

Змістовий модуль 3. Теорема Ербрана

Тема 7. Ербранівський універсум множини диз'юнктив

План лекції.

- Алгоритмічна нерозв'язність задачі виявлення загальнозначущості формул у логіці першого порядку
- Стандартне подання формул
- Ербранівський універсум і ербранівський базис
- Основний приклад диз'юнкта
- H -інтерпретація
- Важливі твердження щодо H -інтерпретації

3.1. Алгоритмічна нерозв'язність задачі виявлення загальнозначущості формул у логіці першого порядку

Ми побачили, як розв'язувати задачі способом доведення теорем. Зараз ми починаємо розглядати процедури пошуку доведення. Насправді пошук загальної процедури для перевірки загальнозначущості формули розпочався давно. Першим намагався знайти таку процедуру Лейбніц (1646 – 1716 р.р.), у подальшому зробили нові спроби Пеано (на самому початку XX століття) і школа Гільберта у 1920-х роках.

Ці спроби продовжувалися допоки Чорч і Тьюрінг у 1936 р. незалежно один від одного довели, що не існує алгоритму для перевірки загальнозначущості формул у логіці першого порядку. Проте існують алгоритми пошуку доведення, які можуть підтвердити, що формула загальнозначуща, якщо вона насправді такою є. Для незагальнозначущих формул ці алгоритми, назагал, не закінчують свою роботу. Беручи до уваги результат Чорча й Тьюрінга, це найкраще, чого ми можемо очікувати від алгоритму пошуку доведення.

Дуже важливий підхід до автоматичного доведення теорем був даний Ербраном (Herbrand J.) у 1930 р. За означенням загальнозначуща формула – це формула, яка є істинною в усіх інтерпретаціях. Ербран розробив алгоритм пошуку інтерпретації, котра *спростовує* дану формулу (тобто у цій інтерпретації формула набуває значення F). Проте, якщо дана формула й справді загальнозначуща, то такої інтерпретації не існує й алгоритм закінчує роботу за скінченну кількість кроків. Метод Ербрана слугує основою для більшості алгоритмів пошуку доведення. Гілмор одним із перших реалізував процедуру Ербрана на комп'ютері. Оскільки формула загальнозначуща тоді й тільки тоді, коли її заперечення є невиконуваною формулою, то його програма призначена виявляти невиконуваність заперечення даної формули. Програма Гілмора справилася з доведенням декількох простих формул, але під час доведення інших формул логіки першого порядку виникли великі труднощі. Девіс і Патнем удосконалили метод Гілмора, але він все ще був незручним для практичної реалізації. І лише у 1965 році Дж. Робінсон (Robinson J.A.) запропонував метод резолюцій, який практично застосовують як метод пошуку доведення теорем.

Головним застосуванням логіки першого порядку у комп'ютерних інформаційних технологіях є автоматизація **пошуку доведень** теорем. Ця проблема актуальна при розв'язуванні задач штучного інтелекту, задач, які виникають у базах даних і знань, при вирішенні питань представлення знань і роботи з ними.

3.2. Стандартне подання формул

Процедури пошуку доведень за Ербраном та за методом резолюцій, які ми розглядатимемо нижче, насправді є процедурами пошуку спростування, тобто замість доведення загальнозначущості формули доводиться, що заперечення формули є невиконувана формула. Це тільки питання зручності: при використанні процедур спростування не відбувається втрати загальності. Окрім того, ці процедури спростування застосовують до формули у сколемівській нормальній формі, яку також називають сколемівською стандартною формою. Ця стандартна форма була уведена Мартіном Девісом і Гіларі Патнемом [14] у 1960 р. Вони довели, що, що будь-яке висловлення в логіці першого порядку (без використання рівності) можна записати у такій формі. Нагадаємо, як отримується ця форма (див. п. 2.14).

1. Формулу логіки першого порядку зводять до випередженої нормальної форми (префікс є послідовністю кванторів а матриця не містить ніяких кванторів).

2. Матрицю зводять до кон'юнктивної нормальної форми (це завжди можна зробити, бо матриця не містить кванторів).

