



# Искусственный интеллект в автоматизированных системах управления и обработки данных

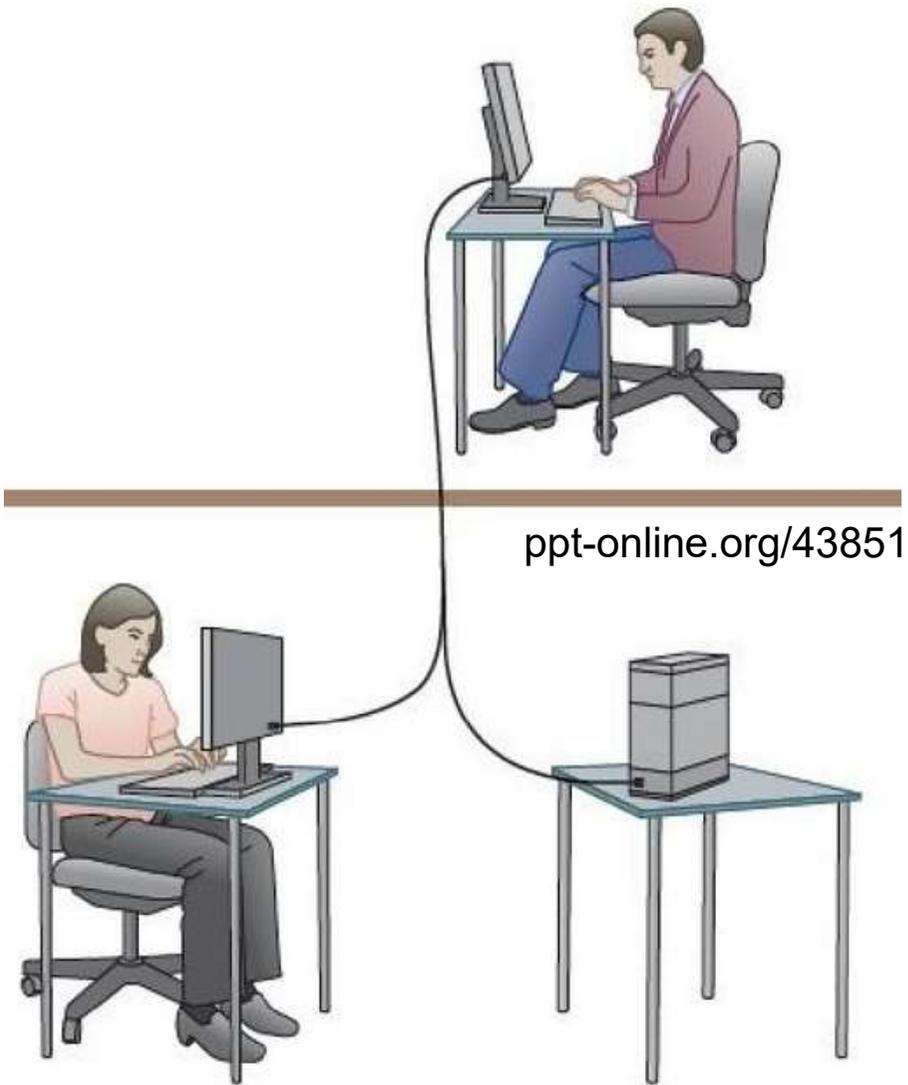
Обладает ли машина знаниями?

Выхованец Валерий Святославович

[valery@vykhovanets.ru](mailto:valery@vykhovanets.ru)

27 апреля 2023 г.

# Может ли машина мыслить?



Тест Тьюринга:

«Человек взаимодействует с одним компьютером и одним человеком на некотором языке.

На основании ответов на вопросы он должен определить, с кем он разговаривает: с человеком или компьютерной программой».

Недостаток теста – отсутствие методики проверки на «умение мыслить».

# Обладает ли машина знаниями?



[ppt-online.org/43851](http://ppt-online.org/43851)



Умение мыслить неразрывно связано с приобретением, накоплением, упорядочиванием и применением знаний.

Тест на обладание знаниями:

«Человек взаимодействует с одним собеседником на некотором языке.

На основании ответов на вопросы он должен определить, обладает ли собеседник знаниями или нет».

Какова методика проверки на «обладание знаниями»?

# Что такое знание?

- **Знание** – психическое образование, которое субъективно (идеально) и не может быть непосредственно изучено.
- **Знание** – результат познавательной деятельности, который как то обоснован и допускает проверку на достоверность.

