Метод резолюций

Ищем доказательство в исчислении предикатов

Хотим научиться проверять доказуемость формул исчисления предикатов. В общем случае невозможно, но человек как-то справляется? Может, для каких-то частных случаев мы сможем предложить метод?

По теореме о полноте можем рассматривать ($\models$) вместо ($\vdash$). Напомним: $\models \alpha$, если для всех $M = \langle D, F, P, E \rangle$ выполнено $M \models \alpha$. Нам мешает:

1. бесконечное множество предметных множеств D и оценок;
2. бесконечный перебор для кванторов;

Будем последовательно упрощать задачу:

1. упростим формулу;
2. заменим произвольное D на рекурсивно-перечислимое, устроенное некоторым фиксированным образом;
3. научимся по этому перечислимому D искать доказательство / противоречие.

Компактность

Определение

Пространство X компактно, если из любого его открытого покрытия U можно выделить конечное подпокрытие:
$X = \cup U$, *существует* $V \subseteq U$, *что* $|V| < \aleph_0$ *и* $X = \cup V$.

Пример

$(0, 1)$ *не компактен. Например,* $U = \{(\varepsilon/2, \varepsilon) \mid \varepsilon \in (0, 1)\}$. *Пусть* $V \subset U$ *и* $|V| < \aleph_0$. *Тогда* $\min\{a \mid (a, b) \in V\} > 0$.

Пример

$[0, 1]$ *компактен. Выберем* U *и покажем, что в нём есть подпокрытие. Рассмотрим все подотрезки вида* $[a, x]$ *где* $a < x$, *имеющие конечное покрытие. Несложно показать, что* $\max x = 1$.

Теорема Гёделя о компактности

Теорема

Если Γ *— некоторое семейство формул, то* Γ *имеет модель тогда и только тогда, когда любое его конечное подмножество имеет модель.*

Сколемизация. Упрощаем формулу.

1. Предварённая форма (поверхностные кванторы):
   $\psi := Qx_1.Qx_2 \dots Qx_n.\varphi(x_1, \dots, x_n)$
2. Для упрощения предполагаем, что кванторы чередуются. Это не сильно уменьшает общность. Например, если $D = \mathbb{N}$, то
   $(\forall x.\forall y.\varphi(x, y)) \leftrightarrow (\forall p.\varphi(\mathsf{plog}_2(p), \mathsf{plog}_3(p))$
3. Убрать кванторы существования:
   $\zeta = \exists x_1.\forall x_2.\exists x_3.\forall x_4 \dots \exists x_{n-1}.\forall x_n.\varphi(x_1, \dots, x_n)$
   Заменим x_{2k+1} функцией Сколема: $e_{2k+1}(x_2, x_4, \dots, x_{2k})$.
   Получим: $\eta = \forall x_2.\forall x_4 \dots \forall x_n.\varphi(e_1, x_2, e_3(x_2), \dots, e_{n-1}(x_2, x_4, \dots, x_{n-2}), x_n)$
   Очевидно, что $\vdash \zeta$ эквивалентно $\models \zeta$ эквивалентно $\models \eta$ и $\vdash \eta$.
4. ДНФ:

$$\forall x_2.\forall x_4 \dots \forall x_n. \bigwedge_c \left(\bigvee_{i=\overline{1,d(c)}} (\neg) P_i(\theta_i) \right)$$

Эрбранов универсум

Определение

Пусть $H_0(\varphi)$ *— все константы в формуле* φ *(либо особая константа a, если констант в* φ *нет)*
$H_1(\varphi)$ *—* $H_0(\varphi)$ *и все функции от значений* $H_0(\varphi)$ *(как строки)*
$H_2(\varphi)$ *—* $H_1(\varphi)$ *и все функции от значений* $H_1(\varphi)$ *(как строки)*
$H = \cup H_n(\varphi)$ *— основные термы.*

Пример

$P(a) \vee Q(f(b))$:
$H_0 = \{a, b\}$
$H_1 = \{a, b, f(a), f(b)\}$
$H_2 = \{a, b, f(a), f(b), f(f(a)), f(f(b))\}$
...
$H = \{f^{(n)}(x) \mid n \in \mathbb{N}_0, x \in \{a, b\}\}$

Выполнимость

Теорема

Формула выполнима тогда и только тогда, когда она выполнима на Эрбрановом универсуме.

