

Математическая логика

mk.cs.msu.ru → Лекционные курсы → Математическая логика (318, 319/2, 241, 242)

Блок 38

Основные свойства аксиоматических теорий

Лектор:

Подымов Владислав Васильевич

E-mail:

valdus@yandex.ru

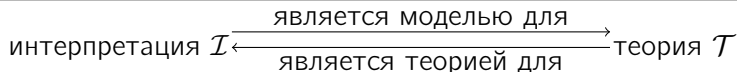
Теории и модели

Каждая теория \mathcal{T} представляет собой множество предложений, а значит,

все интерпретации можно разбить относительно \mathcal{T} на два класса:

- ▶ являющиеся **моделями** \mathcal{T} и
- ▶ не являющиеся моделями \mathcal{T}

Теория \mathcal{T} называется **теорией для интерпретации** \mathcal{I} той же сигнатуры, если \mathcal{I} — модель теории \mathcal{T}



Наряду с «теория для интерпретации» будем также говорить «**теория интерпретации**» (без «для»)

Непротиворечивость

Теория \mathcal{T} называется **непротиворечивой**, если она имеет хотя бы одну модель, а иначе — **противоречивой**

Утверждение. Если теория \mathcal{T} противоречива, то любая формула φ \mathcal{T} -общезначима, \mathcal{T} -невыполнима и не \mathcal{T} -выполнима

Доказательство. Если теория \mathcal{T} не имеет ни одной модели, то по **определению логического следствия**:

- ▶ $\mathcal{T} \models \forall \tilde{x}^n \varphi(\tilde{x}^n)$, то есть $\models_{\mathcal{T}} \forall \tilde{x}^n \varphi(\tilde{x}^n)$,
то есть формула φ **\mathcal{T} -общезначима**
- ▶ $\mathcal{T} \models \forall \tilde{x}^n \neg \varphi(\tilde{x}^n)$, то есть $\models_{\mathcal{T}} \forall \tilde{x}^n \neg \varphi(\tilde{x}^n)$,
то есть формула φ **\mathcal{T} -невыполнима**
и, следовательно, **не является \mathcal{T} -выполнимой** ▼

Таким образом, все противоречивые теории **абсолютно бессмысленны**, и имеет смысл рассматривать только непротиворечивые теории

Непротиворечивость

Пример: теория частичных порядков —

это теория сигнатуры $\langle \emptyset, \emptyset, \{<^{(2)}\} \rangle$, содержащая две аксиомы:

- ▶ аксиома **антирефлексивности**

$$\forall x \neg(x < x)$$

- ▶ аксиома **транзитивности**

$$\forall x \forall y \forall z ((x < y) \& (y < z) \rightarrow (x < z))$$

Легко видеть, что моделями этой теории являются все интерпретации, в которых «<» оценивается как строгий частичный порядок, и только они

Следовательно, теория частичных порядков

- ▶ является теорией любых интерпретаций описанного выше вида и только таких интерпретаций и
- ▶ непротиворечива

Непротиворечивость

Другой пример:

теория \mathcal{T} сигнатуры $\langle \{0, 1, 2\}, \{\mathbf{s}^{(1)}, +^{(2)}\}, \{=^{(2)}\} \rangle$ из блока 37

$$\left\{ \begin{array}{l} 2 = \mathbf{s}(1) \quad 1 = \mathbf{s}(0) \\ \forall x (x + 0 = x) \quad \forall x \forall y (x + \mathbf{s}(y) = \mathbf{s}(x + y)) \\ \forall x (x = x) \quad \forall x \forall y ((x = y) \rightarrow (y = x)) \\ \forall x \forall y \forall z ((x = y) \& (y = z) \rightarrow (x = z)) \\ \forall x \forall y ((x = y) \rightarrow (\mathbf{s}(x) = \mathbf{s}(y))) \\ \forall x \forall y \forall u \forall v ((x = u) \& (y = v) \rightarrow (x + y = u + v)) \end{array} \right.$$

