ПРЕДСТАВЛЕНИЕ ЗНАНИЙ, ЛОГИЧЕСКАЯ МОДЕЛЬ: ФОРМАЛЬНЫЕ ТЕОРИИ

СОДЕРЖАНИЕ

1. Логическая модель ПЗ : *sinopsis*
2. Понятие формальной/логической теории
   - компоненты, семантика, свойства
3. Исчисление высказываний (ИВ)
   - язык, интерпретации, свойства
4. Исчисление предикатов первого порядка (ИП)
   - язык предикатов, семантика, свойства
5. Доказательство теорем: метод резолюций
6. Языки *Prolog* и *Datalog*
7. Заключение, Домашнее задание

ЛОГИЧЕСКАЯ МОДЕЛЬ

- Исторически первая: применение для ПЗ уже существующего языка предикатов (ЯП)
  - ЯП разработан ходе исследования логических основ ЕЯ
- В логической модели знания представляются как совокупность правильных формул какой-либо *формальной логической системы*
- Достоинства формальных логических систем:
  - изученность свойств
  - возможность применения механизма формального вывода (доказательства)
- Операции над знаниями – на основе логического вывода

ФОРМАЛЬНАЯ СИСТЕМА

<u>Формальная логическая система/логика/теория</u>

- Истоки появления: изучение и формализация
  Истинность *Vs.* Доказуемость
  Идея замены доказательства/рассуждения – вычислением, формальными символьными преобразованиями строк
- Формализация, аксиоматизация математических теорий, теоремы – формулы, выводимые из формул-аксиом
- Формальные логические теории отличаются по выразительной мощности и эффективности вывода
  - <u>Исчисление (логика) высказываний</u>
  - <u>Исчисление (логика) предикатов первого порядка</u>
  - Многосортные логики первого порядка
  - <u>Дескриптивные логики</u>
  - Модальные логики (возможность/необходимость)
  - Нечеткая и вероятностная логики

ФОРМАЛЬНАЯ СИСТЕМА: КОМПОНЕНТЫ

Формальная логическая система / логика / теория задается четверкой компонент: $T = < E, R, A, P >$

- E – множество базовых элементов формул
- R – синтаксические правила, на основе которых из элементов Е образуются синтаксически правильные (*правильно построенные*) формулы теории
- A – множество аксиом: подмножество всего множества синтаксически правильных формул, которым априорно приписывается статус истинности.
- P – множество правил формального вывода (оперирования символами), применяя которые к формулам (аксиомам), можно получать новые п.п. формулы (теоремы): $F \vdash G$ (*G* выводима из *F*)

$E+R$ – ***сигнатура теории*** (ее язык)

ФОРМАЛЬНАЯ СИСТЕМА: ИНТЕРПРЕТАЦИЯ

Интерпретация I, или формальная семантика:

- Задаются: Область интерпретации D и Функция, сопоставляющая каждой формуле *F* теории некоторое содержательное высказывание *P* (истинное или ложное) относительно объектов множества D
- Формула *F*:
  - Выполнима, если существует интерпретация, где *P* истинно
  - Общезначима, если истинна в любой интерпретации
  - Противоречива (невыполнима), если ложна в любой интерпретации
- Логическое следствие: $F \models G$, если *G* выполнима (истинна) в любой интерпретации, в которой выполнима *F*
- Формальная семантика называется также модельной: *модель теории* T – интерпретация, в которой ни одна теорема не является противоречивой (все выполнимы)

ФОРМАЛЬНАЯ СИСТЕМА: ОСНОВНЫЕ СВОЙСТВА

Для формальной теории $\mathcal{T}$:

- Полнота (содержательная, относительно интерпретации): Любому истинному высказыванию *P* соответствует теорема в $\mathcal{T}$ (т.е. истинность *P* м.б. обоснована выводом)
- Непротиворечивость
  - Формальная: не выводимы одновременно формулы *F* и ¬*F*
  - Содержательная: существует модель
- Разрешимость: существует алгоритм, который для любой формулы *F* определяет, является ли она теоремой (истинность/ложность формулы)
- Полуразрешимость: существует алгоритм, который для любой формулы-теоремы дает положительный ответ (для нетеорем алгоритм может зациклиться)