Доказательство.

($\Rightarrow$) Пусть $M \models \forall \overline{x}.\varphi$. Тогда построим отображение eval : $H \to M$ (смысл названия вдохновлён языками программирования: eval("$f(f(b))$") перейдёт в $f(f(b))$, где f и b — из M).
Предикатам дадим согласованную оценку: $P_H(t_1, \ldots, t_n) = P_M(h(t_1), \ldots, h(t_n))$.
Очевидно, любая формула сохранит своё значение, кванторы всеобщности по меньшему множеству также останутся истинными.
($\Leftarrow$) Очевидно. $\square$

Противоречивость

Определение

Система дизъюнктов $\{\delta_1, \ldots, \delta_n\}$ противоречива, если для каждой интерпретации M найдётся δ_k и такой набор $d_1 \ldots d_v$, что $[\![\delta_k]\!]^{x_1:=d_1,\ldots,x_v:=d_v} = Л$.

Теорема

Система дизъюнктов противоречива, если она невыполнима на Эрбрановом универсуме.

Доказательство.

Контрапозиция теоремы о выполнимости + разбор определения. □

Основные примеры

Определение

Дизъюнкт с подставленными основными термами вместо переменных называется основным примером. Системой основных примеров назовём множество основных примеров опровержимых дизъюнктов:
Если $M \not\models \delta_k$ для некоторой эрбрановской интерпретации, то возьмём все возможные основные примеры δ_k.

Теорема

Система дизъюнктов S противоречива тогда и только тогда, когда система всевозможных основных примеров E противоречива

Доказательство.

Для некоторой эрбрановой интерпретации дизъюнкт δ_k опровергается тогда и только тогда, когда соответствующая ему подстановка в E опровергается. □

Теорема Эрбрана

Теорема (Эрбрана)

Система дизъюнктов S противоречива тогда и только тогда, когда существует конечное противоречивое множество основных примеров системы дизъюнктов S

Доказательство.

($\Leftarrow$) Пусть $\delta_1[\overline{x} := \overline{\theta}], \ldots, \delta_k[\overline{x} := \overline{\theta}]$ — противоречивое множество примеров дизъюнктов. Тогда интерпретация $\overline{\theta}$ опровергает хотя бы один из δ_k и система противоречива.

($\Rightarrow$) Если S противоречива, то значит, множество основных примеров S противоречиво (по теореме о выполнимости Эрбранова универсума). Тогда по теореме компактности в нём найдётся конечное противоречивое подмножество. $\square$

Правило резолюции (исчисление высказываний)

Пусть даны два дизъюнкта, $\alpha_1 \vee \beta$ и $\alpha_2 \vee \neg\beta$. Тогда следующее правило вывода называется правилом резолюции:

$$\frac{\alpha_1 \vee \beta \qquad \alpha_2 \vee \neg\beta}{\alpha_1 \vee \alpha_2}$$

Теорема

Система дизъюнктов противоречива, если в процессе всевозможного применения правила резолюции будет построено явное противоречие, т.е. найдено два противоречивых дизъюнкта: β *и* $\neg\beta$.

Алгебраические термы

Определение

Алгебраический терм

$$\theta := x | (f(\theta_1, \ldots, \theta_n))$$

где x — переменная, $f(\theta_1, \ldots, \theta_n)$ — применение функции. Напомним, что константы — нульместные функциональные символы, собственно переменные будем обозначать последними буквами латинского алфавита.

