Математическая логика и логическое программирование

mk.cs.msu.ru → Лекционные курсы
→ Математическая логика и логическое программирование (3-й поток)

Блок 24

Эрбрановские интепретации
Теорема об эрбрановских интерпретациях

Лектор:
Подымов Владислав Васильевич
E-mail:
valdus@yandex.ru

ВМК МГУ, 2023/2024, осенний семестр

Краткое вступление

исходная формула
φ

отрицание
$\psi = \neg\varphi$

ПНФ
ψ_{pnf}

ССФ
ψ_{ssf}

система дизъюнктов
$S_{\psi_{ssf}}$

резолютивный вывод
пустого дизъюнкта $\square$
из $S_{\psi_{ssf}}$

$$\models \varphi \quad \Leftrightarrow \quad \not\models \psi \quad \Leftrightarrow \quad \not\models \psi_{pnf} \quad \Leftrightarrow \quad \not\models \psi_{ssf} \quad \Leftrightarrow \quad \not\models S_{\psi_{ssf}}$$

$\Leftarrow$ существует вывод $\square$ из $S_{\psi_{ssf}}$

Последняя недостающая деталь метода — «$\Leftrightarrow$» на месте «$\Leftarrow$»

Краткое вступление

Временно забудем про метод резолюций
и попробуем выделить *общую часть* рассуждений «в лоб»
о невыполнимости системы дизъюнктов S в заданной интерпретации $\mathcal{I}$

Для примера рассмотрим такую систему: $S = \{\mathrm{P}(\mathrm{x}),\ \neg\mathrm{P}(\mathbf{f}(\mathbf{c}))\}$

Для обоснования невыполнимости S достаточно заметить,
что в любой модели $\mathcal{I}$ для S предмет, описываемый термом $\mathbf{f}(\mathbf{c})$,
и обладает свойством P $(\mathcal{I} \models \forall\mathrm{x}\ \mathrm{P}(\mathrm{x}))$, и не обладает $(\mathcal{I} \models \neg\mathrm{P}(\mathbf{f}(\mathbf{c})))$

В этом замечании не используются
природа предметной области $\mathcal{I}$ и устройство оценок $\bar{\mathbf{c}}$ и $\bar{\mathbf{f}}$,
и важно лишь то, каким термом задаётся «противоречивый» предмет

Если в интерпретации $\mathcal{I}$ заменить каждый предмет
множеством основных термов, описывающих этот предмет,
и сохранить устройство оценки P, то в результате получится

эрбрановская интерпретация

Эрбрановские интерпретации

Эрбрановской интерпретацией (или, по-другому, — $\mathcal{H}$-интерпретацией) сигнатуры $\sigma = \langle \mathrm{Const}, \mathrm{Func}, \mathrm{Pred} \rangle$ называется интерпретация $\langle \mathcal{H}_\sigma, \overline{\mathrm{Const}_\mathcal{H}}, \overline{\mathrm{Func}_\mathcal{H}}, \overline{\mathrm{Pred}} \rangle$, где:

- $\mathcal{H}_\sigma$ — эрбрановский универсум ($\mathcal{H}$-универсум): множество это множество всех основных термов сигнатуры
  - σ, если $\mathrm{Const} \neq \emptyset$
  - $\langle \{\mathbf{c}_\mathcal{H}\}, \mathrm{Func}, \mathrm{Pred} \rangle$, если $\mathrm{Const} = \emptyset$ ($\mathbf{c}_\mathcal{H}$ — эрбрановская константа)
- $\overline{\mathrm{Const}_\mathcal{H}}(\mathbf{c}) = \mathbf{c}$ для любой константы $\mathbf{c}$ из Const
- $\overline{\mathrm{Func}_\mathcal{H}}(\mathbf{f})(t_1, \dots, t_n) = \mathbf{f}(t_1, \dots, t_n)$ для любых $\mathbf{f}^{(n)} \in \mathrm{Func}$ и $t_1, \dots, t_n \in \mathcal{H}_\sigma$
- $\overline{\mathrm{Pred}}$ — произвольная оценка предикатных символов

Первые три пункта означают, что $\mathcal{H}$-интерпретация — это интерпретация, построенная над свободной алгеброй термов

Все $\mathcal{H}$-интерпретации заданной сигнатуры отличаются друг от друга только выбором оценки $\overline{\mathrm{Pred}}$

