Интерполяционная теорема для FOCL и родственные результаты

Дмитрий Ярцев

1 декабря 2020

Напоминание: язык FOCL

Под **сигнатурой** в логике первого порядка понимают множество, состоящее из:

- Константных символов;
- Функциональных символов (с арностью), в которые можно подставлять термы и получать термы;
- Предикатных символов (с арностью), в которые можно подставлять термы и получать атомарные формулы.

У нас все три множества будут не более, чем счётными; это делается для простоты и не влияет на истинность рассматриваемых утверждений.

Термы — это константы, переменные или функции от термов.

Формулы — это предикаты от термов или комбинации формул при помощи логических связок и кванторов.

Кванторы нельзя навешивать на переменные, которые находятся в области действия других кванторов.

Теории и модели

Предложение — это формула без свободных переменных, т.е. все её переменные находятся в области действия кванторов.
Теория в сигнатуре σ — это любое множество σ-предложений. Если в сигнатуре есть символ $=$, требуется, чтобы он был двуместным предикатом, а теория должна включать в себя аксиомы отношения эквивалентности на $=$.

Модель $\mathfrak{A}$ теории Γ — это:

- Множество-**носитель** A;
- Интерпретации констант $c_i^{\mathfrak{A}} \in A$;
- Интерпретации функций $f_i^{\mathfrak{A}} : A^{Arity(f_i)} \to A$;
- Интерпретации предикатов $R_i^{\mathfrak{A}} \subset A^{Arity(R_i)}$;
- Равенство обязано интерпретироваться как настоящее равенство на множестве .

Интерпретация стандартно продолжается на все термы и формулы.
Разумеется, чтобы $\mathfrak{A}$ была моделью Γ, необходимо, чтобы интерпертации всех предложений Γ были в ней истинны.

Семантическое следование и свойства FOCL

Говорят, что формула φ **семантически влечёт** ψ ($\varphi \vDash \psi$), если в каждой модели, в которой верна интерпретация φ, верна и интерпретация ψ.
Теория влечёт формулу, если эта формула верна в каждой модели этой теории.
Теория Γ называется **непротиворечивой**, если в Γ из двух формул вида φ и $\neg\varphi$ следует не более одной. Непротиворечивая теория **полна**, если из этих двух формул в ней лежит ровно одна.
Свойство 1: теория непротиворечива тогда и только тогда, когда у неё есть модель.
Свойство 2 — компактность: $\Gamma \vDash \varphi$ тогда и только тогда, когда для некоторого конечного $\Delta \subset \Gamma$ $\Delta \vDash \varphi$.

Первая полезная лемма

Лемма

Пусть Γ *— некоторая* σ*-теория, содержащая формулу вида* $\exists x\ \psi(x)$, φ *— некоторая* σ*-формула,* c *— константа в сигнатуре* σ*, не входящая ни в* Γ*, ни в* φ*. Тогда* $\Gamma \vDash \varphi \Leftrightarrow \Gamma \cup \{\psi(c)\} \vDash \varphi$*.*

Доказательство. Стрелка $\Rightarrow$ очевидна, проверим стрелку $\Leftarrow$. Модели $\Gamma \cup \{\psi(c)\}$ — это в точности те модели Γ, в которых c интерпретируется как свидетель $\exists x\ \psi(x)$. Изменение этой интерпретации, однако, никак не влияет на истинность Γ и φ, а свидетель найдётся в каждой модели. Значит, мы можем брать модели $\Gamma \cup \{\psi(c)\}$, менять в них интерпретацию c и получать все модели Γ. Формула φ при этом всегда остаётся истинной.

Вторая полезная лемма

Лемма

Пусть ни одна формула теории Γ *не содержит в себе константу* c*, при этом* $\Gamma \vDash \varphi(c, \overline{x})$. *Тогда* $\Gamma \vDash \forall y\ \varphi(y, \overline{x})$*, где переменная* y *не входит в формулу* φ.

Доказательство. Любая модель Γ остаётся моделью при изменении интерпретации c, поэтому формула φ верна вне зависимости от интерпретации c. Это и означает истинность формулы с квантором всеобщности.