3. Елімінують (вилучають) квантори існування шляхом використання сколемівських функцій. (При цьому невиконуваність формули зберігається.)

У результаті отримуємо сколемівську нормальну форму.

Раніше (див приклад 2.47) для формули

$$\forall x \exists y \exists z ((\neg P(x, y) \wedge G(x, z)) \vee R(x, y, z)).$$

ми отримали таку стандартну форму:

$$\forall x ((\neg P(x, f(x)) \vee R(x, f(x), g(x))) \wedge (G(x, g(x)) \vee R(x, f(x), g(x)))).$$

Матриця цієї стандартної форми складається з двох диз'юнктив. У подальшому стандартну форму будемо подавати у вигляді множини диз'юнктив S . У нашому прикладі

$$S = \{ \neg P(x, f(x)) \vee R(x, f(x), g(x)), G(x, g(x)) \vee R(x, f(x), g(x)) \}.$$

Надалі вважатимемо, що множина диз'юнктив S подає кон'юнкцію всіх диз'юнктив із S , де кожна предметна змінна в S є під дією квантора загальності.

3.3. Ербранівський універсум і ербранівський базис

У подальшому істотно використовується співвідношення між суперечністю і логічним наслідком, сформульоване у теоремі 1.7 (див. п. 1.12 та п. 2.7). Нагадаємо цей результат: доведення теореми $(A_1 \wedge A_2 \wedge \dots \wedge A_k) \rightarrow B$ еквівалентне доведенню того, що «розширена» множина формул $\{A_1, A_2, \dots, A_k, \neg B\}$, утворена об'єднанням множини аксіом $\{A_1, A_2, \dots, A_k\}$ і заперечення цільового твердження $\{\neg B\}$, невиконувана. Для застосування алгоритмів пошуку доведення теорем кожен з формул множини $\{A_1, A_2, \dots, A_k, \neg B\}$ зводять до сколемівської стандартної форми; в отриманих формах викреслюють квантори загальності і символи операції кон'юнкції. Як результат отримують множину диз'юнктив $S = \{D_1, D_2, \dots, D_m\}$, у якій всі предметні змінні вважаються зв'язаними кванторами загальності.

За означенням множина диз'юнктив S невиконувана тоді й тільки тоді, коли для будь-якої інтерпретації в множині S знайдеться диз'юнкт, фальшивий при цій інтерпретації. Оскільки незручно й неможливо розглядати всі інтерпретації на всіх предметних областях, було би зручно, якби ми могли зафіксувати одну таку спеціальну область H , що множина S невиконувана тоді й тільки тоді, коли вона фальшива в усіх інтерпретаціях на цій області. На щастя, така область існує. Її називають ербранівським універсумом множини S .

Нехай S – множина диз'юнктив. Позначимо через H_0 множину констант, які містяться в диз'юнктах множини S . Якщо диз'юнкти з S не містять констант, то H_0 складається з однієї константи, скажімо a , тобто $H_0 = \{a\}$.

Для $i = 0, 1, 2, \dots$ множина H_{i+1} є об'єднанням множини H_i і множини всіх термів виду $f(t_1, \dots, t_n)$, де $t_1, t_2, \dots, t_n \in H_i$, f – символ n -місної функції, яка зустрічається хоча б в одному з диз'юнктив множини S .

Кожну множину H_i називають *множиною констант i -го рівня* для множини диз'юнктив S , а множину H_∞ називають *ербранівським універсумом* множини S .

Приклад 3.1. Нехай $S = \{P(x), \neg P(x) \vee Q(f(y)), \neg Q(f(a))\}$. Тоді

$$H_0 = \{a\};$$

$$H_1 = \{a, f(a)\};$$

$$H_2 = \{a, f(a), f(f(a))\}$$

$$\dots\dots\dots$$

$$H_\infty = \{a, f(a), f(f(a)), f(f(f(a))), \dots\}.$$

Приклад 3.2. Нехай $S = \{P(x), Q(x) \vee \neg R(y)\}$. Диз'юнкти з множини S не містять констант, тому $H_0 = \{a\}$. Оскільки диз'юнкти в S не містять функціональних символів, то