❖ Вычислительная сложность – важный вопрос

ИСЧИСЛЕНИЕ ВЫСКАЗЫВАНИЙ

- Элементы логических рассуждений – простые высказывания (истинные или ложные)

  МГУ основано в 1755 году

- Пропозиция (*proposio*) – предложение/высказывание, рассматриваемое только с точки зрения истинности

- $\boldsymbol{E}$ – *пропозициональные переменные*: $p, q, r, \ldots$

  логические связки: ¬ (отрицание), → (импликация)

  логические значения (истина и ложь): ⊤ и ⊥

  служебные символы: скобки и запятые

- $\boldsymbol{R}$ – индуктивные правила: п.п.формула –

  1) переменная 2) логическое значение

  3) $\neg F$; $F_1 \rightarrow F_2$, где F, F_1, F_2 – п.п.формулы

- Другие логические связки вводятся как синтаксические сокращения: $x \wedge y \equiv \neg (x \rightarrow \neg y)$ $x \vee y \equiv \neg x \rightarrow y$

- Пример п.п.формулы: $\neg q \wedge (p \rightarrow \neg r)$

ИСЧИСЛЕНИЕ ВЫСКАЗЫВАНИЙ: ИНТЕРПРЕТАЦИЯ

Традиционная: Булева алгебра

(или алгебра высказываний, алгебра булевских функций)

- Каждой пропозициональной переменной ставится в соответствие истинностное значение (истина/ложь – 1 / 0)
- Логические связки трактуются традиционно, как соответствующие булевские функции
- Тогда любая п.п. формула получает истинностное значение
- В такой интерпретации: $p \to \neg p$ выполнима

  $p \vee \neg p$ общезначима; $p \wedge \neg p$ противоречива
- Для Булевой алгебры есть процедура (алгоритм) определения для любой формулы-функции всех ее значений (построение таблицы истинности)

ИСЧИСЛЕНИЕ ВЫСКАЗЫВАНИЙ: ЛОГИЧЕСКИЙ ВЫВОД

Аксиомы и правила вывода ИВ? –

дедуктивная система ($\mathcal{A}$ + $\mathcal{P}$)

Возможно несколько аксиоматизаций, например, КИВ (классическое исчисление высказываний):

- $\mathcal{A}$ – 3 аксиомы, в том числе $(A \rightarrow (B \rightarrow A))$
- $\mathcal{P}$ – правило отделения (*modus ponens*): $\dfrac{A\,,\; A \rightarrow B}{B}$

Пример теоремы:

$$\vdash (\neg q \rightarrow \neg p) \rightarrow (p \rightarrow q)$$

- Теоремы (законы) ИВ – схемы истинных сложных высказываний (утверждений):

$$\vdash (\neg q \rightarrow \neg p) \rightarrow (p \rightarrow q)$$

$$\vdash (\neg p \rightarrow (q \wedge \neg q)) \rightarrow p$$

ИСЧИСЛЕНИЕ ВЫСКАЗЫВАНИЙ: СВОЙСТВА ТЕОРИИ

Основные свойства ИВ :

- Полнота: теоремы и только они являются общезначимыми формулами
- Непротиворечивость (оба вида)
- Разрешимость
  - Стандартная разрешающая процедура проверки общезначимости формулы (по таблице истинности)
  - Алгоритм *DPLL* проверки выполнимости формул ИВ в конъюнктивной нормальной форме (NP-полная задача: правило резолюции + поиск с возвратами)

Эффективность вывода, однако :

Язык (сигнатура) ИВ – слабая выразительность

ИСЧИСЛЕНИЕ ВЫСКАЗЫВАНИЙ: ДРУГАЯ ИНТЕРПРЕТАЦИЯ

Семантика множеств (предикатная семантика)

- Область интерпретации $\mathcal{D}$; пропозицион. переменным (высказываниям) ставятся в соответствие подмножества $\mathcal{D}$
- Интерпретация логических связок и значений:

$$\top - \mathcal{D}; \quad \bot - \varnothing; \quad \neg F - \mathcal{D} \setminus F; \quad A \vee B - A \cup B$$

(*F*, *A*, *B* – формулы) $\quad A \wedge B - A \cap B$

$A \to B - A \supset B$: вложение мн-в, A влечет В (“если A то В”)