Эрбрановские интерпретации

Эрбрановский базис ($B_{\mathcal{H}}$) — это множество всех атомов, построенных над термами $\mathcal{H}$-универсума

$\mathcal{H}$-интерпретация $\mathcal{I}$ всегда и однозначно определяется тем, какие атомы из $B_{\mathcal{H}}$ в ней истинны и какие нет, то есть множеством

$$B^{\mathcal{I}} = \{A \mid A \in B_{\mathcal{H}},\ \mathcal{I} \models A\}$$

Например,

- $B^{\mathcal{I}} = \emptyset$: все основные атомы **ложны** в $\mathcal{I}$
- $B^{\mathcal{I}} = B_{\mathcal{H}}$: все основные атомы **истинны** в $\mathcal{I}$
- $B^{\mathcal{I}} = B^{\mathcal{J}'} \cap B^{\mathcal{J}''}$: в $\mathcal{I}$ истинны те и только те основные атомы, которые истинны в обеих интерпретациях $\mathcal{J}'$, $\mathcal{J}''$

Для удобного использования теоретико-множественной нотации иногда будем отождествлять $\mathcal{H}$-интерпретацию $\mathcal{I}$ с множеством $B^{\mathcal{I}}$

Теорема об эрбрановских интерпретациях

Система дизъюнктов выполнима тогда и только тогда, когда она имеет эрбрановскую модель

Доказательство.

Рассмотрим произвольную систему дизъюнктов S и докажем теорему для этой системы

($\Leftarrow$) Очевидно:
если S выполняется в некоторой эрбрановской интерпретации,
то S выполняется хотя бы в одной интерпретации

($\Rightarrow$)

Пусть система S выполнима

Тогда для неё существует модель $\mathcal{I} = \langle \mathrm{D}, \overline{\mathrm{Const}}, \overline{\mathrm{Func}}, \overline{\mathrm{Pred}} \rangle$

Построим по интерпретации $\mathcal{I}$
эрбрановскую модель $\mathcal{I}_{\mathcal{H}}$ для S
той же сигнатуры σ

Теорема об эрбрановских интерпретациях

Доказательство. $(\Rightarrow)$ $\qquad \mathcal{I} = \langle \mathrm{D}, \overline{\mathrm{Const}}, \overline{\mathrm{Func}}, \overline{\mathrm{Pred}} \rangle \mapsto \mathcal{I}_{\mathcal{H}}$

Если Const $= \emptyset$, то добавим и произвольно оценим константу $\mathbf{c}_{\mathcal{H}}$ в $\mathcal{I}$

Поставим такое соответствие $\alpha : \mathcal{H}_\sigma \to D$:

$$\alpha(t) \text{ — значение терма } t \text{ в } \mathcal{I}$$

Зададим оценку $\overline{\overline{\mathrm{P}}}$ каждого предикатного символа P в $\mathcal{I}_{\mathcal{H}}$ так:

$$\overline{\overline{\mathrm{P}}}(t_1, \dots, t_k) = \overline{\mathrm{P}}(\alpha(t_1), \dots, \alpha(t_k))$$

То же самое другими словами — зададим интерпретацию $\mathcal{I}_{\mathcal{H}}$ так:

$$\mathcal{I}_{\mathcal{H}} = \{\mathrm{P}(t_1, \dots, t_k) \mid \mathrm{P}(t_1, \dots, t_k) \in B_{\mathcal{H}}, \overline{\mathrm{P}}(\alpha(t_1), \dots, \alpha(t_k)) = \mathbb{t}\}$$

Покажем, что такая интерпретация $\mathcal{I}_{\mathcal{H}}$ является моделью для S

Но для начала приведём пример, чтобы стало *понятнее*,
как соотносятся $\mathcal{I}$ и $\mathcal{I}_{\mathcal{H}}$

Теорема об эрбрановских интерпретациях

Доказательство. ($\Rightarrow$) $\qquad \mathcal{I} = \langle \mathrm{D}, \overline{\mathrm{Const}}, \overline{\mathrm{Func}}, \overline{\mathrm{Pred}} \rangle \mapsto \mathcal{I}_{\mathcal{H}}$

