Логический подход к представлению знаний

Артем Таболин

ВМК МГУ 2012

Логический подход к представлению знаний

Логические системы

- Логика высказываний (исчисление высказываний)
- Логика первого порядка (исчисление предикатов первого порядка)
- ...
- Логика высшего порядка.

Высказывания

Высказывание — базовое понятие исчисления высказываний. Соответствует какому-либо простейшему утверждению. Может быть истинным или ложным.

Обозначения

A, B, C, ...

Примеры

- Песок белый
- Идёт снег
- Река замёрзла

Формулы исчисления высказываний

Комбинируя с помощью логических операций $(\vee, \wedge, \Rightarrow, \ldots)$ высказывания можно получать более сложные утверждения. Такие выражения и есть формулы.

Определение

1. Любое высказывание является формулой.
2. Если X и Y — формулы, то $(\neg X)$, $(X \vee Y)$, $(X \wedge Y)$, $(X \Rightarrow Y)$ — тоже формулы.
3. Других формул нет.

Примеры

- $A \wedge (B \vee \neg A)$
- $A \Rightarrow (A \vee B)$
- $(D \wedge \neg C) \vee (\neg D \wedge C)$

База знаний

Факт — это формула исчисления высказываний, про которую известно, что она истинна.

При использовании исчисления высказываний базой знаний будет являться набор фактов.

Перевод предложений ЕЯ на логический язык

Связки ЕЯ	Логические связки
не A	$\neg A$
A и B	$A \wedge B$
Если A, то B; B, потому что A	$A \Rightarrow B$
A или B	$A \vee B$

Примеры

Солнце не светит	$\neg A$
Светит солнце и идёт дождь	$A \wedge B$
Мне тепло, потому что светит солнце	$A \Rightarrow C$
Тренировка состоится если на улице будет тепло или нам разрешат занять спортзал	$(C \vee D) \Rightarrow E$

Трудности перевода

В ЕЯ существует множество других связок. Однако, любые связки (в том числе рассмотренные) не являются такими однозначными, как может показаться на первый взгляд.

Примеры

- Предложения «Ему стало страшно и он убил чужака» и «Он убил чужака и ему стало страшно» не эквивалентны, так как связка «и» здесь выражает причинно-следственную или временную связи.
- В предложении «Завтра я пойду в кино или на каток» связка «или» имеет значение исключающего или ($\oplus$).

Ограничения исчисления высказываний

Все зелёные яблоки кислые, а все красные - сладкие.

Ограничения исчисления высказываний

Все зелёные яблоки кислые, а все красные - сладкие.

- Не позволяет оперировать с отдельными объектами.
- Не позволяет сформулировать утверждение, верное сразу для множества объектов.

Логика первого порядка

- Набор объектов
- Функции, отображающие объекты в объекты
- Предикаты задают отношения между объектами
- Формулы связывают предикаты

Термы

Терм по сути является выражением, значением которого является некоторый объект.

Определение

1. Любая константа является термом (a, b, c, ...).
2. Любая переменная является термом (..., x, y, z).
3. Если f — функциональная константа, а t_1, t_2, ..., t_n — термы, то $f(t_1, t_2, \ldots, t_n)$ — тоже терм.
4. Других термов нет.

Примеры

- x
- $f(g(a), x, y)$
- Цвет(x)
- Объём(Яблоко1)

Атомы

Атом — простейшее отношение между объектами.

Определение

Если P — n-местная предикатная константа, а t_1, t_2, ..., t_n — термы, то $P(t_1, t_2, ..., t_n)$ — атом.

В частности, при $n = 0$ атом P эквивалентен высказыванию исчисления высказываний.

Примеры

- P
- $Q(f(g(a), x, b), x)$
- Голодный(Студент1)
- Сидит(x, Стул1)

Формулы логики первого порядка

Определение

1. Любой атом является формулой.
2. Если X и Y — формулы, то $(\neg X)$, $(X \vee Y)$, $(X \wedge Y)$, $(X \Rightarrow Y)$ — тоже формулы.
3. Если Y — формула, а x — переменная, то $(\forall x)Y$, $(\exists x)Y$ — тоже формулы.
4. Других формул нет.

Примеры

- $P(x, a) \wedge (Q(x) \vee \neg P(x, b))$
- $P(x, y) \Rightarrow P(y, x)$
- $(\forall x)P(x) \Rightarrow (\exists y)Q(x, y)$

Кванторы

Если D — множество объектов, то:

- $(\forall x)(A(x))$ эквивалентно $\bigwedge\limits_{x \in D} A(x)$.
- $(\exists x)(A(x))$ эквивалентно $\bigvee\limits_{x \in D} A(x)$.

Свободные и связанные переменные

Понятие свободной и связанной переменной относится к конкретному вхождению переменной в формулу.