Интерполяционная теорема Крейга

Теорема

Пусть φ *и* ψ *— такие предложения, что* $\varphi \vDash \psi$. *Тогда найдётся такое предложение* θ, *называемое* ***интерполянтом Крейга*** *для предложений* φ *и* ψ, *что:*

1. $\varphi \vDash \theta$ *и* $\theta \vDash \psi$.
2. *Всякий предикатный, функциональный или константный символ, входящий в* θ, *кроме, может быть, равенства, входит также и в* φ, *и в* ψ.

Замечание: исключение для равенства существенно: без него теорема неверна, например, для $\varphi = \exists x (P(x) \wedge \neg P(x)), \psi = \exists x\ Q(x)$.

Доказательство. Предположим, что для предложений φ и ψ не существует интерполянта Крейга, и выведем из этого существование модели для формулы $\varphi \wedge \neg\psi$.

Определим сигнатуры σ_1, состоящую из всех символов, входящих в предложение φ, σ_2 из всех символов ψ, и $\sigma_0 = \sigma_1 \cap \sigma_2$. Можно считать, что σ состоит в точности из символов, входящих в φ или ψ.

Добавим к сигнатуре σ счётное множество констант C и обозначим соответствующее обогащение как σ', аналогично определим множества $\sigma_0', \sigma_1', \sigma_2'$.

Неотделимые теории

Назовём σ_1'-теорию Γ и σ_2'-теорию Δ **неотделимыми**, если не существует такого σ_0'-предложения θ, что $\Gamma \vDash \theta$ и $\Delta \vDash \neg\theta$.
Неотделимые пары теорий — это в некотором смысле обобщение непротиворечивых.
Замечание. Если обе сигнатуры σ_1' и σ_2' содержат равенство, а две непустые теории в этих сигнатурах неотделимы, то они обе непротиворечивы.

Наша цель — построить два неотделимых расширения, одно в сигнатуре σ_1' для $\{\varphi\}$ и одно в сигнатуре σ_2' для $\{\neg\psi\}$, объединение которых непротиворечиво. Делать это мы будем рекурсивно (здесь для простоты нужна счётность сигнатуры): Начнём с теорий $\Gamma_0 = \{\varphi\}$ и $\Delta_0 = \{\neg\psi\}$. Перечислим все σ_1'-предложения $\varphi_0, \varphi_1, \ldots$ и все σ_2-предложения $\psi_0, \psi_1, \ldots$.

На шаге $m+1$ происходит следующее:

1. Если добавление к Γ_m предложения φ_m не нарушает отделимость с Δ_m, положим $\Gamma_{m+1} = \Gamma_m \cup \{\varphi_m\}$, иначе оставим $\Gamma_{m+1} = \Gamma_m$.
2. Аналогично, если добавление к Δ_m предложения ψ_m не нарушает отделимость с новой Γ_{m+1}, положим $\Delta_{m+1} = \Delta_m \cup \{\psi_m\}$, иначе оставим $\Delta_{m+1} = \Delta_m$.
3. Если на этапе 1 мы добавили в Γ_{m+1} формулу вида $\exists x \lambda(x)$, то выберем ещё не задействованную в Γ_{m+1} и Δ_m константу $c \in C$ и добавим туда ещё и формулу $\lambda(c)$.
4. Аналогично, если на этапе 2 в Δ_{m+1} добавилась формула вида $\exists x \delta(x)$, то возьмём ещё не участвовавшую нигде константу $d \in C$ и добавим туда же формулу $\delta(d)$.

На каждом шаге промежуточные теории конечны, поэтому константы никогда не закончатся. Изначально неотделимость теорий означает в точности отсутствие интерполянта Крейга (все новые константы можно заменить на кванторы по второй полезной лемме). Проверим, что теории остаются неотделимы. Для этапов 1 и 2 это так по построению. Убедимся, что для этапа 3 — тоже (для этапа 4 — аналогично).

