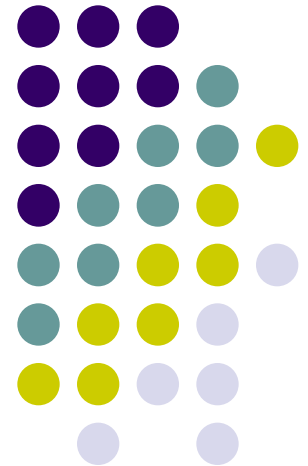


**ПРЕДСТАВЛЕНИЕ
ЗНАНИЙ:**

**ДЕСКРИПТИВНЫЕ
ЛОГИКИ**

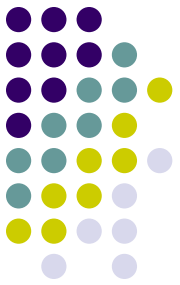


СОДЕРЖАНИЕ



1. Логическая модель и *дескриптивные* (*дескрипционные*) логики
2. Дескриптивная логика $\mathcal{ALC}$
 - Концепты: синтаксис и семантика
 - База знаний: TBox и ABox
 - Задачи логического вывода в БЗ
3. Расширения логики $\mathcal{ALC}$
4. Сравнение вывода в БЗ и БД
5. Заключение, Домашнее задание

ИСТОКИ ДЕСКРИПТИВНЫХ ЛОГИК



- Развитие логической модели, с учетом
 - средств ПЗ, разработанных в рамках других моделей
 - потребностей формального ПЗ
(необходимо для построения онтологий)
- Логика *ALC* (*Attributive Language with Complements*), 1991 г. – первая и одна из базовых дескриптивных логик (ДЛ), на основе которой строятся многие другие
- Ныне на базе разных ДЛ построены **онтологии** в самых различных предметных областях: химия, биология, медицина, биоинформатика и др.
- Реализованы программные системы – механизмы рассуждений, вывода знаний (*reasoners*)

ОСОБЕННОСТИ ДЕСКРИПТИВНЫХ ЛОГИК



- Особенности логической модели ПЗ:
 - знания представляются как совокупность формул
 - операции над знаниями – на базе логического вывода
- Ключевые понятия ДЛ: *концепт* и *роль* соответствуют:
 - ❖ Концепт (понятие) – это множество, класс (объектов), одноместный предикат: *Студент, Автомобиль*
 - ❖ Роль – это свойство, бинарное отношение, предикат: *Учиться (кто, где), Родитель (кто, кого)*
- Выражения языка – *составные концепты* строятся из *атомарных концептов* и *атомарных ролей*, при помощи набора *конструкторов*
- ДЛ позволяют формулировать утверждения не только о классах, но также о конкретных объектах – *индивидах* (экземплярах классов)

КОНСТРУКТОРЫ КОНЦЕПТОВ



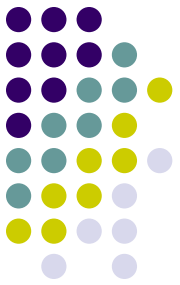
Каждая ДЛ: определенный набор *конструкторов* для построения составных концептов

Стандартные конструкторы:

- *Пересечение* (конъюнкция) концептов: $C \sqcap B$
- *Объединение* (дизъюнкция) концептов: $C \sqcup B$
- *Дополнение* (отрицание) концепта: $\neg C$
- Ограничение на значения роли
квантором всеобщности: $\forall R.C$
- *Экзистенциальное ограничение*
(ограничение квантором существования): $\exists R.C$
где C и B – концепты, а R – роль

Конкретная ДЛ характеризуется *синтаксическими правилами* для построения составных концептов

$\mathcal{LC}$: СИНТАКСИС КОНЦЕПТОВ



Элементы формул-концептов:

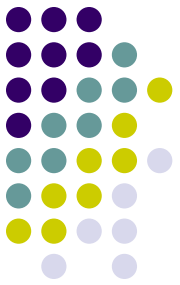
- множество символов (имен) *атомарных концептов*
- множество символов (имен) *атомарных ролей*
(эти множества не пустые, конечные)
- кванторы, служебные символы (точки, скобки)

Синтаксические правила для концептов:

- Всякий *атомарный* концепт A является концептом
- Выражения $\top$ и $\perp$ являются концептами (*истина* и *ложь*)
- Если C есть концепт, то его *дополнение* $\neg C$ – концепт
- Если C и B – концепты, то $C \sqcap B$ и $C \sqcup B$
(*пересечение* и *объединение*) являются концептами
- Если C – концепт, а R – атомарная роль, то выражения
 $\exists R.C$ и $\forall R.C$ являются концептами