- Доказана эквивалентность двух семантик (свойств), т.о.
- ИВ описывает не только законы логических операций над простыми высказываниями, но и законы, которым подчиняются операции над множествами
- Традиционную булевскую семантику можно рассматривать как одноэлементную интерпретацию

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ

Существенно бо́льшая выразительность:

- Элементы логических рассуждений – высказывательные формы, зависящие от переменных:
  Он живет в городе Нижний Новгород
- Входящие в формы объекты могут задаваться не только именами, но и более сложным способом: именные формы, зависящие от переменных и констант:
  sin(x) > 0.5 *Том любит своюсестру*
- Сложный способ именования объектов задается в ИП понятием ***терма***: *сестра (Том)*
- Высказывательная форма – это ***предикат*** (если нет переменных, вырождается в простое высказывание)
- Роль формул ИП – описывать сложные высказывания, состоящие из предикатов, к которым могут быть применены операторы существования и общности

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ЭЛЕМЕНТЫ ФОРМУЛ

Множество $\boldsymbol{E}$ базовых элементов формул включает:

- $\boldsymbol{E_1}$ – предметные постоянные/константы $a, b, c, \ldots$
- $\boldsymbol{E_2}$ – множество функциональных символов $f, g, h, \ldots$
- $\boldsymbol{E_3}$ – множество предикатных символов $P, Q, R, \ldots.$

$\boldsymbol{E_1} \cup \boldsymbol{E_2} \cup \boldsymbol{E_3}$ называется обычно *сигнатурой теории*

Вспомогательные символы:

- $\boldsymbol{E_4}$ – множество предметных переменных $x, y, z, \ldots$
- $\boldsymbol{E_5}$ – логические связки $\wedge, \vee, \neg, \rightarrow$ и

  кванторы: $\forall$ (всеобщность) и $\exists$ (существование)
- $\boldsymbol{E_6}$ – служебные символы: запятые, скобки

Для каждого функционального и предикатного символа определена арность (число аргументов): $n \geq 0$

(при $n=0$ вырождаются в константы и пропозиц. переменные)

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: СИНТАКСИЧЕСКИЕ ПРАВИЛА

- Терм – это либо
  - предметная переменная
  - предметная постоянная/константа
  - выражение вида $f(t_1, \ldots t_n)$, где f – функц. символ, а $t_1, \ldots t_n$ – термы
- Элементарная (атомарная) формула – это выражение вида $P(t_1, \ldots t_k)$, где P – предикатный символ, а $t_1, \ldots t_k$ – термы
- Формула (п.п. формула) – это либо
  - элементарная формула
  - выражение $\neg F$, $F \rightarrow G$, $\forall x\ (F)$, $\exists x\ (F)$, где F и G – п.п.ф.
- Теория первого порядка:
  - кванторы могут связывать только предм. переменные, запрещены кванторы по предикатам и функц. символам
  - предикаты не могут быть аргументами др. предикатов

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ПРИМЕРЫ ФОРМУЛ

(1) $\forall x\, (P(x) \vee \neg P(x))$

(2) $\exists x\ \forall y\, Q(x,y) \rightarrow \forall y\, \exists x\, Q(x,y)$

(3) $\forall x\, (Студент(x) \wedge Старший(курс(x)) \rightarrow \exists y\, Научный\_руководитель(y, x))$

- Термы содержательно интерпретируются как имена объектов/сущностей
- Дизъюнкция и конъюнкция вводятся как синтаксические сокращения
- Скобки указывают порядок операций и область действия кванторов, их можно опускать с учетом установленного приоритета: $\neg, \forall, \exists, \wedge, \vee, \rightarrow$
- Формула замкнута, если в ней нет свободных (т.е. не связанных кванторами) переменных – см. примеры 1-3

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ЧИСТОЕ И ПРИКЛАДНОЕ

Дедуктивная система (Аксиомы и правила вывода) ?
Один из вариантов:

- $\boldsymbol{A}$ – 3 аксиомы КИВ + 2 аксиомы: универсальная конкретизация и экзистенциональное обобщение
- $\boldsymbol{P}$ – правило отделения (*modus ponens*)+ 2 правила связывания квантором общности и существования

Чистое ИП – в нем нет предметных констант, функций, *собственных (*дополнит.) аксиом (слайд 15, примеры 1, 2)
Выражает «чистую» логику (сугубо логические законы)

Прикладное ИП – в сигнатуре есть постоянные и функции (слайд 15, пример 3 , а также:
формальная арифметика, теория равенства и др.)

