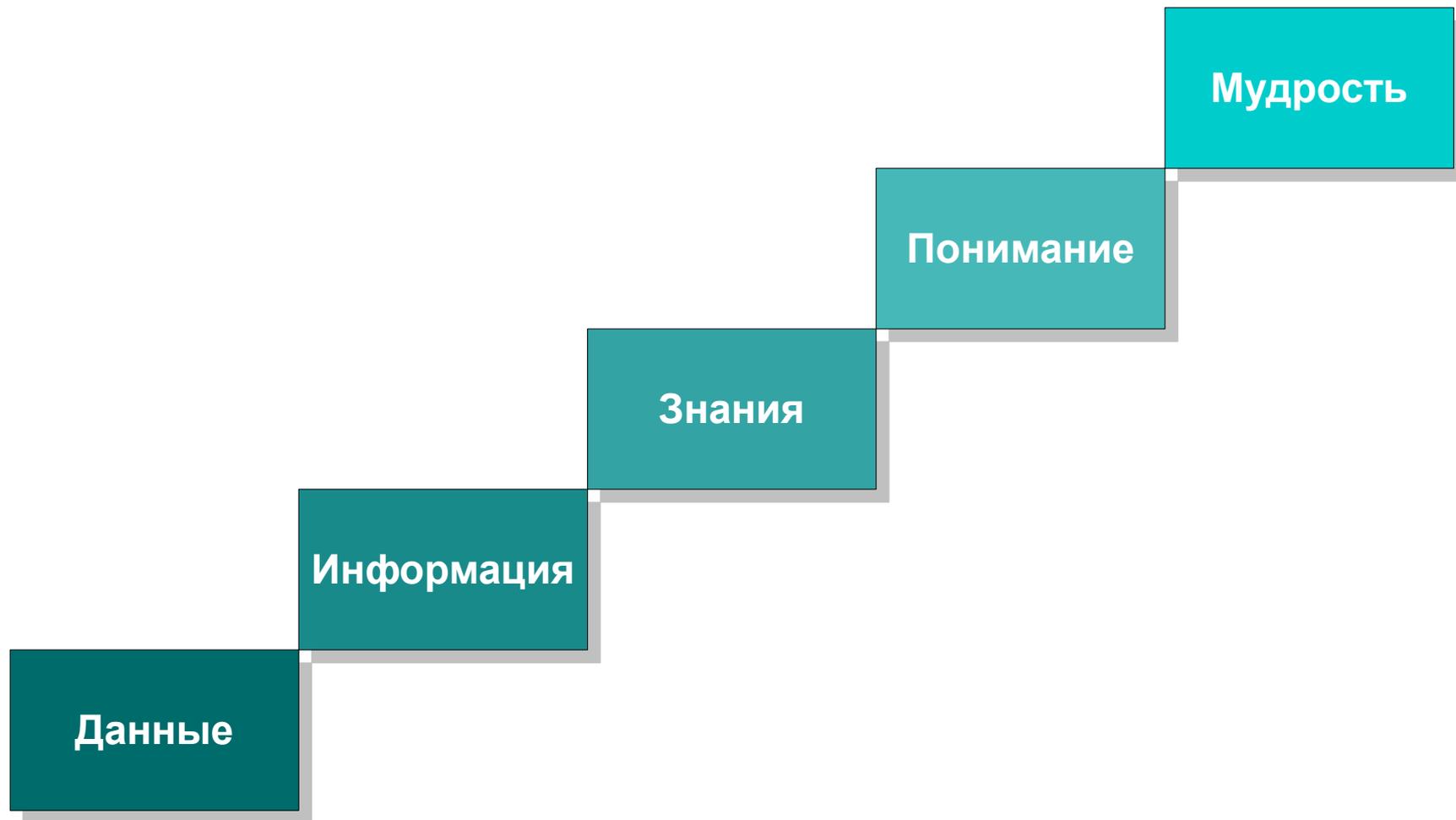


Рис. 3.2. Иерархия: от данных — к мудрости

Согласно иерархии Аккофа:

- **Данные** — это некоторые неупорядоченные символы, рассматриваемые безотносительно к какому-либо контексту.
- **Информация** — это выделенная и упорядоченная часть базы данных, обработанная для использования, то есть отвечающая на вопрос: «Кто?», «Что?», «Где?», «Когда?»...
- **Знание** — это выявленные тенденции или существенные связи между фактами и явлениями, представленные в информации.
- **Понимание** — это осознание закономерностей, содержащихся в разрозненных знаниях, позволяющее ответить на вопрос «Почему?»
- **Мудрость** — взвешенное, оцененное понимание закономерностей с точки зрения прошлого и будущего

Как уже говорилось, **данные** — это набор объективных фактов об объектах, событиях, явлениях, процессах. Это всё то, что регистрируется, описывается и воспринимается человеком. Данные могут быть цифровыми, графическими, аудио, видео и т. п. Они могут описываться на различных языках.

Информация — это данные в определённом контексте (необходимые пользователю, полезные для решения). Также информация — это «**данные, наделённые значимостью и целями**» [Drucker P., 1993].

Не лишено оснований определение, что **информация** — это **данные плюс метаданные**, содержащие их описание (данные о данных).

Продолжая эту логическую цепь рассуждений, знания можно определить как информацию плюс метаинформацию.

Закон РФ «Об информации, информатизации и защите информации» закрепил следующее определение:

«Информация - «сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления».

Данные преобразуются в информацию различными способами.

Согласно работе [Davenport T., Prusak L., 1998] некоторые из них:

- **Контекстуализация** — известно, для какой цели были собраны.
- **Категоризация** — известны единицы анализа или ключевые компоненты данных.
- **Вычисляемость** — данные могут быть проанализированы математически или статически.
- **Корректировка** — устранены ошибки и противоречия в данных.
- **Сжатие** — данные могут быть обобщены в более сжатую форму.

В определении и систематизации знаний особое место занимают воззрения Карла Поппера [Popper K.M., 1972].

В соответствии с этими воззрениями различают три типа знаний:

- **Мир знаний 1** — кодированные знания в физических системах (например, генетический код ДНК), который позволяет объектам адаптироваться к окружающей среде (знания в неодушевлённых объектах).
- **Мир знаний 2** — убеждения и склонности к убеждениям, представления о мире, прекрасном, справедливости, сохраняющаяся в испытаниях, эволюции и опыте (знания в одушевлённых объектах).
- **Мир знаний 3** — совместно используемые лингвистические формулировки, единицы знаний о мире, прекрасном и справедливости, которые помогают выжить в испытаниях и эволюции личности, группе, сообществу, команде, организации, обществу, приобретающему, формулирующему и оценивающему утверждения о знаниях (знания, существующие самостоятельно как артефакты)

Можно сказать, что Мир 1 — это физика, Мир 2 — психология, Мир 3 — продукты человеческого разума.

Согласно Большой Советской Энциклопедии:

«Знание — это проверенный практикой результат познания действительности и верное её отражение в сознании человека»

Многие философы (особенно эмпирики) придерживаются определения:

«Знание — это обоснованное истинное мнение»

Также:

«Знание — это информация в контексте»

«Знание — это способность превращать информацию и данные в эффективные действия» [Applehans W., 1999]

«Знание — это совокупность информации, понимания, осведомлённости, полученных в результате опыта» [Oxford Advanced Learners Dictionary, 1982]

«Знание — это осведомлённость, компетентность или понимание, достигнутое в результате опыта или обучения» и «Знание — это сумма или набор того, что воспринято, обнаружено или выучено» [The American Heritage Dictionary of the English Language, 1992]

Важно отметить наличие **образовательной компоненты** в формировании знания.

Большинство исследователей проблемы интеллектуальных ресурсов отмечают, что знания неотделимы от человека.

Знания могут **осознаваться владельцем** (может объяснить логику своих умозаключений), а могут и **не осознаваться** (действует рационально, эффективно, но не может объяснить, почему он так действует).

Таким образом, можно сказать, что знание — это смесь осознанного и неосознанного в решениях в решениях, действиях и поведении человека.

