

Основы математической логики и логического программирования

ЛЕКТОР: В.А. Захаров

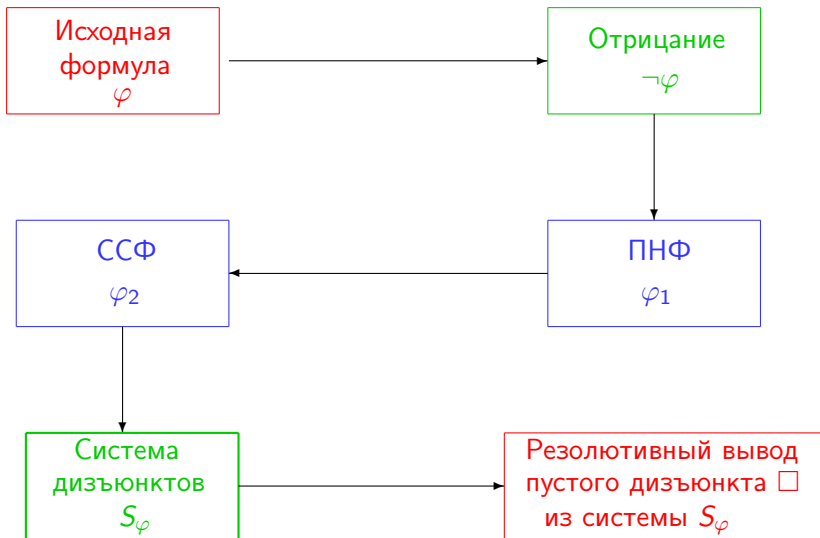
Лекция 7.

Эрбрановские интерпретации.

Теорема Эрбрана.

Задача унификации.

ОБЩАЯ СХЕМА МЕТОДА РЕЗОЛЮЦИИ



ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Проверка общезначимости формулы φ сводится к проверке противоречивости системы дизъюнктов S_φ .

Этап 1. Сведение проблемы общезначимости к проблеме противоречивости: $\varphi \rightsquigarrow \varphi_0 = \neg\varphi$

Этап 2. Построение предваренной нормальной формы (ПНФ).

$$\varphi_0 \rightsquigarrow \varphi_1 = Q_1x_1Q_2x_2 \dots Q_nx_n (D_1 \& D_2 \& \dots \& D_N)$$

Этап 3. Построение сколемовской стандартной формы (ССФ).

$$\varphi_1 \rightsquigarrow \varphi_2 = \forall x_{i_1} \forall x_{i_2} \dots \forall x_{i_k} (D_1 \& D_2 \& \dots \& D_N)$$

Этап 4. Построение системы дизъюнктов.

$$\varphi_2 \rightsquigarrow S_\varphi = \{D_1, D_2, \dots, D_N\},$$

φ общезначимая \iff система дизъюнктов S_φ противоречива.

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Система дизъюнктов $S = \{D_1, D_2, \dots, D_N\}$ противоречива тогда и только тогда, когда

для каждой интерпретации I

в системе S найдется такой дизъюнкт

$$D_j = \forall x_1 \dots \forall x_n (L_{1j} \vee L_{2j} \vee \dots \vee L_{k_j j})$$

и в предметной области D_j найдется такой набор элементов

$$d_1, \dots, d_n,$$

для которых имеют место

$$I \not\models L_{1j}[d_1, \dots, d_n], I \not\models L_{2j}[d_1, \dots, d_n], \dots, I \not\models L_{k_j j}[d_1, \dots, d_n].$$

А можно ли сократить множество рассматриваемых интерпретаций?

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Эрбрановские интерпретаций (H -интерпретации, названные так по имени французского математика **Herbrand**) — это специальная разновидность интерпретаций, в основе которых лежат **свободные алгебры**.

Предметная область эрбрановских интерпретаций называется **эрбрановским универсумом** (H -универсумом).

Определение H -универсума

Пусть задана некоторая сигнатура $\sigma = \langle Const, Func, Pred \rangle$.

