



# AI-технологии NLU и Ontological Semantics

в медицинских  
экспертных системах

*Н.Л. Сомс, AIIRE  
А.В. Добров, AIIRE*

# Экспертные системы и СППР

Экспертные системы в узком смысле (MYCIN, CLIPS, OpenCyc, IBM Watson):

- Rule-based
- структурированные данные
- продукционная модель
- forward chaining
- полная автоматизация
- цель -- заменить эксперта

Системы поддержки принятия решений (частично IBM Watson, Aimedica, AIIRE):

- слабоструктурированные или неструктурированные данные
- может применяться AI (ИСППР)
- частичная автоматизация
- цель -- помочь эксперту



# Проблемы

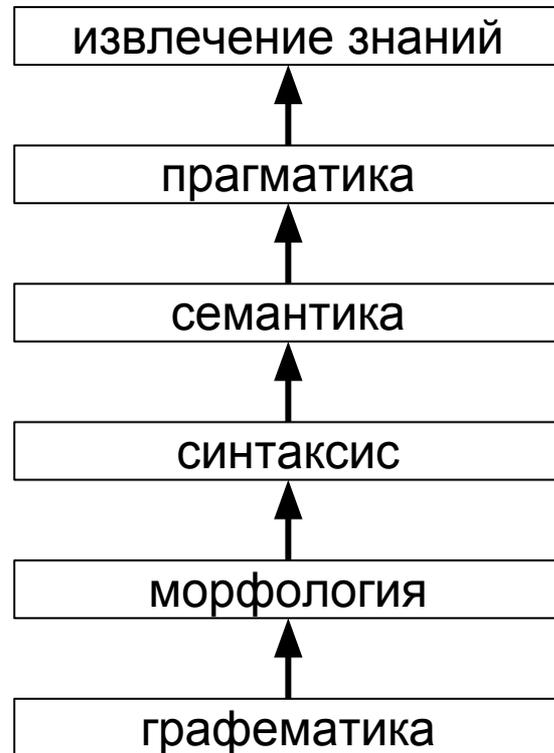
- **Диагностика**
  - Неопределённость информации диагностики, недостоверность наличия симптома, субъективность оценивания специфического симптома
  - Неоднозначность симптомов (о чём свидетельствует сухой кашель?)
  - Неоднозначность естественных языковых описаний (что конкретно означает *"Кружится голова"*?)
  - Противоречивость симптомов при комплексе заболеваний
- **Терапия**
  - Противоречивость назначений при комплексе заболеваний (пример: пациенты с диабетической нефропатией сталкиваются с практически взаимоисключающими диетами)



# Технологии NLU

Автоматическое понимание текстов (Natural Language Understanding, NLU) -- вид автоматической обработки текстов, при котором машина выполняет полную цепочку процедур анализа текста на всех его уровнях, ведущую к интеграции его содержания в используемую базу знаний / онтологию

Технологии NLU успешно выполняют разрешение неоднозначности (disambiguation) и работают именно в условиях неопределённости



# Технологии NLU: графематика

Цель графематического анализа -- выделение атомарных единиц (токенов: словоформ / морфем / знаков пунктуации / чисел и т.д.) в потоке символов. AIIE: wcsplit, Aho-Corrasick

К графематическому анализу также относят выделение границ высказываний, что в общем случае некорректно.

Неоднозначность токенизации наиболее заметна в языках, не имеющих пробелов, но наблюдается также и в особых случаях, когда требуется морфемный анализ или анализ сокращений: *Имидазолидинил*: имидазол + идин + ил / имидазол + и + динил / ...