Определение

1. Если атом $P(t_1, \ldots, t_n)$ зависит от переменной x, то в формулу $P(t_1, \ldots, t_n)$ она входит свободно.
2. Операции $(\neg X)$, $(X \vee Y)$, $(X \wedge Y)$, $(X \Rightarrow Y)$ не меняют свойства свободных и связанных вхождений.
3. Если Y — формула, в которой есть свободные вхождения переменной x, то в формулах $(\forall x)Y$ и $(\exists x)Y$ эти вхождения являются связанными.

Свободные и связанные переменные: Примеры

- $(\forall x)P(x, y)$. Связанные: x. Свободные: y.
- $((\forall x)P(x)) \wedge Q(x)$. Переменная x входит как свободно, так и связанно.
- $((\forall x)P(x)) \wedge ((\exists x)Q(x))$.

База знаний

Факт — это не содержащая свободных переменных формула логики первого порядка, про которую известно, что она истинна.

Базой знаний также являтся набор фактов.

Задача логического вывода

Исходные данные

- База знаний (набор фактов).
- Цель — факт, который надо доказать или опровергнуть.

Задача

Доказать истинность, ложность или неопределённость цели.

Задача логического вывода

Исходные данные

- База знаний (набор фактов).
- Цель — факт, который надо доказать или опровергнуть.

Задача

Доказать истинность, ложность или неопределённость цели.

Методы

- Прямой вывод
- Обратный вывод

- Метод резолюций

Выводы

- $\frac{\alpha \Rightarrow \beta,\ \alpha}{\beta}$
- $\frac{\alpha \wedge \beta}{\alpha}$
- $\frac{\alpha \Leftrightarrow \beta}{(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)}$
- $\frac{\alpha \Rightarrow \beta}{\neg \alpha \vee \beta}$

Метод прямого вывода в исчислении высказываний

Используя некоторый набор правил логического вывода применяем их к имеющемуся набору фактов получая новые факты.

Если в какой-то момент мы получили цель или её отрицание, то можно закончить работу алгоритма.

Если используя правила вывода нельзя получить новые факты, то ничего сказать нельзя

Резолюции

Единичная резолюция

$$\frac{\alpha_1 \vee \ldots \vee \alpha_n, \ \neg\alpha_i}{\alpha_1 \vee \ldots \vee \alpha_{i-1} \vee \alpha_{i+1} \vee \ldots \vee \alpha_n}$$

Полная резолюция

$$\frac{\alpha_1 \vee \ldots \vee \alpha_n, \ \beta_1 \vee \ldots \vee \beta_m}{\alpha_1 \vee \ldots \vee \alpha_{i-1} \vee \alpha_{i+1} \vee \ldots \vee \alpha_n \vee \beta_1 \vee \ldots \vee \beta_{j-1} \vee \beta_{j+1} \vee \ldots \vee \beta_m}$$

Здесь $\alpha_i = \neg\beta_j$.

Метод прямого вывода в исчислении высказываний: резолюции

Если представить набор фактов в виде КНФ, то метод прямого вывода на основе полной резолюции является полным

Определённые выражения

Определённое выражение — дизъюнкция атомов, среди которых положительным является не более чем один.

Можно представить как импликацию:

$$x \vee \neg y \vee \neg z = (y \wedge z) \Rightarrow x$$

Метод прямого вывода в логике первого порядка

Если представить базу знаний в виде фактов, являющихся определёнными выражениями, то следующий алгоритм является полным:

1. Если верны все предпосылки для какого-либо правила — добавляем новый факт в базу.
2. Если ответ найден, завершаем работу.
3. Если невозможно добавить новые факты в базу, то завершаем работу. Иначе — выполняем п.1.

Метод резолюций для исчисления высказываний

Допустим у нас есть база знаний $\{A_1, \ldots, A_n\}$ перед нами стоит задача доказать факт B.

Если система фактов $\{A_1, \ldots, A_n\}$ непротиворечива, то система $\{A_1, \ldots, A_n, \neg B\}$ непротиворечива тогда и только тогда, когда B истинно.

Набор фактов $\{A_1, \ldots, A_n, \neg B\}$, представленный в виде КНФ является противоречивым, если из него можно вывести пустую дизъюнкцию.

Инженерия знаний

1. Постановка задачи
2. Сбор знаний
3. Определение словаря предикатов, функций, констант
4. Регистрация общих знаний о предметной области
5. Описание конкретного экземпляра задачи
6. Передача запросов процедуре логического вывода и получение ответов
7. Отладка базы знаний

Источники

- Стюарт Рассел, Питер Норвиг, Искусственный интеллект: современный подход
- А. Тей, П. Грибомон и др., Логический подход к искусственному интеллекту