В связи с вышесказанным, можно дать следующее определение:

«Знание — это сложная сеть понятий и многообразных отношений (оценки, мнения, причинно-следственные и пространственно-временные связи и зависимости) между ними, котрая сознательно или бессознательно используется нейронной сетью головного мозга при необходимости выработки новых суждений или принятия разнообразных решений» [Тузовский А.Ф., Чириков С.В., Ямпольский В.З., 2005]

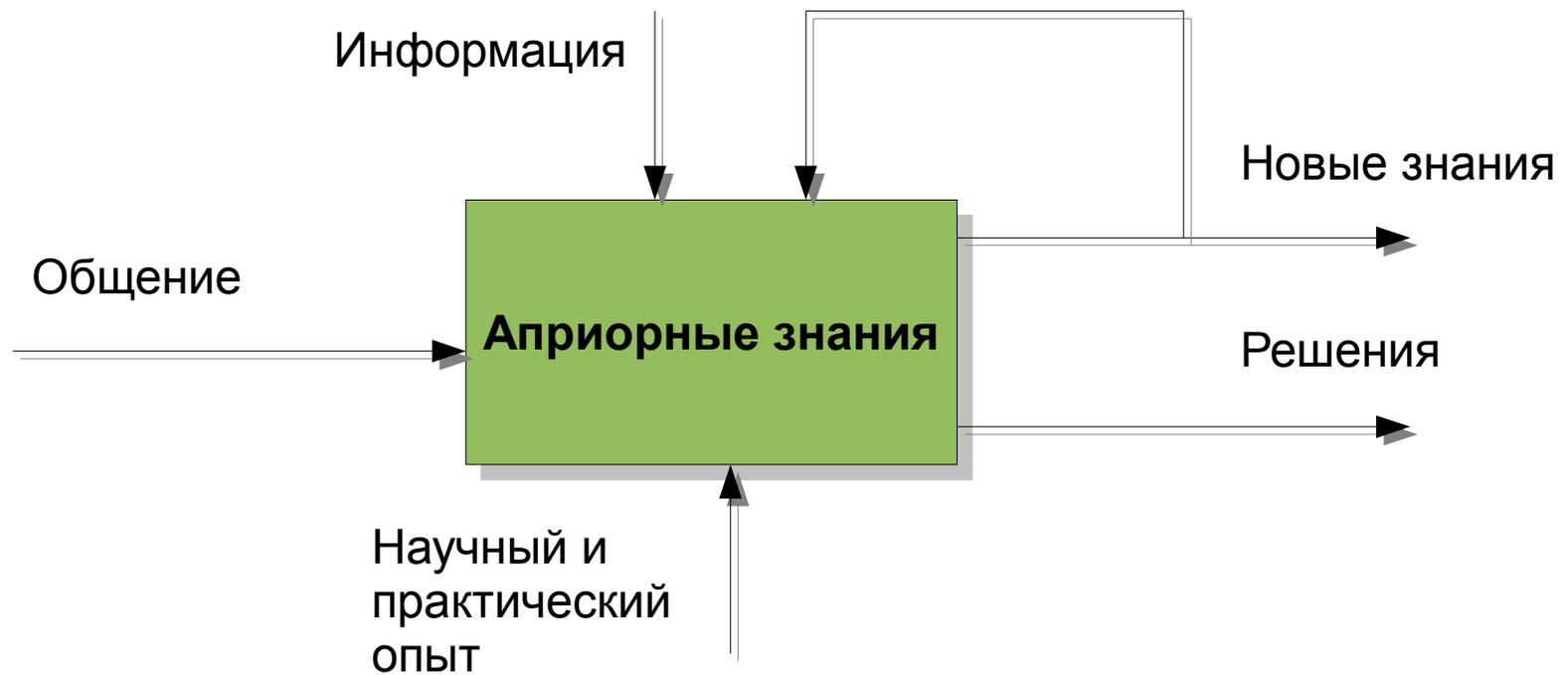


Рис. 3.3. Знания специалиста как модель чёрного ящика
[Тузовский А.Ф., Чириков С.В., Ямпольский В.З., 2005]

«Знание — это комбинация данных и информации, к которым добавлено мнение, мастерство и опыт эксперта, что в результате даёт ценный актив, который может быть использован для оказания помощи в принятии решений» [European Guide to good Practice in Knowledge Management, Part 1]

Классификация знаний.