# Технологии NLU: морфология

Цель морфологического анализа -- определение грамматических признаков и словарных форм атомарных единиц, выделенных при графематическом анализе. AIIRE: морфологические словари, морфосинтаксис

Морфологическая неоднозначность:

*железа* -- желéзо, n gen sg / железá, n nom sg

*отёк* -- отёк, n nom sg / отёк, n acc sg / отечь, v perf m sg

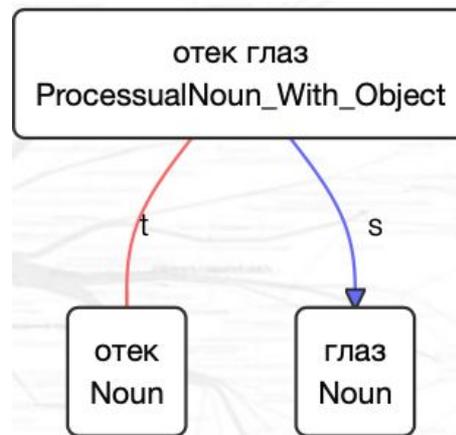
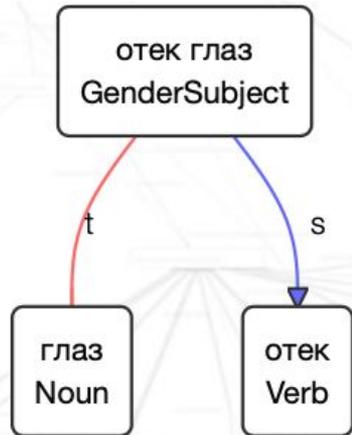
*потом* -- пот, n nom ins / потóм, adv



# Технологии NLU: синтаксис

Цель синтаксического анализа -- определение синтаксической структуры словосочетания, предложения, текста, даже химической формулы. AIIRE: формальная объектно-ориентированная грамматика

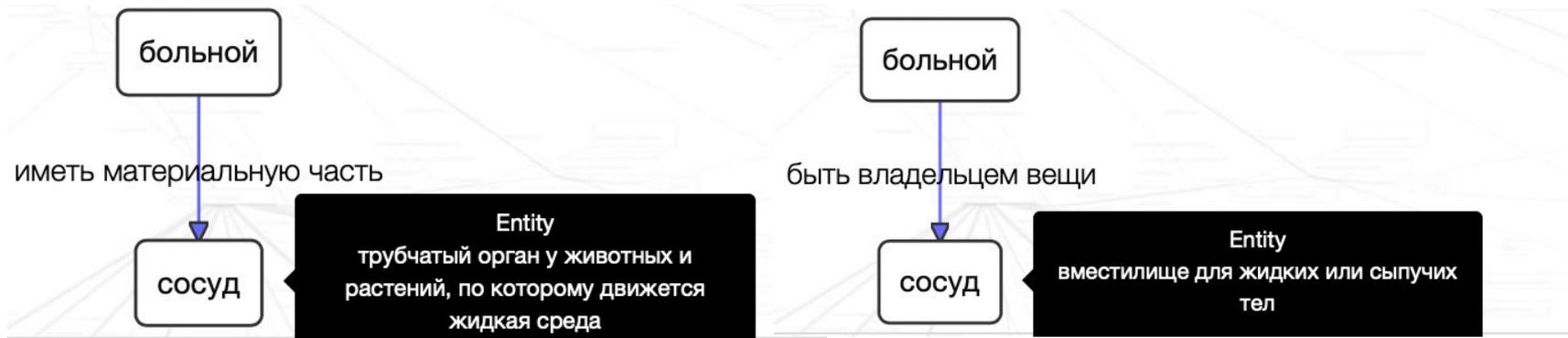
Синтаксическая неоднозначность может возникать в результате морфологической или безотносительно к ней, ср: *отёк глаз*



# Технологии NLU: семантика

Цель семантического анализа -- вычисление значения (смыслового содержания) языковой единицы. AIIRE -- онтологическая семантика.

С точки зрения семантики, неоднозначность может быть лексической или обусловленной неопределённостью связей, ср: *сосуды пациента*



# Технологии NLU: прагматика

Цель прагматического анализа -- устранение неопределённостей семантики путём соотнесения значения предложения с контекстом и прагматической рамкой.

Ср. *Пациент поднял руку* -- с точки зрения семантики, рука могла как принадлежать, так и не принадлежать данному пациенту; с точки зрения прагматики, при отсутствии иного контекста, рука точно принадлежала пациенту

Ср. *Вылечил их семью своими руками* -- с точки зрения семантики, но не прагматики, у врача теоретически могло быть семь рук!