Определение

Система уравнений в алгебраических термах

$$\begin{cases} \theta_1 = \sigma_1 \\ \vdots \\ \theta_n = \sigma_n \end{cases}$$

где θ_i и σ_i — термы

Уравнение в алгебраических термах

Определение

$\{x_i\} = X$ *— множество переменных,* $\{\theta_i\} = T$ *— множество термов.*

Определение

Подстановка — отображение вида: $\pi_0 : X \to T$*, тождественное почти везде.* $\pi_0(x)$ *может быть либо* $\pi_0(x) = \theta_i$*, либо* $\pi_0(x) = x$*.*

Доопределим $\pi : T \to T$, где

1. $\pi(x) = \pi_0(x)$
2. $\pi(f(\theta_1, \ldots, \theta_k)) = f(\pi(\theta_1), \ldots, \pi(\theta_k))$

Определение

Решить уравнение в алгебраических термах — найти такую наиболее общую подстановку π*, что* $\pi(\theta_1) = \pi(\theta_2)$*. Наиболее общая подстановка — такая, для которой другие подстановки являются её частными случаями.*

Задача унификации

Определение

Пусть даны формулы α и β. Тогда решением задачи унификации будет такая наиболее общая подстановка $\pi = \mathcal{U}[\alpha, \beta]$, что $\pi(\alpha) = \pi(\beta)$.
Также, η назовём наиболее общим унификатором.

Пример

- *Формулы $P(a, g(b))$ и $P(c, d)$ не имеют унификатора (мы считаем, что a, b, c, d — нульместные функции, а f — одноместная функция).*

Правило резолюции для исчисления предикатов

Определение

Пусть σ_1 и σ_2 — подстановки, заменяющие переменные в формуле на свежие. Тогда правило резолюции выглядит так:

$$\frac{\alpha_1 \vee \beta_1 \qquad \alpha_2 \vee \neg\beta_2}{\pi(\sigma_1(\alpha_1) \vee \sigma_2(\alpha_2))}\ \pi = \mathcal{U}\big[\sigma_1(\beta_1), \sigma_2(\beta_2)\big]$$

σ_1 и σ_2 разделяют переменные у дизъюнктов, чтобы π не осуществила лишние замены, ведь $\vdash (\forall x.P(x)\ \&\ Q(x)) \leftrightarrow (\forall x.P(x))\ \&\ (\forall x.Q(x))$, но $\nvdash (\forall x.P(x) \vee Q(x)) \rightarrow (\forall x.P(x)) \vee (\forall x.Q(x))$.

Пример

$$\frac{Q(x) \vee P(x) \quad \neg P(a) \vee T(x)}{Q(a) \vee T(x'')}$$ *подстановки:* $\sigma_1(x) = x', \sigma_2(x) = x'', \pi(x') = a$

Метод резолюции

Ищем $\vdash \alpha$.

1. найдём опровержение $\neg\alpha$.
2. перестроим $\neg\alpha$ в ДНФ.
3. будем применять правило резолюции, пока получаем новые дизъюнкты и пока не найдём явное противоречие (дизъюнкты вида β и $\neg\beta$).

Если противоречие нашлось, значит, $\vdash \neg\neg\alpha$. Если нет — значит, $\vdash \neg\alpha$. Процесс может не закончиться.

SMT-решатели

Обычно требуется не логическое исчисление само по себе, а теория первого порядка. То есть, «Satisfability Modulo Theory», «выполнимость в теории» — вместо SAT, выполнимости.

- Иногда можно вложить теорию в логическое исчисление, даже в исчисление высказываний: $\overline{S_2 S_1 S_0} = \overline{A_1 A_0} + \overline{B_1 B_0}$

$$
\begin{array}{ll}
S_0 = A_0 \oplus B_0 & C_0 = A_0 \,\&\, B_0 \\
S_1 = A_1 \oplus B_1 \oplus C_0 & C_1 = (A_1 \,\&\, B_1) \vee (A_1 \,\&\, C_0) \vee (B_1 \,\&\, C_0) \\
S_2 = C_1 &
\end{array}
$$

- А можно что-то добавить прямо на уровень унификации / резолюции: Например, можем зафиксировать арифметические функции — и производить вычисления в правиле резолюции вместе с унификацией.
  Тогда противоречие в $\{x = 1 + 3 + 1, \neg x = 5\}$ можно найти за один шаг.