Вложение теорий: ИВ $\subset$ ЧистИП $\subset$ ПриклИП

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ИНТЕРПРЕТАЦИЯ

Интерпретация I прикладного ИП на области D :

- Каждой предметной постоянной ставится в соответствие элемент из D
- Каждому функциональному n-местному символу – функция (операция) $f : D^n \to D$
- Каждому n-местному предикатному символу – отношение R на множестве D
- Логические связки трактуются традиционно (бул. функции)
- Кванторы для одноместных предикатов :

  $\forall x\, P(x) = \wedge P(c_i) \qquad c_i \in D$ (бескон. мн-во формул)

  $\exists x\, P(x) = \vee P(c_i) \qquad c_i \in D$

  Для многоместных предикатов – аналогично.

Так получается содержательное высказывание на D , и всякая замкнутая формула истинна /ложна в заданной I

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ПОЛНОТА

Общезначимость/Выполнимость/Непротиворечивость формул: $\forall x\, (P(x) \vee \neg P(x))$?

$\exists x\ \forall y\ Q(x,y) \rightarrow \forall y\ \exists x\ Q(x,y)$?

(Замечание: незамкнутая формула истинна в интерпретации I, т. и т. тогда, когда ее универсальное замыкание истинно)

Свойства логики первого порядка?

- Полнота чистого ИП (теорема Гёделя,1929 г.): любому истинному высказыванию P соответствует теорема в T
- Корректность: все теоремы общезначимы (истинны)
- Тем самым имеется однозначная связь логич. истинности высказывания и его выводимости.

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ДРУГИЕ СВОЙСТВА

- Неполнота некоторых прикладных ИП (в частности, Формальная арифметика – Гёдель,1930 г.):
  в них есть недоказуемые истинные утверждения.
  Причина : наличие аксиом с рекурсией
- Непротиворечивость (формальная и содержательная)
- Неразрешимость (теорема Черча,1936 г.):
  не существует общего метода (алгоритма) установления общезначимости (тождественной истинности) произвольной формулы F,
  т.е. является ли она теоремой или нет.
- Полуразрешимость: есть процедура проверки общезначимости, которая для формулы-теоремы дает положительный ответ, а в случае формулы-нетеоремы возможно зацикливание.

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: ДОКАЗАТЕЛЬСТВО ТЕОРЕМ

- Задача доказательства некоторой формулы F (теоремы), т.е. ее вывода из формул-аксиом $A_1, \dots, A_n$ сводится к выяснению логического следования F

  $A_1, \dots, A_n \vDash F$, что равносильно доказательству:

  – общезначимости формулы $A_1 \wedge \dots \wedge A_n \rightarrow F$ или

  – невыполнимости формулы $A_1 \wedge \dots \wedge A_n \wedge \neg F$

- Поскольку ИП полуразрешимо: априорное предположение об общезначимости формулы, которое надо доказать
  - метод Эрбрана, 1930
  - обратный метод Маслова, 1964
  - метод резолюций Робинсона,1965 – обратный вывод,, практически применимый!

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ: МЕТОД РЕЗОЛЮЦИЙ

Доказательство выводимости F из аксиом $A_1 \ldots A_n$ методом доказательства невыполнимости формулы

$$A_1 \wedge \ldots \wedge A_n \wedge \neg F.$$

- Все формулы переводятся в конъюнктивную форму , а кванторы отбрасываются – работа с множеством $\boldsymbol{M}$ формул-дизъюнкций
- Обобщенное правило логического вывода: резолюция
- На каждом шаге вывода из пары дизъюнкций из $\boldsymbol{M}$ строится *резольвента*:

$$B_1 \vee \ldots \vee \boldsymbol{P} \vee \ldots \vee B_k \quad \text{и} \quad C_1 \vee \ldots \vee \boldsymbol{\neg P} \vee \ldots \vee C_m$$