Никакие другие выражения не являются концептами

АЛС : СЕМАНТИКА, ПРИМЕРЫ

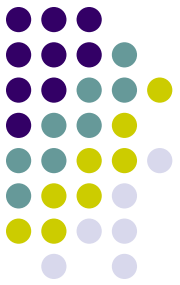


Интерпретация I состоит из непустой области D и функции, интерпретирующей атомарные концепты и роли, а также составные концепты

Примеры: Атомарные концепты $Animal$, Cat , Dog , $Human$, $Male$, $Female$, $Student$; Роли has и $hasChild$ («естественная» интерпретация)

- $Student \sqcap Female$ – множество студенток
- $\forall hasChild . Male$ – все особи, чьи дети мужского пола
- $Human \sqcap \forall hasChild . \neg Student$ – люди, все дети которых не являются студентами
- $Male \sqcap \exists has . ((Cat \sqcup Dog) \sqcap \neg Female)$ – мужчины ...?
- ❖ Если пара (x, z) объектов-индивидов принадлежит отношению-роли R , т.е. верно $x R z$, то говорят, что индивид z является R -последователем индивида x

$\mathcal{ALC}$: СЕМАНТИКА КОНЦЕПТОВ



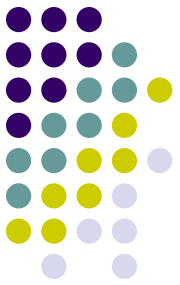
Интерпретация I на области D
(формальная семантика)

- Каждому атомарному концепту C ставится в соответствие подмножество из области D
- Каждой атомарной роли – бинарное отношение в D
- Выражение $\top$ интерпретируется как вся область D
- Выражение $\perp$ – как пустое множество $\emptyset$

Далее индукция по построению концепта (рекурсивно):

- Дополнение $\neg C$ интерпретируется как $D \setminus C$
- Пересечение $C \sqcap V$ и объединение $C \sqcup V$ концептов – как пересечение и объединение соответствующих множеств, т.е. $C \cap V$ и $C \cup V$

СЕМАНТИКА КВАНТОРНЫХ ОГРАНИЧЕНИЙ



- Выражение $\forall R.C$ интерпретируется как множество тех объектов-индивидов, у которых все R -последователи принадлежат интерпретации концепта C . Формально:
$$(\forall R.C) = \{e \in D \mid \forall d \in D : (e, d) \in R \Rightarrow d \in C\}$$
 - Выражение $\exists R.C$ интерпретируется как множество тех объектов индивидов, у которых имеется R -последователь, принадлежащий интерпретации концепта C . Формально:
$$(\exists R.C) = \{e \in D \mid \exists d \in D : (e, d) \in R \wedge d \in C\}$$
- В частности: $\forall R.T \equiv T$ и $\exists R.\perp \equiv \perp$

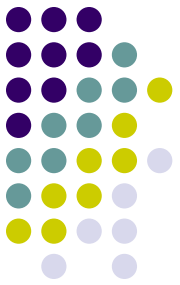
АЛС И ЛОГИКА ПРЕДИКАТОВ



Логику АЛС (и многие ДЛ) можно рассматривать как фрагменты логики предикатов при естественном переводе концептов в предикатные формулы

- Для атомарных концептов C и атомарных ролей R вводятся соответственно одноместные $C(x)$ и двухместные $R(x, y)$ предикатные символы
 - Дополнение, пересечение и объединение концептов переходит в булевы связки: $\neg$, $\wedge$, $\vee$
 - Выражение $\forall R.C$ переходит в $\forall y (R(x, y) \rightarrow C(y))$
 - Выражение $\exists R.C$ переходит в $\exists y (R(x, y) \wedge C(y))$
- (переменная y – новая, а C – перевод соответств. концепта)
- ❖ Составной концепт – формула с одной своб. переменной, например, для $\text{Human} \sqcap \forall \text{hasChild} . \neg \text{Student}$
- формула $\text{Human}(x) \wedge (\forall y (\text{hasChild}(x, y) \rightarrow \neg \text{Student}(y)))$