Пример: рассмотрим интерпретацию $\mathcal{I}$ сигнатуры $\langle \{0\}, \{\mathbf{s}^{(1)}\}, \{<^{(2)}\} \rangle$ с *естественным* арифметическим устройством:

- предметная область — множество всех целых чисел
- $\overline{0}$ — число 0; $\qquad \overline{\mathbf{s}}(n) = n + 1$;
  $\overline{<}$ — отношение строгого неравенства чисел

Тогда:

- $\mathcal{H}_\sigma = \{0, \quad \mathbf{s}(0), \quad \mathbf{s}(\mathbf{s}(0)), \quad \mathbf{s}(\mathbf{s}(\mathbf{s}(0))), \quad \dots\}$
- $\mathcal{I}_{\mathcal{H}} =$
$$\left\{\begin{array}{llll} 0 < \mathbf{s}(0), & \mathbf{s}(0) < \mathbf{s}(\mathbf{s}(0)), & \mathbf{s}(\mathbf{s}(0)) < \mathbf{s}(\mathbf{s}(\mathbf{s}(0))), & \dots \\ 0 < \mathbf{s}(\mathbf{s}(0)), & \mathbf{s}(0) < \mathbf{s}(\mathbf{s}(\mathbf{s}(0))), & \mathbf{s}(\mathbf{s}(0)) < \mathbf{s}(\mathbf{s}(\mathbf{s}(\mathbf{s}(0)))), & \dots \\ 0 < \mathbf{s}(\mathbf{s}(\mathbf{s}(0))), & \mathbf{s}(0) < \mathbf{s}(\mathbf{s}(\mathbf{s}(\mathbf{s}(0)))), & \mathbf{s}(\mathbf{s}(0)) < \mathbf{s}(\mathbf{s}(\mathbf{s}(\mathbf{s}(\mathbf{s}(0))))), & \dots \\ \dots & \dots & \dots & \dots \end{array}\right\}$$

Теорема об эрбрановских интерпретациях

Доказательство. ($\Rightarrow$) $\mathcal{I} = \langle \mathrm{D}, \overline{\mathrm{Const}}, \overline{\mathrm{Func}}, \overline{\mathrm{Pred}} \rangle \mapsto \mathcal{I}_{\mathcal{H}}$

Предположим от противного, что $\mathcal{I}_{\mathcal{H}} \not\models S$

Тогда в S содержится дизъюнкт D, такой что $\mathcal{I}_{\mathcal{H}} \not\models D$

Пусть, для ясности, $D = \forall \widetilde{\mathrm{x}}^{\mathrm{n}}\ (A_1 \vee \dots \vee A_k \vee \neg B_1 \vee \dots \vee \neg B_m)$,
где $A_1, \dots, A_k, B_1, \dots, B_m$ — атомы

Тогда существуют термы $t_1, \dots, t_n \in \mathcal{H}_\sigma$, такие что

$$\begin{array}{lcl} \mathcal{I}_{\mathcal{H}} \not\models A_1[t_1, \dots, t_n] & \dots & \mathcal{I}_{\mathcal{H}} \not\models A_k[t_1, \dots, t_n] \\ \mathcal{I}_{\mathcal{H}} \models B_1[t_1, \dots, t_n] & \dots & \mathcal{I}_{\mathcal{H}} \models B_m[t_1, \dots, t_n] \end{array}$$

По заданию интерпретации $\mathcal{I}_{\mathcal{H}}$, верно и

$$\begin{array}{lcl} \mathcal{I} \not\models A_1[\alpha(t_1), \dots, \alpha(t_n)] & \dots & \mathcal{I} \not\models A_k[\alpha(t_1), \dots, \alpha(t_n)] \\ \mathcal{I} \models B_1[\alpha(t_1), \dots, \alpha(t_n)] & \dots & \mathcal{I} \models B_m[\alpha(t_1), \dots, \alpha(t_n)] \end{array}$$

Следовательно, $\mathcal{I} \not\models \forall \widetilde{\mathrm{x}}^{\mathrm{n}}\ (A_1 \vee \dots \vee A_k \vee \neg B_1 \vee \dots \vee \neg B_m)$,
что *противоречит* выбору $\mathcal{I}$ как **модели** для S

Значит, предположение «$\mathcal{I}_{\mathcal{H}} \not\models S$» неверно, то есть $\mathcal{I}_{\mathcal{H}} \models S$ ▼