их резольвента: $B_1 \vee \ldots \vee B_k \vee C_1 \vee \ldots \vee C_m$

- Полученная резольвента добавляется в $\boldsymbol{M}$
- Цель – получить пустую резольвенту (тем самым доказать невыполнимость, т.е. выводимость F)

МЕТОД РЕЗОЛЮЦИЙ: ДЕТАЛИ

Шаг метода резолюций:

1. Выбор из M пары формул, в которых есть контрарные литералы: предикат P и его отрицание
2. Унификация (обычно – наиболее общая): подстановка вместо переменных выбранных формул определённых термов для совпадения всех аргументов P и $\neg P$
3. Получение резольвенты и добавление её в M

Особенности:

- В общем случае неоднозначность: какую пару формул выбрать? (возможны различные стратегии, перебор)
- Резольвента может быть длиннее исходных формул, в то время как цель – получить пустую резольвенту

- В целом: полная система доказательства формул логики первого порядка для БЗ в конъюнктивной форме

МЕТОД РЕЗОЛЮЦИЙ: ПРИМЕР

Исходные утверждения:

$$Кот(Мурзик) \qquad \forall x(\, Кот(x) \rightarrow Живое\_существо(x))$$

$$\forall y(\, Живое\_существо(y) \rightarrow Дышит(y))$$

Доказать: $Дышит\ (Мурзик)$

Множество M :

$$Кот(Мурзик) \qquad (1)$$

$$\neg Кот(x) \vee Живое\_существо(x) \qquad (2)$$

$$\neg Живое\_существо(y) \vee Дышит\ (y) \qquad (3)$$

$$\neg Дышит(Мурзик) \qquad (4)$$

Резольвента (2) и (3): $\neg Кот(z) \vee Дышит(z)$ (5)

Резольвента (4) и (5): $\neg Кот(Мурзик)$ (6)

Резольвента (1) и (6): пустая резольвента

ЯЗЫК ПРОЛОГ

- Язык Пролог – язык предикатов первого порядка, но особые синтаксис и процедура доказательства.
- Клаузальная форма записи формул (*клаузы Хорна*), при которой кванторы опускаются, например:
  пролог-правило P:- Q, R, T. – это формула $Q \wedge R \wedge T \to P$, равносильная $\neg Q \vee \neg R \vee \neg T \vee P$ – *хорновский дизъюнкт*,
  только одно слагаемое без отрицания.
- Пролог-предложение является хорновским дизъюнктом. Факты и правила Пролога – аксиомы, вопрос – теорема

  Дышит(y):- Живое_существо(y).

- Стратегия доказательства – линейная резолюция (*SLD*) для хорновских дизъюнктов, более эффективная
- Проблемы обратного вывода (от доказываемого факта): незавершаемые и избыточные вычисления

ЯЗЫК *DATALOG*

- Синтаксически: подмножество Пролога:
  ЯП без функциональных символов (функций)
- *DataLog* ориентирован на работу с БД: язык запросов
- Отличающаяся от Пролога семантика:
  - Операции над конечными множествами
  - Порядок предложений не важен: декларативность !
  - Прямой вывод: доступен для понимания и сочетается с реляционными операциями БД
  - Различение экстенсиональных и интенсиональных предикатных символов (разделение БЗ)
  - Ограничения на рекурсию, логическое отрицание и др.
- Важно: разрешимость вывода (полиномиальное время?)
- Язык запросов: – алгоритмически не полон (по Тьюрингу)
  – мощнее других языков запросов к БД (рекурсия)

ЗАКЛЮЧЕНИЕ

- ИВ и ИП – две формальные системы вывода (дедукции), отражающие общелогические законы, но отличающиеся выразительностью языка.
- Модельная семантика строится на базе непосредственного установления соответствия синтаксических понятий языка логики с понятиями объектного мира, описываемого этим языком.
- Формальная логическая система дает возможность вывода из известных утверждений (знаний) – новых утверждений (знаний), позволяя более экономно хранить информацию в БЗ.
- ПЗ в логической модели – на основе прикладной (конкретизированной) формальной теории.

СПАСИБО ЗА ВНИМАНИЕ!

ДОМАШНЕЕ ЗАДАНИЕ

Повторить основные понятия и синтаксические правила формул исчисления предикатов первого порядка