# Технологии NLU: извлечение знаний

Data mining, извлечение ситуаций, fact extraction в случае применения пайплайна NLU -- выделение семантических графов, соответствующих заданным шаблонам. AIIR: AiFactEval.

Примеры выделения ситуаций по шаблону “гипергликемия”:

*Глюкоза в сыворотке была 10.5*

*Сахар в крови колебался между 11 и 17*

*Повышенный сахар натощак*

*Уровень глюкозы подскочил до небес*

глюкоза = 10.5 ммоль/л

глюкоза = 11.0 ммоль/л

глюкоза > 5.5 ммоль/л

глюкоза >> 5.5 ммоль/л



# Онтологическая семантика

Компьютерная онтология -- база знаний, содержащая концепты и отношения между ними

Лингвистическая онтология -- модель функционирования системы понятий, стоящих за значениями языковых единиц

Онтология предметной области -- модель функционирования системы понятий предметной области

Онтологическая семантика -- моделирование значения естественного языкового выражения средствами компьютерной онтологии



✓ сосуд (трубчатый орган у животных и растений, по которому движется жидкая среда)

✓ орган (часть организма, выполняющая определенную функцию)

✓ плоть (плотная часть живого тела)

✓ часть организма (#)

✓ часть материального объекта (#)

✓ материальное тело (трехмерное тело, состоящее из материи, обладающее массой и объемом, существующее в пространстве и во времени)

✓ объект реального мира (любой объект, существующий в реальном мире)

✓ объект (понятие о сущности, существующей в объективной или вымышленной реальности)

✓ объект или процесс (что-либо, являющееся объектом мышления - процессом или сущностью)

✓ объект, свойство или процесс (#)

✓ понятие (отображённое в мышлении единство существенных свойств, связей и отношений предметов или явлений)

✓ изменяемое тело (непустой измеряемый изменяемый объект)

✓ тело (непустой измеряемый объект)

✓ n-мерный объект (#)

✓ объект (понятие о сущности, существующей в объективной или вымышленной реальности)

✓ объект или процесс (что-либо, являющееся объектом мышления - процессом или сущностью)

✓ объект, свойство или процесс (#)

✓ понятие (отображённое в мышлении единство существенных свойств, связей и отношений предметов или явлений)

✓ изменяемый измеряемый объект (локализуемый в каком-либо пространстве и имеющий измерения изменяемый объект)

✓ изменяемый объект (объект, который можно изменить, имеющий разные состояния)

✓ изменяемое понятие (понятие, свойства которого могут меняться)

1. **расстройство здоровья живого организма, нарушение его деятельности**



2С



базовый

Asia Dobrova (утверждено)

Передать

ЧР: существительное; лемма: болезнь [Entity]

Атрибуты

Предки

Связи

транзакции(ряд)

транзакции(связи)

**синонимы**

1. [заболевание \( расстройство здоровья живого организма, нарушение ... \)](#) (Википедия)
2. [недуг \( расстройство здоровья живого организма, нарушение ... \)](#) (Википедия)
3. [недомогание \( расстройство здоровья человека, болезненное ... \)](#) (Викисловарь)
4. [хворь \( расстройство здоровья живого организма, нарушение ... \)](#) (Викисловарь)
5. [состояние нездоровья \( расстройство здоровья живого организма, нарушение ... \)](#) (Шведова)

**гиперонимы**

1. [ненаправленное неадресованное состояние живого существа \(#\)](#) (AIIRE) — гиперонимы: 9

**гипонимы**

1. [вирусное заболевание \(#\)](#) (AIIRE) — гипонимы: 1
2. [инфекция \( проникновение в организм болезнетворных ... \)](#) (Викисловарь) — гипонимы: 83
3. [кома \( остро развивающееся тяжёлое патологическое ... \)](#) (Википедия)
4. [морок \( устар., перен. что-н. одуряющее, очаровывающее, ... \)](#) (Викисловарь) — гипонимы: 1
5. [психическое заболевание \( состояние психики, отличное от нормального, ... \)](#) (Википедия) — гипонимы: 6
6. [соматическое заболевание \( телесное заболевание \)](#) (Википедия) — гипонимы: 2