Знания можно классифицировать по различным критериям, однако рассмотрим следующую классификацию:

По способу существования:

- факты (хорошо известные обстоятельства)
- эвристики (знания из личного опыта экспертов)

По способу использования (в ЭС):

- факты: знания типа «*X – это Y*» («*яблоко — это фрукт*»)
- правила: знания для принятия решений («*ЕСЛИ А, ТО В*»)
- метазнания: указывают способы использования знаний (*пословица «Семь раз отмерь, один — отрежь», очевидно применима не только в отраслях, связанных с отрезанием чего-либо*)

По формам представления :

- декларативные: факты в виде структурированных данных
- процедуральные: алгоритмы в виде процедур обработки фактов

...а также знания классифицируют:

По способу приобретения:

- **научные:** полученные в результате систематического изучения
- **бытовые (житейские):** полученные в результате опыта

По отношению к **классам объектов** и **отдельным объектам** знания делят на

- **Интенциональные знания** — знания, характеризующие или относящиеся к некоторому классу объектов.
- **Экстенциональные знания** — знания, относящиеся к конкретному объекту из какого-либо класса.

Также выделяют:

- **Морфологические и синтаксические знания** — знания о правилах построения структуры описываемого явления (например, правила построения предложений, склонения существительных, спряжения глаголов)
- **Семантические знания** — знания о смысле и значении описываемых явлений
- **Предметные знания** — знания о предметной области, объектах из этой области, их отношениях, действиях над ними и др.

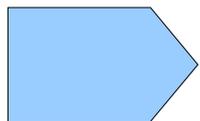
Основные этапы жизненного цикла знаний

В соответствии с методологией системного анализа [Перегудов Ф.И., 1976], в жизненном цикле знаний, как и в жизненном цикле любой сферы деятельности, можно выделить три основных процесса:

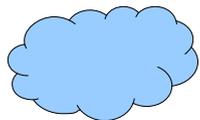
- Выявление потребности в знаниях
- Производство (создание) знаний
- Потребление (утилизация) знаний

Жизненный цикл знаний, предложенный [McElroy M.W., 2003], встроенный в бизнес-среду организационной системы, имеет следующий вид (следующий слайд):

Обозначения:



- Работа со знаниями



- Наборы значений

КУЗ

— кодированные утверждения о знании

РБОЗ

— распределённая база организационных знаний

ОУЗ

— опровергнутое утверждение о знании

ОЗ

— организационное знание

ВУЗ

— подтверждённое утверждение о знании

НУЗ

— неопределённое утверждение о знании

Методы представления знаний.

При проектировании модели представления знаний следует учитывать два требования:

- однородность представления
- простота понимания

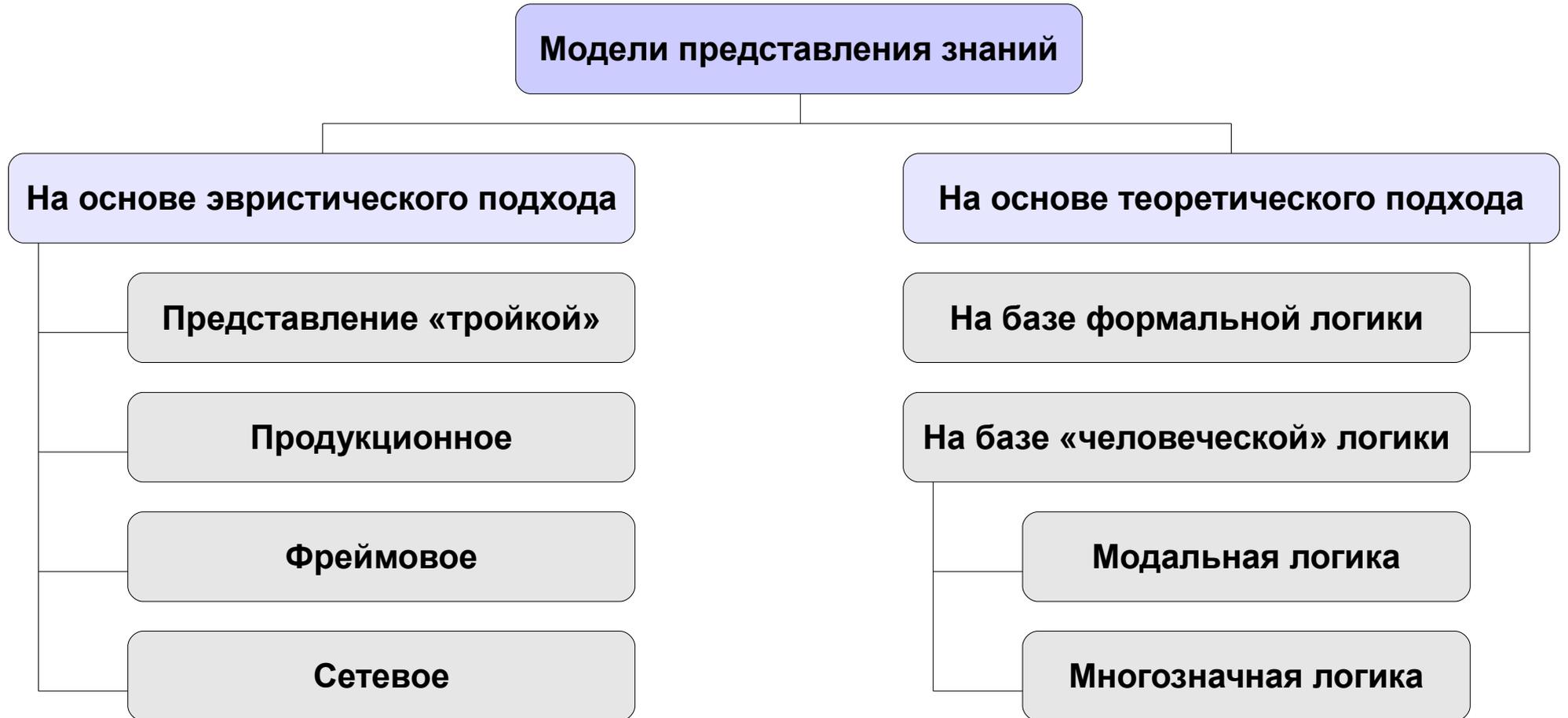


Рис. 3.3. Модели представления знаний

- **Представление знаний «тройкой»:** «*объект-атрибут-значение*» - является одним из первых методов представления знаний и используется для представления фактических знаний в простых системах.
- **Продукционная модель** (модель правил, модель продукций). На данный момент наиболее проработанная и распространённая модель представления знаний, в частности, в ЭС.

Данная модель предусматривает разработку системы продукционных правил вида:
«**ЕСЛИ** A_1 **И** A_2 **И**... A_n , **ТО** B_1 **ИЛИ** B_2 **ИЛИ**... B_n »,

где **A** и **B** – некоторые высказывания, к которым применены логические операции **И** и **ИЛИ**. Если высказывание в левой части (условие) истинно, то истинно высказывание в правой части (следствие).

Полнота базы правил (базы знаний) определяет возможности системы по удовлетворению потребностей пользователей в этой системе.

Пример представления знаний продукционной моделью:

Правила в базе знаний:

*ЕСЛИ в стране падение курса национальной валюты,
ТО материальное положение населения ухудшается.
ЕСЛИ объёмы производства в стране падают,
ТО курс национальной валюты снижается.
ЕСЛИ материальное положение населения ухудшается,
ТО уровень смертности в стране возрастает.*

Если на вход системе поступит факт:

«в стране высокий уровень падения объёмов производства»,

то по цепочке рассуждений будет получен **вывод**

«в стране возрастает уровень смертности»

В современных ЭС в базе знаний (БЗ) хранятся тысячи правил.

Достоинствами продукционных систем являются:

- простота пополнения и изъятия правил
- простота реализации механизма вывода
- наглядность объяснений результатов работы

Основной недостаток:

Связан с **большим количеством правил**, описывающих предметную область, а именно:

- трудность **обеспечения непротиворечивости правил** при их большом числе (требуется создания специальных метаправил для разрешения возникающих противоречий)
- время формирования итогового заключения может быть достаточно большим