Тогда **эрбрановским универсумом** σ называется множество

термов $H_\sigma = \bigcup_{i=0}^{\infty} H_i$, где

$$i = 0 \quad H_0 = \begin{cases} Const, & \text{если } Const \neq \emptyset, \\ \{c\}, & \text{если } Const = \emptyset \text{ (эрбрановская константа)}; \end{cases}$$

$$i \rightarrow i + 1 \quad H_{i+1} = H_i \cup \{f^{(k)}(t_1, \dots, t_k) : f^{(k)} \in Func, t_1, \dots, t_k \in H_i\}.$$

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Эрбрановский универсум — это множество всех термов, которые можно построить из констант и функциональных символов заданной сигнатуры. Термы эрбрановского универсума не содержат переменных и называются **основными термами** .

Пример

Пусть $Const = \emptyset$, $Func = \{f^{(1)}, g^{(2)}\}$. Тогда

$i = 0$ $H_0 = \{c\}$ (эрбрановская константа, т.к. $Const = \emptyset$);

$i = 1$ $H_1 = \{c, f(c), g(c, c)\}$;

$i = 2$ $H_2 = \{c, f(c), g(c, c),$
 $f(f(c)), f(g(c, c)), g(f(c), c), g(c, f(c)),$
 $g(f(c), f(c)), g(c, g(c, c)), g(g(c, c), c),$
 $g(g(c, c), g(c, c)), g(f(c), g(c, c)), g(g(c, c), f(c))\}$;

$i = 3$ и. т. д.

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Определение H -интерпретации

Эрбрановская интерпретация $I_H = \langle H_\sigma, \overline{Const}_H, \overline{Func}_H, \overline{Pred} \rangle$ сигнатуры $\sigma = \langle Const, Func, Pred \rangle$ состоит из

- ▶ **стандартной** предметной области — эрбрановского универсума H_σ ;
- ▶ **стандартной** оценки констант: $\overline{Const}_H(c) = c$, т. е. значением каждого константного символа c является его собственное изображение;
- ▶ **стандартной** оценки функциональных символов: $\overline{Func}_H(f^{(n)}) = \mathbf{f} : \mathbf{f}(t_1, t_2, \dots, t_n) = f^{(n)}(t_1, t_2, \dots, t_n)$, т. е. каждый функциональный символ f играет роль конструктора термов эрбрановского универсума;
- ▶ **произвольной** оценки предикатных символов.

Таким образом, разные H -интерпретации отличаются только истолкованием предикатных символов.

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Преимущества эрбрановских интерпретаций.

- ▶ Не нужно заботиться о выборе области интерпретации — у всех H -интерпретаций одна и та же предметная область — H -универсум.
- ▶ Не нужно заботиться об оценке функциональных символов — у всех H -интерпретаций одна и та же стандартная оценка $Const$ и $Func$.
- ▶ И, как будет показано, для проверки противоречивости систем дизъюнктов достаточно ограничиться рассмотрением H -интерпретаций.

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Теорема о H -интерпретациях

Система дизъюнктов S выполнима тогда и только тогда, когда S имеет эрбрановскую модель, т.е. выполнима хотя бы в одной H -интерпретации.

Доказательство

(\Leftarrow) Очевидно.

(\Rightarrow) Пусть $S = \{D_1, \dots, D_N\}$, и $I = \langle \mathcal{D}_I, \overline{Const}_I, \overline{Func}_I, \overline{Pred}_I \rangle$ — некоторая модель для S , т. е. для любого дизъюнкта $D_i = \forall x_1 \dots \forall x_n (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})$ из S и для любого набора элементов d_1, \dots, d_n из \mathcal{D}_I имеет место

$$I \models (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})[d_1, \dots, d_n].$$

Покажем, что существует H -интерпретация J_H , в которой выполняется каждый дизъюнкт D_i из S .

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Доказательство

Пусть σ — сигнатура системы дизъюнктов S . Рассмотрим эрбрановский универсум H_σ и отображение

$$\alpha : H_\sigma \rightarrow \mathcal{D}_I,$$

которое каждому основному терму $t \in H_\sigma$ сопоставляет элемент $\alpha(t) = d_t \in \mathcal{D}_I$, равный значению терма t в интерпретации I .