**Приобретение знаний** осуществляется в результате педагогического процесса, самообразования и жизненного опыта.

**Применение знаний** – адекватная реакция на окружающий мир для достижения поставленных целей.

**Представление знаний** – это отчуждение знания от его носителя во внешней (объективной) форме.

**Объективация знаний** – это признание знаний, представленных в некоторой внешней форме, объективно истинными и социально значимыми.

# Формы рационального познания

- **Понятие** – форма мысли, что-либо выделяющая и называющая: субъект  $s$ , предикат  $p$ , объект  $o$ , атрибут  $a$ .
- **Суждение** – форма мысли, что-либо утверждающая:  $A = s p o a$ .
- **Умозаключение** – умственное действие, связывающее суждения-посылки с суждениями-следствиями:  $A \vee B \rightarrow A$ .
- **Рассуждение** – мыслительная деятельность, направленная на получение из суждений-посылок суждений-следствий.
- **Теория** – суждения-следствия, полученные в процессе рассуждения из некоторых суждений-посылок с помощью умозаключений заданного вида.
- **Знание** – результат отражения реальности, представленный в сознании понятиями и суждениями, утвержденными некоторым множеством умозаключений.

# Что значит «обладать знаниями»?

- **Сужение** области знания (выделение предметной области).
- **Владение** единичными сущностями предметной области (называть, распознавать, различать).
- **Определение** имеющихся понятий и создание новых (интенционально, экстенционально, генетически).
- **Оперирование** суждениями о сущностях и понятиях предметной области (высказывать, запоминать, проверять).
- **Выполнение** умозаключений по правилам, присущим выбранной предметной области.
- **Получение** скрытого знания путем рассуждений на имеющихся знаниях (вывод на знаниях, запросы к знаниям).
- **Приобретение** нового знания путем обучению с учителем, с подкреплением, без учителя.

# Проблемы обработки знаний

- **Искусственные нейронные сети** оказались непригодны для представления и обработки знаний:
  - невозможность объяснения получаемых результатов,
  - отсутствие гарантии верного (полного) решения задачи,
  - непредсказуемость прогностической способности сети,
  - трудности в решении задач по шагам и вычислительных задач,
  - трудность получения обучающей выборки,
  - требуется длительное обучение,
  - катастрофическая забывчивость;
  - риск переобучения сети, и т. д.
- **Символьные методы** оказались мало эффективными для представления и обработки знаний большого объема.

# Модели знаний

- **Модель** знаний – средства формального описания знаний (язык) и методы оперирования ими (операции).
- **Модели знаний:**
  - *логические* (язык логики): исчисление предикатов, дескрипционные логики;
  - *продукционные* (язык продукций): логические, грамматические;
  - *сетевые* (язык ориентированных графов): семантические сети, концептуальные графы, фреймы, «сущность-связь», объектные;
  - *понятийные* (понятийный язык).

# Логические модели

## Логика первого порядка

- **Высказывания** ( $\varphi, \psi$ ):
  - атомарные  $a, b, \dots$ ;
  - $(\varphi), \neg \varphi, \varphi \rightarrow \psi$ ;
  - $\exists a \varphi(a), \forall a \varphi(a)$ .
- **Аксиомы:**
  - $(\varphi \rightarrow (\psi \rightarrow \gamma)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \gamma)),$
  - $(\varphi \rightarrow \psi) \rightarrow (\neg \psi \rightarrow \neg \varphi),$
  - $\varphi \rightarrow (\psi \rightarrow \varphi),$
  - $\varphi(a) \rightarrow \exists b(\varphi(b)),$
  - $\forall a(\varphi(a)) \rightarrow \varphi(b).$
- **Пример (утверждения и факты):**
  - $\exists x S(x), \exists x \exists y P(x, y), \exists x \exists y \exists z M(x, y, z);$
  - $\exists x \neg S(x), \exists x \exists y \neg P(x, y);$
  - $\forall x \exists y \exists z (M(x, y, z) \rightarrow z \wedge \neg P(x, y));$
  - $\forall x \forall y (z \vee S(x) \rightarrow M(x, y, z));$
  - $\exists x \forall y (S(x) \rightarrow \neg P(x, y));$
  - $S(1), S(2), P(0, A), P(1, D), M(0, E, s).$