$$H_0 = H_1 = H_2 = \dots = H_\infty = \{a\}.$$

Приклад 3.3. Нехай $S = \{P(f(x), a, g(y), b) \vee Q(x)\}$. Тоді

$$H_0 = \{a, b\};$$

$$H_1 = \{a, b, f(a), f(b), g(a), g(b)\};$$

$$\dots\dots\dots$$

$$H_\infty = \{a, b, f(a), f(b), g(a), g(b), f(f(a)), f(f(b)), f(g(a)), f(g(b)), g(f(a)), g(f(b)), g(g(a)), g(g(b)), f(f(f(a))), \dots\}.$$

Ербранівський універсум, як ми побачили, визначений не всією множиною диз'юнктив S , а тільки символами функцій і константами, які входять у диз'юнкти з множини S .

Множину атомарних формул виду $P(t_1, \dots, t_n)$, де $t_1, \dots, t_n \in H_\infty$, а P – символ n -місного предиката, який входить у диз'юнкти з S , називають *ербранівським базисом множини диз'юнктив S* .

Для множини диз'юнктив S з прикладу 3.1 ербранівським базисом є множина атомарних формул

$$B = \{P(a), Q(a), P(f(a)), Q(f(a)), P(f(f(a))), Q(f(f(a))), \dots\}.$$

Ербранівським базисом множини диз'юнктив S з прикладу 3.2 є множина

$$B = \{P(a), Q(a), R(a)\}.$$

Ербранівським базисом множини диз'юнктив S з прикладу 3.3 є множина

$$B = \{P(t_1, t_2, t_3, t_4), Q(t) \mid t_1, t_2, t_3, t_4, t \text{ пробігають елементи ербранівського універсуму множини } S\}.$$

Зауваження. Звернемо увагу, що у формулах ербранівського базису B множини диз'юнктив S відсутні предметні змінні.

Нехай термін «*вираз*» означає терм, атом, літерал або диз'юнкт. Тоді *основним виразом* називатимемо вираз, який не містить предметних змінних.

Очевидно, що ербранівський базис множини S складається з *основних атомів (тобто атомів, які не містять предметних змінних)*.

3.4. Основний приклад диз'юнкту

Нехай D – диз'юнкт із множини диз'юнктив S . Основним прикладом диз'юнкту D називають диз'юнкт, отриманий із D заміною предметних змінних на елементи ербранівського універсуму H_∞ .

Нехай S – множина диз'юнктив із прикладу 3.1. Тоді диз'юнкт $\neg Q(f(a))$ має один основний приклад – це сам цей диз'юнкт; множина основних прикладів диз'юнкту $\neg P(x) \vee Q(f(y))$ нескінченна:

$$\{\neg P(a) \vee Q(f(a)), \neg P(f(a)) \vee Q(f(a)), \neg P(a) \vee Q(f(f(a))), \dots\}.$$

Якщо S – множина диз'юнктив із прикладу 3.2, то кожний диз'юнкт цієї множини має один основний приклад.

Приклад 3.4. Нехай множина диз'юнктив $S = \{P(x), Q(f(y)) \vee R(y)\}$.

Ербранівський універсум $H_\infty = \{a, f(a), f(f(a)), f(f(f(a))), \dots\}$.

Нехай диз'юнкт $D(x) = P(x)$. Тоді $P(a), P(f(a)), P(f(f(a))), P(f(f(f(a))))$, ... є основними прикладами диз'юнкту $D(x)$ (їх нескінченне число).

Нехай диз'юнкт $D(y) = Q(f(y)) \vee R(y)$. Тоді $Q(f(a)) \vee R(a), Q(f(f(a))) \vee R(f(a)), Q(f(f(f(a)))) \vee R(f(f(a))), \dots$ є основними прикладами диз'юнкту $D(y)$ (їх також нескінченне число).

Зауваження. Основний приклад диз'юнкту $D(x_1, \dots, x_k)$ не містить предметних змінних. Він отримується заміною предметних змінних x_1, \dots, x_k деякими елементами ербранівського універсуму.