Оценку \overline{Pred}_H предикатных символов в H -интерпретации J определим так:

для любого предикатного символа P и любого набора основных термов $t_1, \dots, t_m \in H_\sigma$ положим

$$\overline{P}_H(t_1, \dots, t_m) = \mathbf{true} \iff \overline{P}_I(\alpha(t_1), \dots, \alpha(t_m)) = \mathbf{true}.$$

Эта оценка однозначно определяет H -интерпретацию J_H

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Доказательство

Тогда для любого функционального символа f , предикатного символа P и любого набора основных термов $t_1, \dots, t_m \in H_\sigma$ верно

$$\alpha(f(t_1, \dots, t_m)) = \bar{f}_I(\alpha(t_1), \dots, \alpha(t_m)),$$

$$J_H \models P[t_1, \dots, t_m] \iff I \models P[\alpha(t_1), \dots, \alpha(t_m)]$$

Отображение α — это **гомоморфизм** интерпретации J_H в интерпретацию I .

Значит, для **любого** дизъюнкта $D' = \forall x_1 \dots \forall x_m (L_1 \vee \dots \vee L_k)$ и любого набора основных термов $t_1, \dots, t_m \in H_\sigma$ верно

$$J_H \models (L_1 \vee \dots \vee L_k)[t_1, \dots, t_m]$$



$$I \models (L_1 \vee \dots \vee L_k)[\alpha(t_1), \dots, \alpha(t_m)]$$

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Доказательство

Итак, для любого дизъюнкта

$D_i = \forall x_1 \dots \forall x_n (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})$ из S и любого набора основных термов $t_1, \dots, t_n \in H_\sigma$ имеем

$$J_H \models (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})[t_1, \dots, t_n]$$



$$I \models (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})[\alpha(t_1), \dots, \alpha(t_n)].$$

Поскольку

$$I \models (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})[\alpha(t_1), \dots, \alpha(t_n)],$$

все дизъюнкты из S выполнимы в построенной эрбрановской интерпретации J_H . □

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Следствие

Система дизъюнктов S противоречива тогда и только тогда, когда S невыполнима ни в одной эрбрановской интерпретации.

Значит, для проверки противоречивости систем дизъюнктов достаточно исследовать только эрбрановские интерпретации.

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Как задавать H -интерпретации?

Эрбрановские интерпретации полностью определяются оценкой предикатных символов. Значит, достаточно указать оценку всех предикатных символов.

Пусть $P^{(m)} \in \text{Pred}$, и $t_1, \dots, t_m \in H_\sigma$ — набор основных термов. Тогда формула $P^{(m)}(t_1, \dots, t_m)$ называется **основным атомом**. Множество всех основных атомов называется **эрбрановским базисом** и обозначается B_H .

Всякая H -интерпретация I задается подмножеством B^I истинных основных атомов:

$$B^I = \{P^{(m)}(t_1, \dots, t_m) : I \models P^{(m)}(t_1, \dots, t_m), \{t_1, \dots, t_m\} \subseteq H\}.$$

ЭРБРАНОВСКИЕ ИНТЕРПРЕТАЦИИ

Примеры

- ▶ $B^I = \emptyset$ соответствует интерпретации I , в которой все \bar{P} тождественно ложны.

$$I \models P^{(m)}[t_1, \dots, t_m] \iff P^{(m)}(t_1, \dots, t_m) \in \emptyset.$$

- ▶ $B^I = B_H$ соответствует интерпретации I , в которой все \bar{P} тождественно истинны.
- ▶ Пересечение $B^{I_1} \cap B^{I_2}$ задает интерпретацию, в которой истинны те и только те отношения, которые истинны в обеих интерпретациях I_1 и I_2 .

Значит, это очень удобный способ задания интерпретаций, позволяющий проводить над ними теоретико-множественные операции.

В дальнейшем все H -интерпретации мы будем ассоциировать с подмножествами эрбрановского базиса B_H .