Пусть после присоединения к Γ_{m+1} предложения $\lambda(c)$ нашлось такое σ_0'-предложение $\theta = \theta(c, \overline{x})$, что $\Gamma_{m+1} \cup \{\lambda(c)\} \vDash \theta$ и $\Delta_{m+1} \vDash \neg\theta$.
С одной стороны, $\Gamma_{m+1} \cup \{\lambda(c)\} \vDash \exists y\ \theta(y, \overline{x})$ для новой переменной y. По первой полезной лемме $\lambda(c)$ можно выкинуть и оставить $\Gamma_{m+1} \vDash \exists y\ \theta(y, \overline{x})$.
С другой стороны, по второй полезной лемме не только $\Delta_{m+1} \vDash \neg\theta(c, \overline{x})$, но и $\Delta_{m+1} \vDash \forall y \neg\theta(y, \overline{x})$.
Две формулы в правых частях являются отрицаниями друг друга, поэтому теории Γ_{m+1} и Δ_{m+1} должны были быть отделимы ещё до этапа 3.

Возьмём теперь

$$\Gamma_\omega = \bigcup_{m\in\mathbb{N}} \Gamma_m, \Delta_\omega = \bigcup_{m\in\mathbb{N}} \Delta_m.$$

В силу компактности две теории остались неотделимы. Тогда они обе непротиворечивы. Проверим, что непротиворечиво и их объединение.
Сначала заметим, что Γ_ω полна в σ_1', а Δ_ω полна в σ_2'.
Действительно, пусть нашлось такое предложение φ_m в σ_1', что $\varphi_m \notin \Gamma_\omega, \neg\varphi_m \notin \Gamma_\omega$.

Поскольку мы не добавили φ_m на этапе 2, для какого-то σ_0'-предложения θ верно

$$\Gamma_m \vDash \varphi_m \to \theta, \Delta_m \vDash \neg\theta,$$

$$\Gamma_\omega \vDash \varphi_m \to \theta, \Delta_\omega \vDash \neg\theta.$$

Аналогично для некоторого σ_0'-предложения θ' верно

$$\Gamma_\omega \vDash \neg\varphi_m \to \theta', \Delta_\omega \vDash \neg\theta'.$$

Вместе эти два утверждения влекут

$$\Gamma_\omega \vDash \theta \vee \theta', \Delta_\omega \vDash \neg(\theta \vee \theta'),$$

что противоречит отделимости.

Для теории Δ_ω полнота доказывается аналогично.
Чтобы избежать отделимости, теории Γ_ω и Δ_ω должны содержать одни и те же σ_0'-предложения, поэтому $\Gamma_\omega \cap \Delta_\omega$ будет полной теорией в сигнатуре σ_0'.
Заметим, что для каждого функционального или константного символа $f \in \sigma_1'$ и вектора $\overline{c} \in C^{Arity(f)}$ к теории Γ_ω в какой-то момент присоединилась формула $\exists x\ f(\overline{c}) = x$ (если нет, то из полноты присоединилось отрицание этой формулы, а тогда итоговая теория не может иметь моделей). Вместе с ней туда попала и формула $f(\overline{c}) = d$ для некоторой константы $d \in C$. Это означает, что в теории Γ_ω множество констант обязательно замкнуто под действием функциональных символов, а все старые константы равны каким-то новым. Для Δ_ω рассуждения аналогичны.

Возьмём теперь любую модель $\mathfrak{A}$ теории Γ_ω. Из замкнутости от её носителя можно оставить только образы констант из C и снова получить модель в сигнатуре σ_1'. Заметим, что эта подмодель $\mathfrak{A}'$ всё ещё служит моделью Γ_ω. Это можно понять индукцией по сложности предложения.

- Предложения, зависящие только от констант, сохранили истинность.
- Предложения, полученные при помощи логических связок, получились из более простых предложений, истинность которых не поменялась.
- Ложные предложения вида $\exists x\ \Phi(x, \overline{y})$ не могли стать истинными, а у истинных по построению есть свидетель c для x из множества C. Предложение $\Phi(c, \overline{y})$ осталось истинно, поэтому истинным будет и более сложное.
- Наконец, предложения вида $\forall x\ \Phi$ можно переписать как $\neg\exists x\ \neg\Phi$ и применить предыдущие доводы.