## Дескрипционная логика

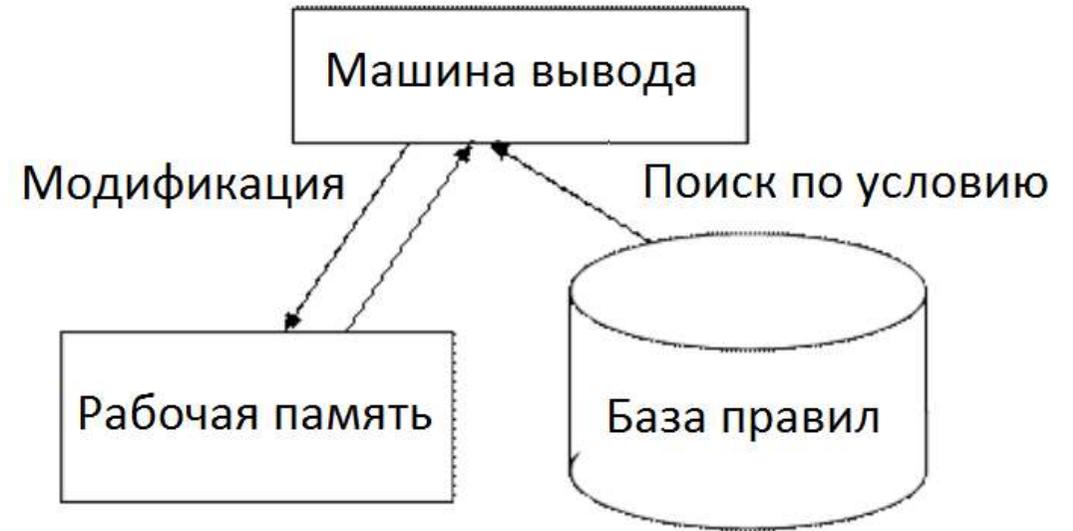
- **Концепты** ( $C, D$ ) и **роли** ( $R: \{(c, d)\}$ ):
  - атомарные концепты  $A, B, \dots$ ;
  - $\top$  (домен),  $\perp$  (пустой концепт);
  - $\neg C, C \sqcap D, C \sqcup D$ ;
  - $\forall R.C, \exists R.C$ .
- **Пример Tbox (утверждения):**
  - Women = Person  $\sqcap$  Female;
  - Mother = Women  $\sqcap \exists \text{hasChild}.\top$ ;
  - $\forall \text{hasChild}.\text{Person} \subseteq \text{Person}$ ;
  - Doctor  $\subseteq$  Person.
- **Пример Abox (факты):**
  - Mary: Woman  $\sqcap \neg \text{Doctor}$ ;
  - Mary:  $\exists \text{hasChild}.\text{Female}$ ;
  - Mary hasChild Peter;
  - Peter: Doctor  $\sqcap \forall \text{hasChild}.\perp$ .

# Недостатки логических моделей

- Монотонность вывода – невозможно пересмотра промежуточных данных вывода (все факты, а не гипотезы).
- Невозможность использовать в качестве переменных предикаты.
- Неразрешимость формул логики первого порядка.
- Сложность вывода на знаниях (логика первого порядка) или низкая выразительность языка (дескрипционная логика).
- Необходимость использования эвристик при выводе.
- Нет средств для представления процедурных знаний.
- Обучение модели только с учителем.

# Продукционная модель

$P = \langle S, A \rightarrow B, Q \rangle$ ,  
 $S$  - предусловие;  
 $A \rightarrow B$  - ядро продукции;  
 $Q$  - постусловие.



**Прямой вывод:** от фактов к заключению.

**Обратный вывод:** от заключения к фактам.

## Пример

Утверждение 1: Если (отдых летом) И (человек активный), ТО (ехать в горы).

Утверждение 2: Если (любит солнце), ТО (отдых летом).

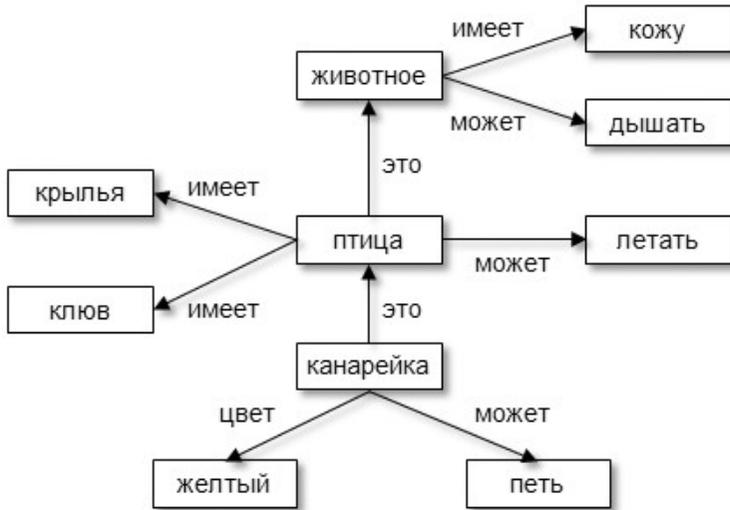
Факты: (человек активный), (любит солнце).