ТЕОРЕМА ЭРБРАНА

Вернемся к системам дизъюнктов

Пусть имеется дизъюнкт $D = \forall x_1 \dots \forall x_m (L_1 \vee \dots \vee L_k)$ и набор основных термов t_1, \dots, t_m из эрбрановского универсума H_σ .

Тогда дизъюнкт $D' = (L_1 \vee \dots \vee L_k)\{x_1/t_1, \dots, x_m/t_m\}$, полученный из D подстановкой основных термов t_1, \dots, t_m вместо всех переменных дизъюнкта D называется **основным примером** дизъюнкта D .

Пример

$D = \forall x_1 \forall x_2 (\neg R(x_1, x_2) \vee P(h(x_1)))$ — дизъюнкт,

$D' = \neg R(g(c'), c'') \vee P(h(g(c'))) = D\{x_1/g(c'), x_2/c''\}$ — основной пример.

ТЕОРЕМА ЭРБРАНА

Система дизъюнктов $S = \{D_1, D_2, \dots, D_N\}$ противоречива



для каждой ***H*-интерпретации** I в S найдется такой дизъюнкт $D_i = (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})$ и такой набор основных термов t_1, \dots, t_m , для которых имеют место

$$I \not\models (L_{1i} \vee L_{2i} \vee \dots \vee L_{k_i i})[t_1, \dots, t_m]$$



для каждой ***H*-интерпретации** I существует основной пример $D' = D_i\{x_1/t_1, \dots, x_m/t_m\}$ дизъюнкта из системы S , для которого

$$I \not\models D'.$$

ТЕОРЕМА ЭРБРАНА

Рассмотрим множество основных примеров дизъюнктов

$$\mathcal{G}_S = \{D' : D' \text{ — основной пример дизъюнкта из } S, \\ \text{существует эрбрановская интерпретация } I \subseteq B_H : \\ I \not\models D'\}.$$

Тогда система дизъюнктов S противоречива



система основных примеров \mathcal{G}_S противоречива.

ТЕОРЕМА ЭРБРАНА

Вспомним теорему компактности Мальцева

Множество замкнутых формул T имеет модель тогда и только тогда, когда каждое **конечное** подмножество T' , $T' \subseteq T$, имеет модель.

Таким образом, множество основных примеров дизъюнктов \mathcal{G}_S противоречиво



существует **конечное** противоречивое подмножество \mathcal{G}' , $\mathcal{G}' \subseteq \mathcal{G}_S$.

И в результате проведенных рассуждений мы приходим к
Теореме Эрбрана

ТЕОРЕМА ЭРБРАНА

Теорема Эрбрана

Система дизъюнктов $S = \{D_1, \dots, D_N\}$ противоречива



существует конечное противоречивое множество \mathcal{G}'
основных примеров дизъюнктов системы S .

В чем состоит главная особенность теоремы Эрбрана?

Основной пример дизъюнкта не содержит **предметных переменных**, и поэтому является простой **булевой формулой**

Таким образом проблема общезначимости формул логики предикатов (сложная проблема!!!) сводится к проблеме выполнимости булевой формулы (простой проблеме??!!).

ТЕОРЕМА ЭРБРАНА

Маленькое лукавство.

Увы, теорема Эрбрана не сообщает нам ничего о том, КАКИЕ основные примеры дизъюнктов образуют это противоречивое множество \mathcal{G}' .

Нам нужно придумать механизм поиска этого противоречивого множества \mathcal{G}' .

Дэвис и Патнем предложили использовать компьютер для перебора всех эрбрановских интерпретаций в поиске противоречивой системы основных примеров дизъюнктов. Но Дж. Робинсон поступил хитрее...

Из дизъюнктов системы S нужно устранять потенциальные противоречия, пока не будут устранены все источники противоречия или не будет получено неустранимое (явное) противоречие.

ТЕОРЕМА ЭРБРАНА

Если в системе S содержатся дизъюнкты $D_1 = L$ и $D_2 = \neg L$, то, очевидно, S противоречива (явное противоречие).