Проведём ту же процедуру с какой-нибудь моделью $\mathfrak{B}$ теории Δ_ω и оставим от неё подмодель $\mathfrak{B}'$. У $\mathfrak{A}'$ и $\mathfrak{B}'$ изоморфны L_0'-обеднения, потому что эту изоморфность можно описать средствами σ_0'. Тогда носители $\mathfrak{A}'$ и $\mathfrak{B}'$ можно отождествить (так, чтобы образы одной и той же константы были сопоставлены друг другу). То, что получилось, будет моделью теории $\Gamma_\omega \cup \Delta_\omega$. А это невозможно, потому что в этой модели истинно $\varphi \wedge \neg\psi$, и φ не могло влечь ψ.

Определение отношений

Поговорим теперь о применениях теоремы Крейга. Первое из них касается способов определения отношений. Пусть P и P' — два n-местных предикатных символа, не входящие в сигнатуру σ, а $\Gamma(P)$ — какое-то множество предложений сигнатуры $\sigma \cup \{P\}$, а $\Gamma(P')$ — его аналог, получающийся заменой всех вхождений P на P'. Будем говорить, что $\Gamma(P)$ **определяет** P **неявно**, если

$$\Gamma(P) \cup \Gamma(P') \vDash \forall x_1, \ldots, x_n (P(x_1, \ldots, x_n) \leftrightarrow P'(x_1, \ldots, x_n)),$$

где следствие берётся в сигнатуре $\sigma \cup \{P, P'\}$. Множество $\Gamma(P)$ **явно определяет** P, если для некоторой формулы $\varphi(x_1, \ldots, x_n)$ сигнатуры σ верно

$$\Gamma(P) \vDash \forall x_1, \ldots, x_n (P(x_1, \ldots, x_n) \leftrightarrow \varphi(x_1, \ldots, x_n)).$$

Теорема Бета

Легко видеть, что если $\Gamma(P)$ явно определяет P, то оно определяет его и неявно. Оказывается, верно и обратное.

Теорема

Множество $\Gamma(P)$ явно определяет P тогда и только тогда, когда оно определяет его неявно.

Доказательство. Пусть $\Gamma(P)$ неявно определяет P. Добавим к σ n новых констант $c_1, \dots, c_n$. Из компактности можно считать, что $\Gamma(P)$ конечно. Тогда его можно заменить на одну конъюнкцию всех предложений — $\psi(P)$. Итак, имеем

$$\psi(P) \wedge \psi(P') \vDash P(c_1, \dots, c_n) \to P'(c_1, \dots, c_n),$$

$$\psi(P) \wedge P(c_1, \dots, c_n) \vDash \psi(P') \to P'(c_1, \dots, c_n).$$

По теореме Крейга найдётся промежуточное предложение $\theta(c_1, \dots, c_n)$, не содержащее ни P, ни P', что

$$\psi(P) \wedge P(c_1, \dots, c_n) \vDash \theta(c_1, \dots, c_n);$$

$$\psi(P) \vDash P(c_1, \dots, c_n) \to \theta(c_1, \dots, c_n);$$

$$\theta(c_1, \dots, c_n) \vDash \psi(P') \to P'(c_1, \dots, c_n);$$

$$\theta(c_1, \dots, c_n) \vDash \psi(P) \to P(c_1, \dots, c_n).$$

Получилось, что

$$\psi(P) \vDash P(c_1, ..., c_n) \leftrightarrow \theta(c_1, ..., c_n).$$

Заменив здесь константы на кванторы по второй лемме, мы видим, что множество $\{\psi(P)\}$ явно определяет отношение P при помощи формулы θ. Поскольку при расширении множества предложений оно не перестаёт определять P (как явно, так и неявно), это же было верно и для исходного множества $\Gamma(P)$.

Теорема Робинсона о непротиворечивости

Второе применение теоремы Крейга касается расширений полных теорий. Получившаяся теорема Робинсона — это обобщение того, что мы делали на последнем шаге доказательства теоремы Крейга.

Теорема

Пусть σ_1 *и* σ_2 *— две сигнатуры,* $\sigma_0 = \sigma_1 \cap \sigma_2$. *Предположим, что* Γ *— полная теория в сигнатуре* σ_0, Γ_1 *и* Γ_2 *— её непротиворечивые расширения в* σ_1 *и* σ_2 *соответственно. Тогда теория* $\Gamma_1 \cup \Gamma_2$ *непротиворечива в сигнатуре* $\sigma_1 \cup \sigma_2$.