# Недостатки продукционной модели

- Трудности проверки на непротиворечивость и полноту.
- Недетерминированность выбора из активированных продукций.
- Необходимость использования эвристик при выводе.
- Нет средств для представления процедурных знаний.
- Обучение модели только с учителем.

# Сетевые модели

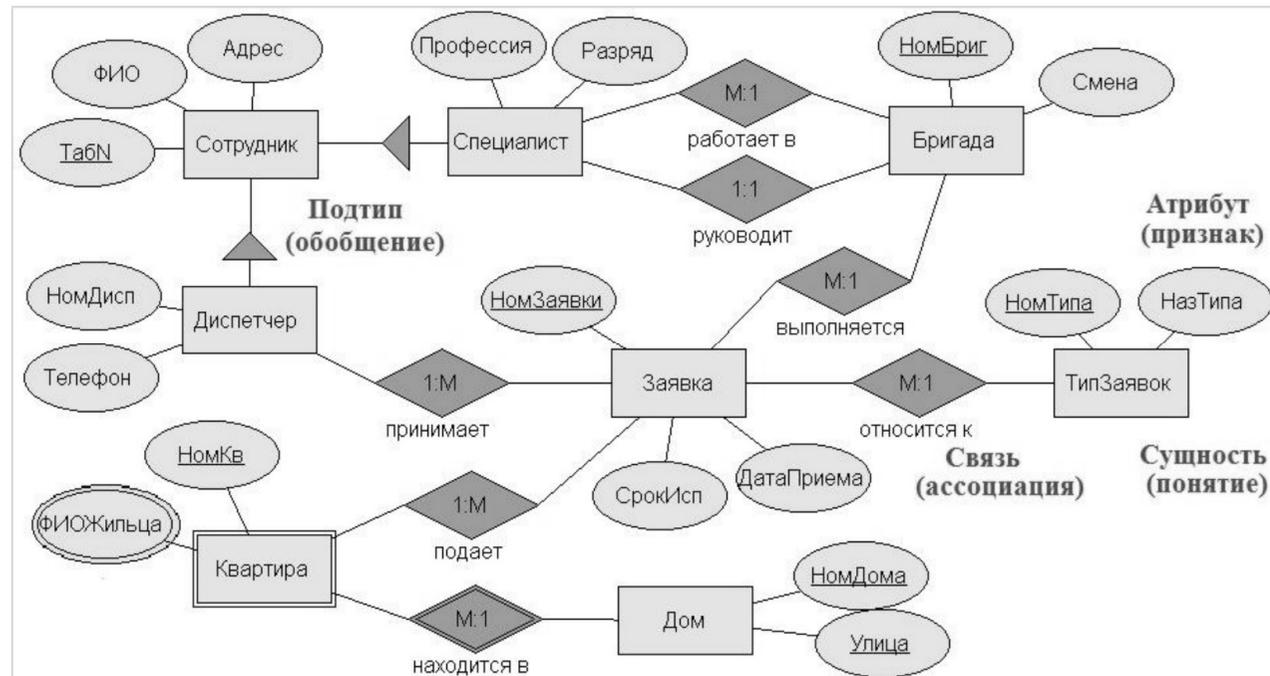
## Семантическая сеть



## Концептуальный граф



## Модель «сущность-связь»



## Фреймовая модель

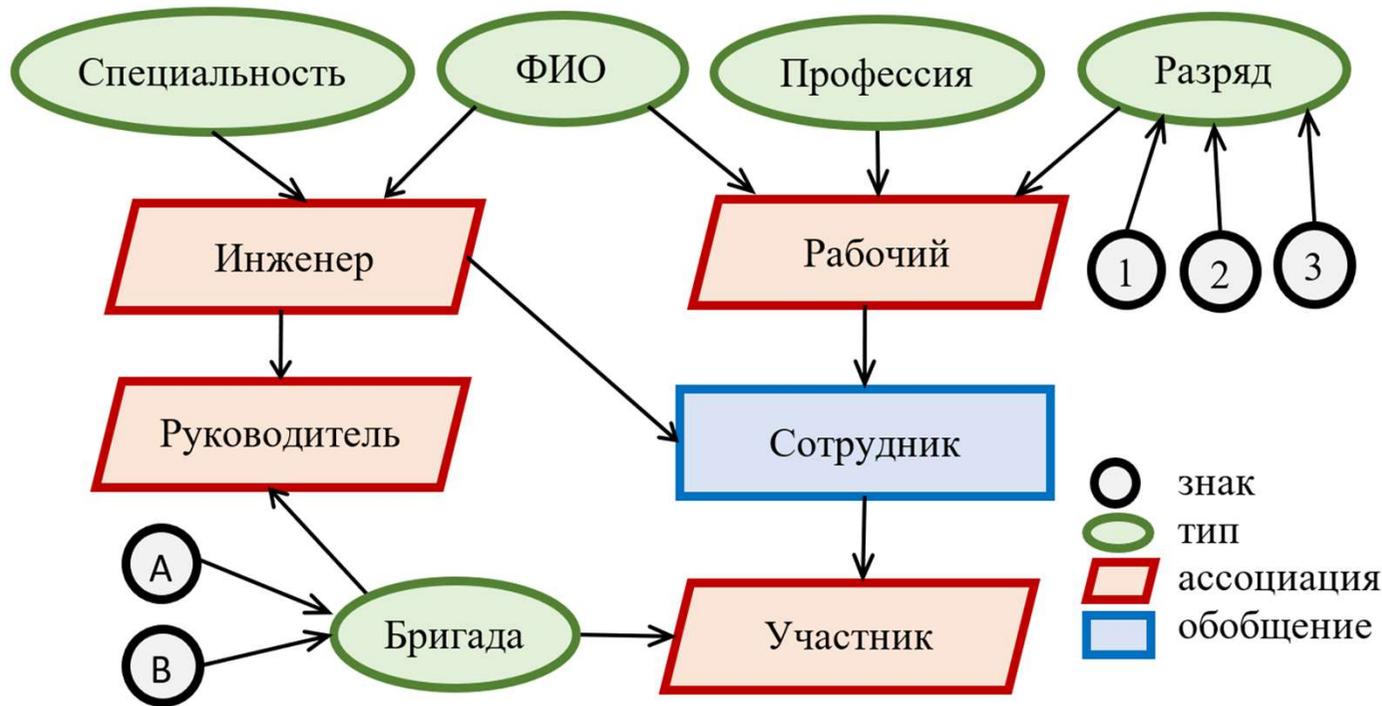
РУКОВОДИТЕЛЬ		
Имя слота	Значение слота	Тип значения слота
Имя	Иванов И. И.	Строка символов
Рожден	01.01.1965	Дата
Возраст	age(Рожден)	Процедура
Специальность	Юрист	Строка символов
Отдел	Отдел кадров	Строка символов
Зарплата	80000	Число
Адрес	ДОМ_АДРЕС	Фрейм

# Недостатки сетевых моделей

- Сложность модификации и актуализации.
- Слабая выразительность по сравнению с логическими моделями.
- Большое число отношений, результат запроса – фрагмент сети (поиск изоморфного подграфа – NP-полная задача).
- Сложность вывода, необходимость использования эвристик.
- Вывод на сетевых моделях не может гарантировать достоверность результата, не имеет логического обоснования.
- Разметка дуг задает предикаты (понятия), однако использование таких понятий требует формализации в логике второго порядка.
- Нет средств для представления процедурных знаний.
- Обучение модели только с учителем.

# Понятийная модель

- **Понятийная структура** задает способы образования понятий.



- **Узлы:**
  - понятие-знак;
  - понятие-тип;
  - понятие-ассоциация;
  - понятие-обобщение.
- **Связи** понятия – указание понятий, использованных для его образования.

- **Понятийная модель** – понятийная структура и описание сущностей понятий предметной области (задание перечислимых или разрешимых множеств сущностей).

# Конкретные понятия

○ **Понятие-знак** (элементарное понятие) образуется при мысленном *выделении* в предметной области уникального представления (целостного образа) и присвоения ему имени (абстракция *идентификации*).

**Примеры:** Зеленый, Много, Красиво, Всегда, Движение, Любовь.

**В информационных системах:** 'с', 12, 1.3e-12, 18.06.2019 15:45, "abc" (значения простых типов данных).

○ **Понятие-тип** (простое понятие) образуется путем *объединения* элементарных понятий, сходных в некотором смысле (абстракция *типизации*).