СВЯЗЬ $\mathcal{ALC}$ с ЛОГИКОЙ $\mathcal{L}^2$



- Перевод в предикаты согласуется семантикой ДЛ, т.е. при любой I , если атомарные концепты и роли интерпретированы так же, как и соответствующие им предикаты, то и любой составной концепт интерпретируется тем же самым множеством, что и соответствующая ему предикатная формула
- Не всякая формула логики предикатов является переводом какого-либо концепта; например, формула $\forall x P(x)$
- При переводе могут получиться лишь формулы, в которых квантор навешивается на второй аргумент предиката
- $\mathcal{ALC}$ погружена в логику $\mathcal{L}^2$ предикатов с двумя переменными, которая является разрешимой
- Важно: перевод позволяет переносить результаты о разрешимости, вычислительной сложности и т. п. из области логики предикатов в область ДЛ

КОМПОНЕНТЫ *АЛС* и БЗ



Знания записываются в виде аксиом,
все множество аксиом разбивается на:

- **TBox** (*terminological box*) — набор утверждений общего вида — терминология, общие знания о понятиях (концептах) и их взаимосвязях (ролях), интенциональные знания
- **ABox** (*assertional box*) — набор утверждений частного вида, знания об объектах-индивидах, их свойствах и связях с другими объектами, экстенциональные знания

Первые более стабильны и постоянны, тогда как
вторые более подвержены модификациям

База знаний= **TBox** + **ABox**

КОМПОНЕНТА TBox



Терминологические аксиомы:

- *Вложенность концептов*: выражение вида $C \sqsubseteq B$
- *Эквивалентность концептов*: выражение $C \equiv B$
где C и B – произвольные концепты,
а $\sqsubseteq$ – символ вложенности (*subsumption*), на основе вложенности строится *иерархия концептов*

Возможны также аксиомы (в расширениях $\mathcal{ALC}$):

- *Вложенность ролей*: $R \sqsubseteq S$
- *Эквивалентность ролей*: $R \equiv S$
где R и S – произвольные роли.
- ❖ *Терминология* – конечный набор терминологических аксиом перечисленных видов
- ❖ Семантика терминологии определяется естественным образом

СЕМАНТИКА ТЕРМИНОЛОГИИ



Пусть дана интерпретация I

- Аксиома $C \sqsubseteq B$ истинна в интерпретации I , если $C \subseteq B$, при этом I называется *моделью* этой аксиомы
- Аксиома $C \equiv B$ истинна в I , если $C = B$ (множества равны)
- Концепт C *выполним*, если существует интерпретация I , в которой множество $C \neq \emptyset$ (I называется *моделью* C)

Для заданной терминологии T :

- Терминология является *выполнимой* в интерпретации I , если I является моделью всех входящих в нее аксиом (I является *моделью* терминологии)
- Концепт C *выполним в терминологии* T , если существует модель этой терминологии, в которой множество $C \neq \emptyset$

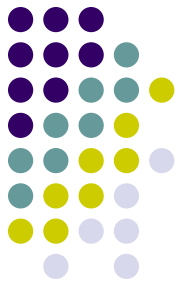
Важно: рассмотрение терминологий не выводит нас за рамки логики предикатов и логики $\mathcal{L}^2$

БЗ: ТЕРМИНОЛОГИЧЕСКИЕ ЗАДАЧИ



Задачи логического анализа и вывода новых знаний для терминологической компоненты БЗ:

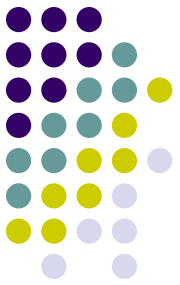
- Выполнимость атомарных концептов относительно терминологии (если этого нет, то ошибка ПЗ)
- Проверка (аксиом) вложенности концептов
- Совместимость аксиом терминологии: проверка того, что терминология вообще имеет хотя бы одну модель
- Вывод новых вложений для составных концептов
- Классификация терминологии, т.е. построение **иерархии (таксономии) концептов** – частично упорядоч. множества всех атомарных концептов относительно вложения
- Обнаружение эквивалентных концептов
- Проверка **ацикличности терминологии**
- ❖ В логиках, содержащих $\mathcal{ALC}$, проблема вложенности концептов сводится к выполнимости концепта



АЦИКЛИЧЕСКИЕ ТЕРМИНОЛОГИИ

- Аксиомы вида $C \equiv B$ можно считать определениями:
 $Woman \equiv Human \sqcap Female$
 $Mother \equiv Woman \sqcap \exists hasChild.T$
- Ацикличность – разумное требование:
 смысл новых, *производных* концептов, стоящих в левых частях равенств-определений, однозначным образом определяется через смысл *базовых* концептов, входящих в правую часть и не встречающихся в левых частях
- Пример нарушения ацикличности: $Mother \equiv Parent \sqcap Woman$
 $Parent \equiv Mother \sqcup Father$
- Вследствие нарушения возможны несколько интерпретаций этих концептов:
 - $Mother$ – множество матерей, $Parent$ – множество родителей
 - $Mother$ – множество всех женщин, $Parent$ – объединение множеств отцов и женщин
 - $Mother$ – пустое множество, $Parent$ – множество отцов

КОМПОНЕНТА ABox



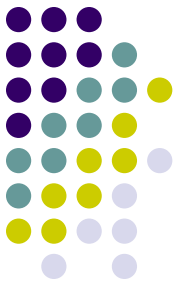
Содержит утверждения-факты.