Як було зазначено на початку підрозділу 3.3 (і буде доведено в подальшому), для розв'язування питання про невиконуваність множини диз'юнктив S , як предметну область достатньо розглядати тільки ербранівський універсум. *Виявляється, можна ще обмежити й саму інтерпретацію до так званої H -інтерпретації.*

3.5. H -інтерпретація

Почнемо з означення.

Нехай H_∞ – ербранівський універсум множини диз'юнктив S . Інтерпретацію ***IH*** з предметною областю H_∞ називають *H -інтерпретацією множини S* , якщо вона задовольняє таким умовам:

- 1) інтерпретація ***IH*** відображає всі константи, які входять у диз'юнкти множини S , у самих себе;
 - 2) якщо f – символ n -місної функції, яка входить у диз'юнкти множини S , та h_1, \dots, h_n – елементи H_∞ , то в ***IH*** через f позначається функція, яка відображає (h_1, \dots, h_n) – елемент $(H_\infty)^n$ у $f(h_1, \dots, h_n)$ – елемент H_∞ ;
 - 3) значення предикатів в S на елементах із H_∞ не обмежено нічим.
-

Якщо $B = \{B_1, B_2, \dots, B_n, \dots\}$ – ербранівський базис множини диз'юнктив S , то H -інтерпретацію ***IH*** зручно подати як множину літералів

$$\{L_1, L_2, \dots, L_n, \dots\},$$

де $L_i \in B_i$, якщо ***IH***(B_i) = Т, і $L_i \in \neg B_i$, якщо ***IH***(B_i) = F (тут Т і F – значення істинності).

Зауваження. **Задати ербранівську інтерпретацію** – означає задати істинні значення для атомів ербранівського базису.

Приклад 3.5. Нехай S – множина диз'юнктивів із прикладу 3.1. Нижче наведено деякі H -інтерпретації:

$$IH_1 = \{P(a), Q(a), P(f(a)), Q(f(a)), P(f(f(a))), \dots\},$$

$$IH_2 = \{P(a), Q(a), \neg P(f(a)), \neg Q(f(a)), P(f(f(a))), Q(f(f(a))), \neg P(f(f(f(a))))\}, \dots\},$$

$$IH_3 = \{P(a), \neg Q(a), P(f(a)), \neg Q(f(a)), P(f(f(a))), \neg Q(f(f(a))), \dots\}.$$

Останнє, що залишилося зробити для доведення головного результату, – увести поняття H -інтерпретації I^*H , яка відповідає (будь-якій) інтерпретації I множини диз'юнктивів S .

Нехай I – інтерпретація на області M . H -інтерпретацією I^*H , яка відповідає I , є інтерпретація, яка задовольняє таку умову. Нехай h_1, \dots, h_n – елементи ербранівського універсуму H_∞ . Нехай кожний h_i відображається в якійсь d_i із M . Якщо $P(d_1, \dots, d_n)$ отримує в інтерпретації I логічне значення Т (або F), то $P(h_1, \dots, h_n)$ також отримує логічне значення Т (відповідно F) в інтерпретації I^*H для будь-якого символу предиката P . Іншими словами, для довільних елементів h_1, \dots, h_n ербранівського універсуму й будь-якого символу предиката P виконується еквівалентність

$$(I^*HP)(h_1, \dots, h_n) = (IP)(I(h_1), \dots, I(h_n)).$$

Приклад 3.6. Нехай S – множина диз'юнктивів із прикладу 3.1. Нагадаємо, що

$$S = \{P(x), \neg P(x) \vee Q(f(y)), \neg Q(f(a))\}.$$

Розглянемо інтерпретацію I на області $M = \{1, 2\}$, яку визначено рівністю $I(a) = 1$ і таблицею