**Примеры:** Цвет (Красный, Зеленый, ...), Целое число (1, 2, 3, ...).

**В информационных системах:** Символ, Целое, Число с плавающей запятой, Дата и время, Строка (простые типы данных).

# Абстрактные понятия

 **Понятие-ассоциация** образуется путем *соединения* нескольких понятий в одно составное (абстракция *ассоциации*).

**Пример:** понятие-ассоциация Погода как соединение понятий Место, Дата, Температура, Влажность, Ветер, Облачность и т.п.

**В информационных системах:** массив структур, таблица базы данных.

 **Понятие-обобщение** образуются путем *объединения* нескольких понятий в одно общее (абстракция *обобщения*).

**Пример:** Понятие-обобщение Фрукт как объединение понятий Яблоко, Груша, Персик, Абрикос и др.

**В информационных системах:** массив базовых объектов, union-запрос к базе данных.

# Содержание понятий

**Понятие-знак** – значение встроенного типа данных.

## Понятие-тип

Entity	Aspect
65000000000	
65000000001	География
65000000002	Проекты
65000000003	Шаблоны
65000000004	Опросы

## Понятие-ассоциация

Entity	District	Capital
72000000000		
72000000001	Центральный федеральный округ	г. Москва
72000000002	Южный федеральный округ	г. Ростов-на-Дону
72000000003	Северо-Западный федеральный округ	г. Санкт-Петербург
72000000004	Дальневосточный федеральный округ	г. Владивосток
72000000005	Сибирский федеральный округ	г. Новосибирск
72000000006	Уральский федеральный округ	г. Екатеринбург
72000000007	Приволжский федеральный округ	г. Нижний Новгород
72000000008	Северо-Кавказский федеральный округ	г. Пятигорск

## Понятие-обобщение

Entity	Title	Capital
72000000000		
72000000001	Центральный федеральный округ	г. Москва
72000000002	Южный федеральный округ	г. Ростов-на-Дону
72000000003	Северо-Западный федеральный округ	г. Санкт-Петербург
72000000004	Дальневосточный федеральный округ	г. Владивосток
72000000005	Сибирский федеральный округ	г. Новосибирск
72000000006	Уральский федеральный округ	г. Екатеринбург
72000000007	Приволжский федеральный округ	г. Нижний Новгород
72000000008	Северо-Кавказский федеральный округ	г. Пятигорск
73000000001	Амурская область	г. Благовещенск
73000000002	Еврейская автономная область	г. Биробиджан
73000000003	Камчатский край	г. Петропавловск-Камчатский
73000000004	Магаданская область	г. Магадан
73000000005	Приморский край	г. Владивосток
73000000006	Республика Саха (Якутия)	г. Якутск
73000000007	Сахалинская область	г. Южно-Сахалинск
73000000008	Хабаровский край	г. Хабаровск
73000000009	Чукотский автономный округ	г. Анадырь

# ПОНЯТИЙНЫЙ ЯЗЫК

- **Факты** - истинные суждения с логическими связями И ( $\wedge$ ), ИЛИ ( $\vee$ ), НЕ ( $\neg$ ) и двумя видами элементарных выражений:
  - предикатами принадлежности сущности  $E$  понятию  $N$  вида  $N(E)$ ;
  - отношениями вида  $E[N] \circ V$ , где  $E[N]$  - функтор, возвращающий сущность понятия  $N$ , входящую в сущность  $E$ ,  $\circ$  - знак отношения ( $=$ ,  $<$ ,  $>$  и т.п.), некоторая сущность  $V$ .
- **Примеры:**
  - Разряд(2) – 2 есть сущность понятия-типа Разряд;
  - Рабочий( $X$ ) –  $X$  есть сущность понятия-ассоциации Рабочий;
  - $X[\text{Разряд}] = 3$  – Рабочий  $X$  имеет Разряд, равный 3;
  - Сотрудник( $Y$ ) –  $Y$  есть сущность понятия-обобщения Сотрудник.

**Понятийный язык** – одноместное исчисление предикатов (разрешим и непротиворечив на счетных и полон на конечных множествах)

## Правила вывода

$$a: N(E) \rightarrow \bigwedge_{\forall N_i \in H} E[N_i]; \quad g: N(E) \leftrightarrow \bigvee_{\forall N_i \in H} N_i(E),$$

где  $a$  – правило вывода из понятия-ассоциации  $N$ ,  $g$  – правило вывода из понятия-обобщения  $N$ ,  $H$  - схема понятия  $N$ .