Вводится множество *имён индивидов*: $a, b, c, \dots$

- Два вида утверждений об индивидах:
 - утверждение о принадлежности индивида a концепту C записывается как $C(a)$ или $a: C$
 - утверждение о связи двух индивидов a и b ролью R записывается как $R(a,b)$ или $(a,b): R$ или $a R b$
- Система фактов ABox – конечный набор таких утверждений
- Примеры:

$\text{Ann: Female} \sqcap \neg \text{Student}$
 Ann has Rex
 $\text{Rex: Dog} \sqcap \forall \text{hasChild}.\perp$

СЕМАНТИКА ABox



Расширение интерпретации I : каждому имени a индивида сопоставлен некоторый элемент из области D

- Утверждение-факт $C(a)$ или $R(a,b)$ выполняется в интерпретации I , если имеет место соответственно $a \in C$ или $(a, b) \in R$,
при этом I называется *моделью* этого факта
- ABox является *выполнимым* в интерпретации I , если в ней выполняются все факты набора (т.е. I - *модель* ABox)
- ABox называется *выполнимым* в терминологии T , если существует модель ABox, являющаяся одновременно моделью терминологии (т.е. общая *модель БЗ*)
- Индивид a является *экземпляром* концепта C в БЗ, если в любой модели БЗ имеет место $a \in C$

Обычно рассматриваются интерпретации, удовлетворяющие *соглашению об уникальности имён* (*unique name assumption*): разным именам индивидов сопоставлены различные элементы D (инъективное отображение)

ЗАДАЧИ по БЗ с ABox



Логические и алгоритмические задачи:

- Выполнимость базы знаний: имеет ли заданная пара (TBox, ABox) хотя бы одну модель?
- Принадлежность объекта-индивида заданному концепту БЗ
- Извлечение экземпляров: найти все экземпляры заданного концепта относительно заданной БЗ
- Классификация индивида: поиск для него наиболее узкого, т.е. наименьшего (по вложению) атомарного концепта
- Ответ на запрос к БЗ: выдать все наборы индивидов, удовлетворяющие заданным ограничениям на факты
- ❖ В настоящее время изучены т.н. *конъюнктивные запросы* к БЗ ДЛ (а также их дизъюнкции), которые похожи на аналогичные запросы из области БД
- ❖ Для запросов более общего вида задача быстро приобретает высокую вычислительную сложность или даже становится неразрешимой

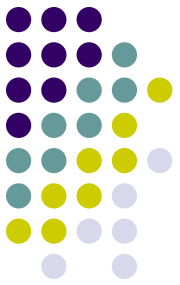
ЛОГИКА $\mathcal{ALC}$: ХАРАКТЕРИСТИКА



- Строго говоря, $\mathcal{ALC}$ – это не одна логика, а семейство дескриптивных логик, где каждая логика задается конкретным набором атомарных концептов и ролей, а также индивидов (сигнатурой)
- **Сигнатура логики** – конечные непустые множества
 - атомарных концептов $CN = \{C_1, C_2, \dots, C_n\}$
 - атомарных ролей $RN = \{R_1, R_2, \dots, R_m\}$
 - индивидных объектов $IN = \{I_1, I_2, \dots, I_k\}$
- Свойства логики $\mathcal{ALC}$:
 - ❖ Разрешимость: наличие алгоритма установления выполнимости концепта, совместимости БЗ (следует из вложимости в логику $\mathcal{L}^2$)
 - ❖ Вычислительная сложность: полиномиальная

В дескрипт. логиках вывод по компонентам $Tbox$ и $ABox$ может существенно различаться по производительности

РАСШИРЕНИЯ ЛОГИКИ $\mathcal{ALC}$

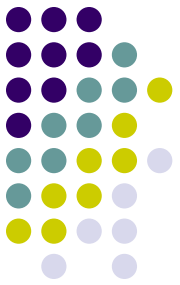


Многочисленные расширения логики $\mathcal{ALC}$ новыми выразительными возможностями:

- Ограничения на роли (численные, функциональные, симметричность, рефлексивность ролей и др.)
- Дополнительные конструкторы для составных концептов
- Конструкторы для построения *составных ролей* из простых ролей с помощью операций:
 - инверсии, пересечения, объединения, дополнения ролей
 - композиции ролей, транзитивного замыкания
- Многоместные роли (обозначающие n -арные отношения)
- Дополнит. виды аксиом в TBox и ABox, например: $\neg R(a,b)$

Неформальное соглашение об именовании таких логик: обычно путем добавления к имени $\mathcal{ALC}$ букв, отвечающих добавленным в язык конструкторам и др. средствам (происхождение букв – из англ. названий конструкторов)

ЧАСТЫЕ РАСШИРЕНИЯ $\mathcal{ALC}$

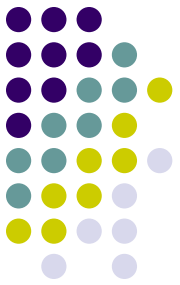


Наиболее известные расширения:

- $\mathcal{F}$ – Функциональность ролей: концепты вида $(\leq 1 R)$: существует не более одного R -последователя
- $\mathcal{N}$ – Ограничения кардинальности ролей: концепты $(\leq n R)$: существует не более n R -последователей
- $\mathcal{Q}$ – Качественные ограничения кардинальности ролей: концепты вида $(\leq n R.C)$: , означающие, что существует не более n R -последователей в C
- $\mathcal{I}$ – Обратные (инверсные) роли: если R есть роль, то R^{-} – тоже роль, обращение исходного отношения
- $\mathcal{O}$ – Номиналы: если a есть имя индивида, то $\{a\}$ есть концепт, означающий одноэлементное множество

Например, расширение инверсными ролями, номиналами и ограничениями кардинальности ролей: логика $\mathcal{ALCIOQ}$

ДРУГИЕ РАСШИРЕНИЯ $\mathcal{ALC}$



- $\mathcal{H}$ – Иерархия ролей: в $\mathcal{TBox}$ допускаются аксиомы вложенности ролей $R \sqsubseteq S$
- $\mathcal{R}$ – Составные аксиомы вложенности ролей в $\mathcal{TBox}$
- $\mathcal{S}$ – Транзитивные роли: в $\mathcal{TBox}$ возможны аксиомы транзитивности вида $Tr(R)$
- $\mathcal{D}$ – Расширение языка конкретными типами ПО
- ❖ Иногда аксиомы для ролей выделяются в отдельный набор, который называют *иерархией ролей* или $\mathcal{RBox}$
- ❖ Буква $\mathcal{S}$ (*system*) не добавляется к имени логики, а замещает в нём буквы $\mathcal{ALC}$, например:
логика $\mathcal{SHIQ}$, с инверсными ролями (буква $\mathcal{I}$),
качественными ограничениями кардинальности ролей ($\mathcal{Q}$),
транзитивными ролями ($\mathcal{S}$) и иерархией ролей ($\mathcal{H}$)

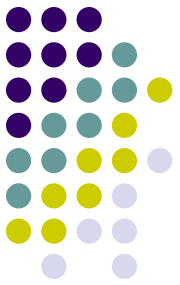
Фундаментальные логические проблемы дескрипт. логик:
разрешимость проблем выполнимости концепта,
совместимости БЗ, ответа на конъюнктивные запросы

РАСШИРЕНИЯ $\mathcal{LS}$: ПРИМЕРЫ



- Функциональное ограничение роли: (≤ 1 *Мать*)
(у любого живого существа одна мать)
- Ограничение кардинальности роли: ($\leq n$ *Игрок*)
(в команде не более n игроков)
- Пары инверсных ролей: *Родитель* – *Ребёнок*,
Раньше – *Позже* ...
- Транзитивное замыкание роли *Родитель* – роль *Предок*, роли *Ребёнок* – роль *Потомок*
- Вложенность ролей: Мать $\sqsubseteq$ Родитель
- Симметричные роли: *Равенство*(=) , *Друг*
- Рефлексивные роли: *Равенство*(=) ,
Эквивалентность ($\equiv$)

СРАВНЕНИЕ БЗ с БД



Базы Знаний и Базы Данных формулируются на разных языках, однако сравнение возможно