«Арифметическая» интерпретация \mathcal{I} является моделью для \mathcal{T} ,
то есть теория \mathcal{T} непротиворечива

Но помимо этого моделью \mathcal{T} является и такая интерпретация \mathcal{J} :

- ▶ В предметной области содержится один предмет d
- ▶ $\bar{0} = \bar{1} = \bar{2} = \bar{\mathbf{s}}(d) = \bar{+}(d, d) = d$
- ▶ $\equiv(d, d) = \mathfrak{t}$

Так как $\mathcal{I} \models \mathcal{T}$ и $\mathcal{I} \not\models (1 + 1 = 1)$, то верно $\not\models_{\mathcal{T}} (1 + 1 = 1)$

Так как $\mathcal{J} \models \mathcal{T}$ и $\mathcal{J} \models (1 + 1 = 1)$, то верно $\not\models_{\mathcal{T}} \neg(1 + 1 = 1)$

Полнота и элементарность

Теория \mathcal{T} из последнего примера оказалась **не самой лучшей**:

- ▶ эта теория придумывалась для обоснования арифметических утверждений,
- ▶ но нашлось утверждение φ (« $1 + 1 = 1$ »), которое невозможно ни обосновать ($\not\models_{\mathcal{T}} \varphi$), ни опровергнуть ($\not\models_{\mathcal{T}} \neg\varphi$)

Для каждой интерпретации \mathcal{I} (**очевидно, что**) справедливо следующее свойство: для любого предложения φ верно либо $\mathcal{I} \models \varphi$, либо $\mathcal{I} \models \neg\varphi$

Значит, и для **самой лучшей** теории, предназначенной для обоснования утверждений, смысл которых задаётся интерпретацией \mathcal{I} , должно быть справедливо аналогичное свойство

Теория \mathcal{T} называется **полной**, если для любого предложения φ выполняется хотя бы одно из соотношений

$$\models_{\mathcal{T}} \varphi, \quad \models_{\mathcal{T}} \neg\varphi$$

Полнота и элементарность

Каждой интерпретации \mathcal{I} соответствует такая «очевидная» теория:

- ▶ Переберём всевозможные предложения φ
(об алгоритме речь не идёт, так что бесконечность числа предложений не считаем препятствием для перебора)
- ▶ Если $\mathcal{I} \models \varphi$, то добавим его в теорию (объявим аксиомой), а иначе не будем добавлять

Такая теория называется **элементарной теорией интерпретации** \mathcal{I} :

$$\text{Th}(\mathcal{I}) = \{\varphi \mid \varphi \in \text{CForm}, \mathcal{I} \models \varphi\}$$

Утверждение. Любая интерпретация является моделью своей элементарной теории

Доказательство. По определению элементарной теории, каждая аксиома элементарной теории $\text{Th}(\mathcal{I})$ выполняется в \mathcal{I} ▼

Следствие

Элементарная теория любой интерпретации непротиворечива

Полнота и элементарность

Утверждение. Элементарная теория любой интерпретации полна

Доказательство.

Покажем (согласно определению полноты теории), что для любых интерпретации \mathcal{I} предложения φ верно хотя бы одно из соотношений

$$\models_{\text{Th}(\mathcal{I})} \varphi, \quad \models_{\text{Th}(\mathcal{I})} \neg\varphi$$

Случай 1: $\mathcal{I} \models \varphi$

По определению элементарной теории: $\varphi \in \text{Th}(\mathcal{I})$

По очевидному свойству логического следования: $\text{Th}(\mathcal{I}) \models \varphi$

По определению общезначимости в теории: $\models_{\text{Th}(\mathcal{I})} \varphi$

Случай 2: $\mathcal{I} \not\models \varphi$

По семантике \neg : $\mathcal{I} \models \neg\varphi$

По случаю 1: $\models_{\text{Th}(\mathcal{I})} \neg\varphi \blacktriangledown$

Полнота и элементарность

У любой непротиворечивой теории существует бесконечно много моделей

Но оказывается, что модели **полной** теории в некотором смысле одинаковы

Интерпретации \mathcal{I} и \mathcal{J} называются **элементарно эквивалентными**, если их элементарные теории равны

Утверждение

Теория полна \Leftrightarrow все её модели элементарно эквивалентны

Полнота и элементарность

Утверждение

Теория полна \Leftrightarrow все её модели элементарно эквивалентны

Доказательство.