А если в S есть дизъюнкты $D_1 = D'_1 \vee L$ и $D_2 = D'_2 \vee \neg L$, то $I \models D_1$ и $I \models D_2$ влечет $I \models D'_1 \vee D'_2$. Значит, добавление дизъюнкта $D_0 = D'_1 \vee D'_2$ к S не нарушает ее выполнимости. Зато устраняется потенциальное противоречие L и $\neg L$.

Правило вывода, разрешающее противоречия,

$$\frac{D'_1 \vee L, D'_2 \vee \neg L}{D'_1 \vee D'_2}$$

называется **правилом резолюции**. Это правило можно применять до тех пор, пока не возникнет неустранимое противоречие $D_1 = L$ и $D_2 = \neg L$. Это и будет служить признаком противоречивости системы S .

ТЕОРЕМА ЭРБРАНА

Ну, а если $D_1 = P(x)$ и $D_2 = \neg P(f(y))$, то как быть тогда?
Ведь здесь нет явного противоречия.

Противоречие здесь скрытое. Дизъюнкты D_1 и D_2 имеют основные примеры

$$P(f(c)) = D_1\{x/f(c)\} \text{ и } \neg P(f(c)) = D_2\{y/c\},$$

которые образуют противоречие. Значит, D_1 и D_2 тоже противоречивы. Но как отыскать это скрытое противоречие?

Нужно постараться привести литеры $P(x)$ и $P(f(y))$ к общему виду. Приведение выражений к общему виду называется **унификацией**.

Нам нужен алгоритм унификации!

ЗАДАЧА УНИФИКАЦИИ

Приведение двух атомов A' и A'' к общему виду достигается применением такой подстановки θ , для которой $A'\theta = A''\theta$.
Значит, для решения задачи унификации придется изучить алгебраические свойства подстановок.

Воспоминания

Подстановка — это отображение $\theta : Var \rightarrow Term$.

Конечная подстановка $\theta = \{x_1/t_1, x_2/t_2, \dots, x_n/t_n\}$.

$E\theta$ — **применение подстановки** θ к выражению E .

ЗАДАЧА УНИФИКАЦИИ

Композиция подстановок

Пусть $\theta, \eta \in \text{Subst}$. **Композиция подстановок** $\theta\eta$ — это подстановка μ , которая определяется следующим соотношением:

$$\mu(x) = (x\theta)\eta \text{ для любой } x \in \text{Var}.$$

Утверждение

Пусть $\theta = \{x_1/t_1, \dots, x_n/t_n\}$, $\eta = \{y_1/s_1, \dots, y_m/s_m\}$. Тогда подстановка

$$\begin{aligned} \mu = & \{x_1/t_1\eta, \dots, x_n/t_n\eta\} \cup \\ & \{y_i/s_i : y_i \notin \{x_1, x_2, \dots, x_n\}\} - \\ & \{x_j/t_j\eta : x_j = t_j\eta\}. \end{aligned}$$

является композицией $\theta\eta$.

ЗАДАЧА УНИФИКАЦИИ

Доказательство

$$\begin{aligned} \mu = & \{x_1/t_1\eta, \dots, x_n/t_n\eta\} \cup \\ & \{y_i/s_i : y_i \notin \{x_1, x_2, \dots, x_n\}\} - \\ & \{x_j/t_j\eta : x_j = t_j\eta\}. \end{aligned}$$

Рассмотрим произвольную $z \in \text{Var}$. Возможны 3 варианта.

1. $z = x_j \in \text{Dom}_\theta$. Тогда $z(\theta\eta) = (x_j\theta)\eta = t_j\eta$.

Если $t_j\eta \neq x_j$, то $x_j \in \text{Dom}_\mu$, и $x_j\mu = t_j\eta$.

Если $t_j\eta = x_j$, то $x_j \notin \text{Dom}_\mu$, и $x_j\mu = x_j$.

В любом случае $x_i(\theta\eta) = x_i\mu$.

2. $z = y_i \in \text{Dom}_\eta \setminus \text{Dom}_\theta$. Тогда $z(\theta\eta) = (y_i\theta)\eta = y_i\eta = s_i$,

и $z\mu = s_i$.