Доказательство. Если $\Gamma_1 \cup \Gamma_2$ противоречива, то от них можно оставить конечные подмножества Δ_1 и Δ_2 так, чтобы $\Delta_1 \cup \Delta_2$ осталось непротиворечиво. Пусть σ_1 — конъюнкция всех предложений из Δ_1, а σ_2 — конъюнкция всех предложений из Δ_2. Из противоречивости $\sigma_1 \vDash \neg\sigma_2$. По теореме Крейга найдётся промежуточное предложение θ в сигнатуре $\sigma_1 \cap \sigma_2 = \sigma_0$, что $\sigma_1 \vDash \theta, \theta \vDash \neg\sigma_2$. Тогда, с одной стороны, $\Gamma_1 \vDash \theta$, а с другой, $\sigma_2 \vDash \neg\theta$ и $\Gamma_2 \vDash \neg\theta$. Это противоречит тому, что Γ_1 и Γ_2 являются расширением одной и той же полной σ_0-теории.

Что ещё можно сделать с теоремой Крейга?

Много других приложений интерполяционной теоремы есть в https://math.stanford.edu/~feferman/papers/craigtransps.pdf. Мы взглянем на самые простые из них (и ещё одно).

Мы рассмотрели интерполяционную теорему для символов сигнатуры. Можно искать интерполянты по отношению к другим свойствам, например: свободным переменным, позитивно и негативно входящим предикатам, переменным, по которым ведутся кванторы, ... Один из интересных результатов в этой области — теорема Эрбранда:

Теорема

*Если φ — **универсальная формула** (т.е. все кванторы после вынесения наружу становятся кванторами $\forall$), а ψ — **экзистенциальная формула** (то же для $\exists$), причём $\varphi \vDash \psi$, то у φ и ψ есть бескванторный интерполянт.*

Применение для классов моделей

Проективный класс моделей — это класс K всех моделей, в которых верна формула $\exists\overline{x}\varphi(\overline{x},\overline{y})$, где φ не зависит от модели. **Элементарный класс** — это класс моделей, в которых верно конкретное предложение $\theta(\overline{y})$. В новых терминах теорему Крейга можно переписать так:

Теорема

Любые два непересекающихся проективных класса можно разделить элементарным.

Это похоже на теорему Крейга, потому что отсутствие пересечения можно записать в виде $\varphi(\overline{x}, \overline{y}) \to \neg\psi(\overline{z}, \overline{y})$. У этой теоремы есть

Следствие

Δ-интерполяционная теорема: класс моделей элементарен тогда и только тогда, когда он и его дополнение проективны.

Применение в model checking

Последнее в этом докладе применение интерполяции связано с верификацией моделей. Рассмотрим модель $\mathcal{M}$, состоящую из множества состояний S, некоторые из которых начальные и образуют множество I, и отношения перехода T. Пусть φ_I и φ_T — формулы, определяющие I и T соответственно.

Мы хотим проверить, можно ли из начального состояния достичь такого, в котором верна формула ψ. Если это делать по-честному, придётся вычислять образ I под действием степеней T, а это сложно.
Идея: рассмотрим вместо оператора образа его аппроксимацию сверху. Если даже под действием его из I никогда не попасть туда, где верна ψ, то это невозможно и под действием образа.
Такой аппроксимацией может выступать интерполянт Крейга (который, как легко показать, в пропозициональной логике вычисляется явно за экспоненциальное время)

Достижимость ψ за k шагов можно записать как-то наподобие

$$\varphi_I(s_0) \wedge \varphi_T(s_0, s_1) \wedge \ldots \wedge \varphi_T(s_{k-1}, s_k) \wedge (\psi(s_0) \vee \ldots \vee \psi(s_k)).$$

Если выделить отсюда $A = \varphi_I(s_0) \wedge \varphi_T(s_0, s_1)$ и $B = \neg$ [всё остальное], то можно исключительно при помощи $\varphi_I, \varphi_T, \sigma_0$ и σ_1 описать интерполянт Крейга $A'(s_0, s_1)$. Если заменить отношение перехода на этот интерполянт, можно проще пересчитать множество достижимых состояний, а ложное срабатывание случится не раньше, чем через k шагов.

Спасибо!
Вопросы?