(\Rightarrow) Рассмотрим произвольную полную теорию \mathcal{T}

Предположим, что не все модели \mathcal{T} элементарно эквивалентны

Тогда существуют модели \mathcal{I} , \mathcal{J} и формула φ , такие что $\mathcal{I} \models \varphi$ и $\mathcal{J} \not\models \varphi$

Раз верно $\mathcal{J} \not\models \varphi$, то верно и $\not\models_{\mathcal{T}} \varphi$

Раз верно $\mathcal{I} \models \varphi$, то верно и $\mathcal{I} \not\models \neg\varphi$, а значит, и $\not\models_{\mathcal{T}} \neg\varphi$

Следовательно, теория \mathcal{T} неполна (*противоречие*)

Полнота и элементарность

Утверждение

Теория полна \Leftrightarrow все её модели элементарно эквивалентны

Доказательство.

(\Leftarrow) Рассмотрим произвольную теорию \mathcal{T} , все модели которой элементарно эквивалентны

Покажем (согласно определению полноты теории), что для любого предложения φ верно хотя бы одно из соотношений

$$\models_{\mathcal{T}} \varphi, \quad \models_{\mathcal{T}} \neg\varphi$$

Пусть \mathcal{I} — произвольная модель теории \mathcal{T}

Случай 1: $\mathcal{I} \models \varphi$

По элементарной эквивалентности моделей, для любой модели \mathcal{J} теории \mathcal{T} верно $\mathcal{J} \models \varphi$

Следовательно, верно и $\models_{\mathcal{T}} \varphi$

Случай 2: $\mathcal{I} \not\models \varphi$

По семантике \neg , верно $\mathcal{I} \models \neg\varphi$

По случаю 1, верно и $\models_{\mathcal{T}} \neg\varphi \blacktriangledown$

Полнота и элементарность

Вернёмся к последнему примеру:

теория \mathcal{T} сигнатуры $\langle \{0, 1, 2\}, \{\mathbf{s}^{(1)}, +^{(2)}\}, \{=^{(2)}\} \rangle$

$$\left\{ \begin{array}{l} 2 = \mathbf{s}(1) \quad 1 = \mathbf{s}(0) \\ \forall x (x + 0 = x) \quad \forall x \forall y (x + \mathbf{s}(y) = \mathbf{s}(x + y)) \\ \forall x (x = x) \quad \forall x \forall y ((x = y) \rightarrow (y = x)) \\ \forall x \forall y \forall z ((x = y) \& (y = z) \rightarrow (x = z)) \\ \forall x \forall y ((x = y) \rightarrow (\mathbf{s}(x) = \mathbf{s}(y))) \\ \forall x \forall y \forall u \forall v ((x = u) \& (y = v) \rightarrow (x + y = u + v)) \end{array} \right\}$$

Модель \mathcal{I} : «арифметическая» интерпретация

Модель \mathcal{J} с одним предметом d :

- ▶ $\bar{0} = \bar{1} = \bar{2} = \bar{\mathbf{s}}(d) = \bar{+}(d, d) = d$
- ▶ $\bar{\equiv}(d, d) = \bar{\mathfrak{t}}$

Верно $\mathcal{I} \not\models (1 + 1 = 1)$ и $\mathcal{J} \models (1 + 1 = 1)$

Значит, модели \mathcal{I} и \mathcal{J} не являются элементарно эквивалентными, и теория \mathcal{T} неполна