- В БД могут фигурировать многочестные отношения (предикаты), однако есть ДЛ с многочестными отношениями, которые сводятся к традиционным
- В БЗ есть терминологический компонент (ТBox), его слабый аналог – схема БД, она менее выразительна, т.к. допускаются лишь аксиомы вида $R_1 \subseteq R_2$ (R_i – предикаты с одинаковым числом аргументов)
- В БД в качестве запросов по сути допускаются произвольные формулы логики предикатов, а в БЗ ДЛ запросы ограниченных видов, менее выразительны даже конъюнктивные запросы (представляют некоторый фрагмент логики предикатов)
- Однако более важное отличие – в семантике: что именно является логическим следствием для заданной БЗ и БД

ГЛАВНОЕ ОТЛИЧИЕ БЗ от БД



- ❖ В БД принято **предположение о замкнутости мира - CWA** (если некоторое утверждение не является истинным, то оно принимается ложным)
 - ❖ В БЗ ДЛ – **предположение об открытости мира - OWA** (подобное утверждение считается ни истинным, ни ложным)
- Это кардинальным образом влияет на то, какие факты считаются логически выводимыми
- БД формально представляет собой одну модель – ту, в которой каждое отношение состоит в точности из тех кортежей, что записаны в БД, и по этой причине
 - ❖ проблема произвольных логических запросов разрешима (хотя не разрешима в логике предикатов), т.к. сводится к проверке утверждения на этой фиксированной модели
 - ❖ БД – монотонный вывод?
 - БЗ представляет целое семейство моделей, и для ответа на запрос требуется их все перебрать
 - ❖ Ответ на запрос в БЗ есть всегда подмножество ответа БД

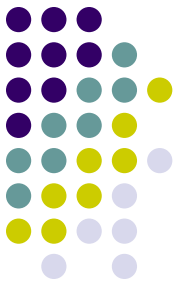
Пример: **hasChild (Ann, Mark)** – сколько детей у **Ann**?

ДЕСКРИПТИВНЫЕ ЛОГИКИ: РЕЗЮМЕ



- Deskriptivnye logiki razlichayutsya po vyrazitel'nym sredstvam, no stroyatsya takim obrazom, chtoby BZ byli razreshimymi teoriyami
- DL ne imeyut deduktivnoy sistemy (pravil vyvoda)
- Dva ponimaniya termina *deskriptivnaya logika*:
 - konkretnaya logicheskaya sistema
 - nauka, izuchayushchaya takie sistemy
- Osnovnaya zadacha nauki DL – izucheniye razreshimosti osnovnykh problem, postroyeniye algoritmov ix resheniya, otsenka ix vychislitel'noy slozhnosti
- Dlya DL razrabotany programmnye sistemy logicheskogo vyvoda (*reasoners*), oni razlichayutsya
 - po podderzhivaemому vidu DL
 - po tipu realizovannoy v nih razreshayushchey protsedury (*tablo-algoritm*, rezolyutsiya i t. p.),
 - po formatam iskhodnykh dannyyh, po YP dlya realizatsii

ЗАКЛЮЧЕНИЕ



Онтологии, Дескриптивные логики и *Semantic Web*

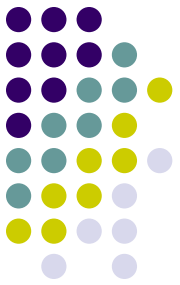
Для того, чтобы содержащиеся в онтологиях знания были пригодны для машинной обработки, (в частности, для автоматизированного логического вывода новых знаний из уже имеющихся)

- Язык, на котором формулируются онтологии, должен иметь точную, формальную семантику.
- Соответствующие логические проблемы должны быть разрешимы и иметь допустимую на практике вычислительную сложность.
- Желательно, чтобы язык имел довольно большую выразительную силу, пригодную для формулировки на нём практически значимых фактов.



СПАСИБО ЗА ВНИМАНИЕ!

ДОМАШНЕЕ ЗАДАНИЕ № 7



Дополнение к Практическому заданию №2

по разработке формальной декларативной онтологии на основе системы *PROTÉGÉ*

- На основе логики $\mathcal{ALC}$ или ее расширения записать несколько (5-12) утверждений-аксиом разного типа о понятиях (классах) построенной онтологии (3-4 балла)
 - Срок выполнения – 3 недели, до **21 апреля**
 - **Отчет по заданию** (1-2 стр.) должен включать указание использованной дескриптивной логики и перечень аксиом с комментариями (включая использованные обозначения)
 - Отчет м. б. приложен к отчету по Пр3 №2
- Сдавать в распечатанном виде или высылать на почту