**Схема** понятия – набор абстрагируемых понятий (набор понятий, из которых образовано определяемое понятие).

- **Примеры** правил вывода:

- $a: \text{Инженер}(X) \rightarrow X[\text{Специальность}] \wedge X[\text{ФИО}];$
- $a: \text{Рабочий}(X) \rightarrow X[\text{ФИО}] \wedge X[\text{Профессия}] \wedge X[\text{Разряд}];$
- $g: \text{Сотрудник}(X) \leftrightarrow \text{Инженер}(X) \vee \text{Рабочий}(X);$
- $a: \text{Участник}(X) \rightarrow X[\text{Сотрудник}] \wedge X[\text{Бригада}].$

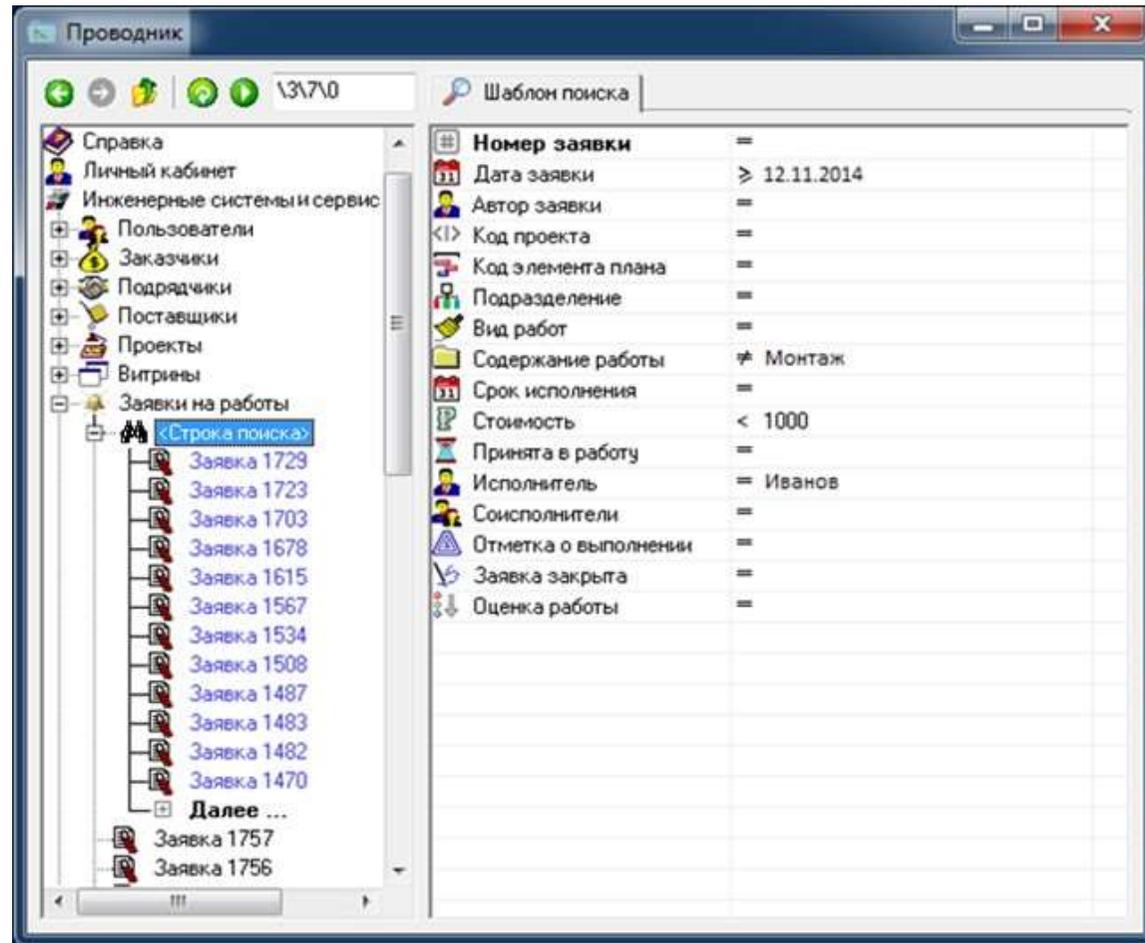
# Язык запросов

- **Вопрос:** «Кто из участников бригады В не является рабочим с третьим разрядом?»

- **Запрос:**

Участник( $X$ )  $\wedge$   $X$ [Бригада] = В  $\wedge$   
 $\wedge$   $\forall Y$  ( $X$ [Сотрудник] =  $Y$   $\wedge$   
 $\wedge$  Рабочий( $Y$ )  $\wedge$   
 $\wedge$   $\neg Y$ [Разряд] = 3).