Полнота и элементарность

Другой пример: теория частичных порядков

$$\mathcal{T}_{<} = \left\{ \begin{array}{l} \forall x \neg(x < x) \\ \forall x \forall y \forall z ((x < y) \&(y < z) \rightarrow (x < z)) \end{array} \right\}$$

Модель \mathcal{I} : множество чисел $\{0, 1, \dots, 10\}$ с естественным порядком

Модель \mathcal{J} : множество чисел \mathbb{N}_0 с естественным порядком

Предложение φ : $\forall x \exists y (x < y)$ («не существует максимального числа»)

Тогда верны соотношения

- ▶ $\mathcal{I} \not\models \varphi$ (10 — максимальное число)
- ▶ $\mathcal{J} \models \varphi$ (в \mathbb{N}_0 нет максимального числа)

Значит, модели \mathcal{I} и \mathcal{J} не являются элементарно эквивалентными, и теория $\mathcal{T}_{<}$ неполна

А примеры полных теорий появятся позже

Разрешимость

В блоке 13 коротко обсуждался вопрос автоматизации доказательства теорем (проверки общезначимости формул) логики предикатов

В блоке 36 было показано, что проблема общезначимости формул алгоритмически неразрешима (теорема Чёрча)

Но аналогичная проблема общезначимости в аксиоматической теории бывает и разрешимой

Теория \mathcal{T} называется разрешимой, если проблема \mathcal{T} -общезначимости формул алгоритмически разрешима

Проверка того, является ли теория разрешимой, бывает и трудной, но иногда, если теория достаточно проста и достаточно хороша, можно утверждать её разрешимость — например:

Утверждение. Любая конечная полная теория разрешима

Разрешимость

Утверждение. Любая конечная полная теория разрешима

Доказательство.

Рассмотрим произвольную конечную полную теорию $\mathcal{T} = \{\psi_1, \dots, \psi_k\}$

Предложим *какой-нибудь* алгоритм проверки \mathcal{T} -общезначимости произвольной формулы $\varphi(\tilde{x}^n)$

По **определению полноты**, верно хотя бы одно из соотношений

$$\mathcal{T} \models \forall \tilde{x}^n \varphi, \quad \mathcal{T} \models \neg \forall \tilde{x}^n \varphi$$

По **теореме о логическом следствии**, хотя бы одна из формул

$\chi_+ = (\psi_1 \& \dots \& \psi_k \rightarrow \forall \tilde{x}^n \varphi)$, $\chi_- = (\psi_1 \& \dots \& \psi_k \rightarrow \neg \forall \tilde{x}^n \varphi)$
общезначима

По **теореме о табличном выводе**, хотя бы одна из таблиц

$$T_+ = \langle \mid \chi_+ \rangle, \quad T_- = \langle \mid \chi_- \rangle$$

невыполнима

Разрешимость

Утверждение. Любая конечная полная теория разрешима

Доказательство.

$$(T_+ = \langle \mid \psi_1 \& \dots \& \psi_k \rightarrow \forall \tilde{x}^n \varphi \rangle, T_- = \langle \mid \psi_1 \& \dots \& \psi_k \rightarrow \neg \forall \tilde{x}^n \varphi \rangle)$$

Алгоритм проверки \mathcal{T} -общезначимости формул можно устроить так:

- ▶ Будем одновременно (*параллельно*) строить табличные выводы для T_+ и T_- согласно **стратегии** из доказательства **теоремы о полноте табличного вывода**
- ▶ Стратегией гарантируется, что за конечное число действий для одной из таблиц T_+ , T_- будет построен успешный табличный вывод
- ▶ Если успешный вывод построен для T_+ , то формула φ \mathcal{T} -общезначима
- ▶ Если успешный вывод построен для T_- , то формула φ не \mathcal{T} -общезначима ▼

Для самостоятельного размышления: а существует ли такая хорошая теория — непротиворечивая, конечная и полная?