3. $z = y_i \notin \text{Dom}_\eta \cup \text{Dom}_\theta$. Тогда $z(\theta\eta) = z$ и $z\mu = z$. □

ЗАДАЧА УНИФИКАЦИИ

Пример

$$\theta = \{x/f(x, c), y/g(u), z/y\},$$

$$\eta = \{x/g(y), y/z, u/c\}.$$

$$\theta\eta =$$

ЗАДАЧА УНИФИКАЦИИ

Пример

$$\theta = \{x/f(x, c), y/g(u), z/y\},$$

$$\eta = \{x/g(y), y/z, u/c\}.$$

$$\theta\eta =$$

ЗАДАЧА УНИФИКАЦИИ

Пример

$$\theta = \{x/f(x, c), y/g(u), z/y\},$$

$$\eta = \{x/g(y), y/z, u/c\}.$$

$$\theta\eta = \{x/f(x, c)\eta, y/g(u)\eta, z/y\eta\} \cup \{u/c\}$$

ЗАДАЧА УНИФИКАЦИИ

Пример

$$\theta = \{x/f(x, c), y/g(u), z/y\},$$

$$\eta = \{x/g(y), y/z, u/c\}.$$

$$\theta\eta = \{x/f(x, c)\eta, y/g(u)\eta, z/y\eta\} \cup \{u/c\} \\ \{x/f(g(y), c), y/g(c), z/z, u/c\}$$

ЗАДАЧА УНИФИКАЦИИ

Пример

$$\theta = \{x/f(x, c), y/g(u), z/y\},$$

$$\eta = \{x/g(y), y/z, u/c\}.$$

$$\begin{aligned}\theta\eta &= \{x/f(x, c)\eta, y/g(u)\eta, z/y\eta\} \cup \{u/c\} \\ &\quad \{x/f(g(y), c), y/g(c), z/z, u/c\} \\ &\quad \{x/f(g(y), c), y/g(c), u/c\}\end{aligned}$$

ЗАДАЧА УНИФИКАЦИИ

Определение

Пусть E_1 и E_2 — два логических выражения (термы, атомы, формулы и др.)

Подстановка θ называется **унификатором** выражений E_1 и E_2 , если $E_1\theta = E_2\theta$.

Подстановка θ называется **наиболее общим унификатором (НОУ)** выражений E_1 и E_2 , если

1. θ — унификатор выражений E_1 и E_2 ;
2. для любого унификатора η выражений E_1 и E_2 существует такая подстановка ρ , для которой верно равенство

$$\eta = \theta\rho$$

$НОУ(E_1, E_2)$ — множество наиболее общих унификаторов выражений E_1 и E_2 .

ЗАДАЧА УНИФИКАЦИИ

Пример

$$E_1 = P(f(x_1), x_2)$$

$$E_2 = P(y_1, c)$$

$\eta = \{x_1/f(c), x_2/c, y_1/f(f(c))\}$ — унификатор E_1 и E_2 , т.к.

$$E_1\eta = P(f(f(c)), c) = E_2\eta.$$

$\theta = \{x_2/c, y_1/f(x_1)\}$ — наиболее общий унификатор E_1 и E_2 , т.к.

$$E_1\theta = P(f(x_1), c) = E_2\theta,$$
$$\eta = \theta\{x_1/f(c)\}.$$

Выражения E_1 и E_2 **унифицируемы**, и $\theta \in \text{НОУ}(E_1, E_2)$.

ЗАДАЧА УНИФИКАЦИИ

Пример

$$E_1 = R(f(x_1), x_1)$$

$$E_2 = R(y_1, f(y_1))$$

Выражения E_1 и E_2 не имеют унификаторов, и
 $НОУ(E_1, E_2) = \emptyset$.

ЗАДАЧА УНИФИКАЦИИ

Задача унификации

состоит в том, чтобы для двух выражений E_1 и E_2 выяснить, являются ли эти выражения унифицируемыми, и, в случае их унифицируемости, вычислить наиболее общий унификатор.

КОНЕЦ ЛЕКЦИИ 7.