- **Время:** поиск сущности в таблице понятия с  $n$  сущностями имеет асимптотику  $\ln n$ .



# Представление процедурных знаний

**Процедурные знания** представляются в виде процедур и функций, которые, являются сущностями соответствующих понятий, а само понятие – сущностью понятия понятий.

**Процедура** повышения разряда рабочим бригады В.

```
// Определение функции повышения разряда
Повышение = (Разряд) => Разряд < 3 ? 1 : 0;
// Создание индексной переменной
Индекс = 0;
// Открытие транзакции с именем Разряды
begin Разряды;
// Запрос сущностей локального понятия Рабочие
Рабочие = (X)(Участник(X) != ' ' && X[Бригада] == ' В '
&& Рабочий(X[Сотрудник]) != ' ');
```

```
// Цикл по сущностям локального понятия Рабочие
while (Индекс < Рабочие{})
{
    // Создание локальной сущности Рабочий
    Рабочий = Рабочие{Индекс};
    // Увеличение атрибута Разряд
    Рабочий[Разряд] += Повышение(Рабочий[Разряд]);
    // Приращение индекса сущности понятия Рабочие
    Индекс = Индекс + 1;
}
// Обновление измененных сущностей
()Рабочие;
// Успешное завершение транзакции
commit Разряды;
```

# Обучение понятийной модели

- **Обучение без учителя.** Структуризации потока входных понятий-знаков и их конкатенаций на основе текущей модели, выделение сущностей известных и неизвестных понятий, пополнение существующих понятий новыми сущностями или образования новых понятий.
- **Обучение с подкреплением.** Подкрепление ранее созданных сущностей и понятий или их «забывание» при редком их повторении во время накопления «жизненного опыта».
- **Обучение с учителем.** Создание, изменение и удаление понятий и принадлежащих им сущностей непосредственно пользователем.

Иванов И.И.\*Технология машиностроения | Петров П.П.\*А | Сидоров С.С.\*Сварщик\*4 | Новиков В.В.\*Программное обеспечение

# Машина обладает знаниями

- **Сужение** области знания путем использования понятия в заданном аспекте или множестве аспектов.
- **Владение** единичными сущностями предметной области в виде понятий-знаков и их соединений (конкатенаций).
- **Определение** имеющихся понятий и создание новых путем использования операций идентификации, ассоциации и обобщения.
- **Оперирование** суждениями о сущностях и понятиях в виде логических высказываний одноместного исчисления предикатов.
- **Выполнение** умозаключений по правилам дедукции, абдукции и индукции для достоверного и правдоподобного вывода.
- **Получение** скрытого знания путем рассуждений, где правила вывода на знаниях задаются схемами понятий.
- **Представление** процедурных знаний и **обучение** без учителя.

# Заключение

- **Трудности** представления и обработки знаний в логических, сетевых и продукционных моделях. Наилучшая модель – понятийная.
- **Понятийное мышление** – предельно общий и единственный механизм рационального познания, а понятийная модель – адекватная форма фиксации результатов понятийного мышления.
- **Семантическая инвариантность** понятийной модели – для ее интерпретации не надо знать семантику предикатов: все, что можно узнать о предметной области, содержится в самой модели.
- **Высокая эффективность** обработки знаний в формализме понятийной модели: логарифмическая асимптотика.
- **Сильный искусственный интеллект** не сводится к умению мыслить и обладанию знаниями. Необходимо использование не только рациональных, но и иррациональных формы познания, а также эмоций.

# Литература

- 1. Baral C. Knowledge Representation, Reasoning and Declarative Problem Solving. – Cambridge: Cambridge University Press, 2004. – 541 p.
- 2. The description Logic Handbook: Theory, Implementation and Applications / Eds. F. Baader, D. Calvanese, D. L. McGuinness, et al. – New York: Cambridge University Press, 2003. – 573 p.
- 3. Chein M., Mugnier M.-L. Graph-based Knowledge Representation: Computational Foundations of Conceptual Graphs. – Springer, 2009. – 425 p.
- 4. Klahr D., Langley P., Neches R. Production System Models of Learning and Development. Cambridge. – Cambridge: The MIT Press, 1987. – 466 p.
- 5. Выхованец В.С. Понятийный анализ и понятийное моделирование // Управление большими системами. Вып. 92. – М.: ИПУ РАН, 2021. – С. 64-109. – <https://doi.org/10.25728/ubs.2021.92.4>.