**объекты отношения** [соответствие процесса действию \(#\)](#)

1. [болеть \( страдать болезнью, хворать, быть больным \)](#) (Баранов)

**предметные области понятия**

1. [медицина \( область практической деятельности, связанная с ... \)](#) (Википедия) — гипонимы: 11



✓ грипп (мед. острое вирусное заболевание, характеризующееся воспалением дыхательных путей и лихорадочным состоянием)

✓ вирусное заболевание (#)

✓ заболевание (расстройство здоровья живого организма, нарушение его деятельности)

✓ ненаправленное неадресованное состояние живого существа (#)

✓ ненаправленное неадресованное состояние (#)

✓ ненаправленное состояние (#)

✓ состояние (процесс, являющийся состоянием, не предполагающий конкретного обязательного итога, являющийся характеристикой объекта, основанной на совокупности обстоятельств его существования)

✓ действие или состояние (#)

✓ процесс (последовательность стадий или состояний чего-либо в ходе каких-либо действий)

✓ физический объект или процесс (#)

✓ объект или процесс (что-либо, являющееся объектом мышления - процессом или сущностью)

✓ объект, свойство или процесс (#)

✓ понятие (отображённое в мышлении единство существенных свойств, связей и отношений предметов или явлений)

✓ нематериальный объект или процесс (#)

✓ объект или процесс (что-либо, являющееся объектом мышления - процессом или сущностью)

✓ объект, свойство или процесс (#)

✓ понятие (отображённое в мышлении единство существенных свойств, связей и отношений предметов или явлений)

✓ фазовая ситуация (ситуация, развивающаяся во времени или являющаяся фазой процесса)

✓ ситуация (совокупность объектов, связанных некоторыми отношениями)

✓ объект или процесс (что-либо, являющееся объектом мышления - процессом или сущностью)

✓ объект, свойство или процесс (#)

✓ понятие (отображённое в мышлении единство существенных свойств, связей и

# Диалоговый интерфейс

Чат-боты (онлайн-консультанты) приобрели высокую популярность, способствующую развитию экспертных систем и интеллектуальных систем поддержки принятия решений. Возможный диалог с использованием AIIRE / AiFactEval при соответствующей проработке онтологии:

*Пользователь: Добрый день ! На фоне беременности применяла лозап 50 мг 1 раз утром, отменила только на сроке в 21 день беременности ( беременность ЭКО). очень переживаю за последствия, как это может отразиться на детях. будет двойня.*

*Чат-бот: Уточните, пожалуйста, каково ваше давление?*

*Пользователь: 156/121*



# Диалоговый интерфейс

*Чат-бот: Наблюдается ли у Вас гестационный диабет?*

*Пользователь: Да, но сахар невысокий — 6-7 натощак. Сладкое не ем.*

*Чат-бот: Фиксируются ли проблемы с почками?*

*Пользователь: А как это определить?*

*Чат-бот: Посмотрите последний анализ суточной мочи — каков диурез, каково содержание белка? Посмотрите также анализ крови — каково содержание креатинина, калия, мочевины?*

*Пользователь: Я не делала эти анализы. Надо сделать?*



# Диалоговый интерфейс

*Чат-бот: Да, чем скорее, тем лучше.*

*Пользователь: Это снова я. Суточный диурез 700мл, белок 0.952, креатинин 247, калий 5.9, мочевины 22*

*Чат-бот: Мы рекомендуем Вам НЕМЕДЛЕННУЮ госпитализацию!!! Ваши показатели могут говорить о тяжёлой почечной недостаточности; на фоне артериальной гипертензии и диабета её прогрессирование может быть ускоренным; кроме того, возможна эклампсия. Лозартан отменять не рекомендуем, но в Вашей ситуации необходима срочная госпитализация и консультация с нефрологом.*



Спасибо!