$f(1)$	$f(2)$	$P(1)$	$P(2)$	$Q(1)$	$Q(2)$
2	1	F	T	T	T

Для такої інтерпретації I можна визначити H -інтерпретацію I^*H , яка відповідає інтерпретації I . Це виконують так. Спочатку будують ербранівський базис множини диз'юнктивів S :

$$B = \{P(a), Q(a), P(f(a)), Q(f(a)), P(f(f(a))), Q(f(f(a))), \dots\}.$$

Тепер, користуючись інтерпретацією I , задають значення кожного члена ербранівського базису B :

$$\begin{aligned} P(a) &= P(1) = F; \\ Q(a) &= Q(1) = T; \\ P(f(a)) &= P(f(1)) = P(2) = T; \\ Q(f(a)) &= Q(f(1)) = Q(2) = T; \\ P(f(f(a))) &= P(f(f(1))) = P(f(2)) = P(1) = F; \\ Q(f(f(a))) &= Q(f(f(1))) = Q(f(2)) = Q(1) = T; \\ &\dots \end{aligned}$$

Останнє можна записати так

$P(a)$	$Q(a)$	$P(f(a))$	$Q(f(a))$	$P(f(f(a)))$	$Q(f(f(a)))$...
F	T	T	T	F	T	...

Інакше це можна записати так:

$$I^*H = \{\neg P(a), Q(a), P(f(a)), Q(f(a)), \neg P(f(f(a))), Q(f(f(a))), \dots\}.$$

Побудову інтерпретації множини S на ербранівському універсумі закінчують, коли з кожною атомарною формулою ербранівського базису зв'язано відповідне значення

істинності. Нагадаємо, що остаточно H -інтерпретацію I^*H подано як множину літералів $\{L_1, L_2, \dots, L_n, \dots\}$, де $L_i \in B_i$, якщо $I^*H(B_i) = T$, і $L_i \in \neg B_i$, якщо $I^*H(B_i) = F$ (T і F – значення істинності).

Розглянемо тепер випадок, коли диз'юнкти з множини S не містять констант. Нехай I – інтерпретація множини S на області M і a – константа, яку використано, щоб почати ербранівський універсум H_∞ . У цьому випадку значення $I(a)$ не визначено. Для отримання H -інтерпретації I^*H розширимо функцію I на a , для чого візьмемо $I(a)$ рівним будь-якому елементу з M . Далі поступаємо так, як щойно описано. Якщо множина M неодноеlementна, то можна отримати не одну H -інтерпретацію I^*H , яка відповідає I . Неважко навести приклад, коли H -інтерпретацій I^*H стільки ж, скільки елементів у множині M .

Приклад 3.7. Нехай $S = \{P(x), Q(y, f(y, z))\}$ і для S вибрана така інтерпретація I :

$$M = \{1, 2\};$$

$f(1, 1)$	$f(1, 2)$	$f(2, 1)$	$f(2, 2)$
1	2	2	1

$P(1)$	$P(2)$	$Q(1, 1)$	$Q(1, 2)$	$Q(2, 1)$	$Q(2, 2)$
T	F	F	T	F	T

Тоді інтерпретації I відповідатимуть дві H -інтерпретації:

$$I^*H_1 = \{\neg P(a), Q(a, a), P(f(a, a)), \neg Q(a, f(a, a)), Q(f(a, a), a), \neg Q(f(a, a), f(a, a)), \dots\}, \text{ якщо } a = 1;$$

$$I^*H_2 = \{P(a), \neg Q(a, a), P(f(a, a)), \neg Q(a, f(a, a)), \neg Q(f(a, a), a), \neg Q(f(a, a), f(a, a)), \dots\}, \text{ якщо } a = 2.$$

Лема 3.1. Якщо інтерпретація I на деякій області M задовольняє множину диз'юнктив S , то довільна H -інтерпретація I^*H , яка відповідає інтерпретації I , також задовольняє множину S .

Доведення. Нехай інтерпретація I задовольняє множину диз'юнктив S на області M . Це означає, що довільна формула з S приймає значення T на області M при інтерпретації I . Але тоді за побудовою інтерпретації I^*H , яка відповідає I , кожна з формул множини S теж приймає значення T при інтерпретації I^*H , тобто інтерпретація I^*H задовольняє множину S .

Теорема 3.1 (про H -інтерпретацію). Множина диз'юнктив S невиконувана (не виконується при всіх інтерпретаціях) тоді й тільки тоді, коли S не виконується при всіх H -інтерпретаціях множини S .

Доведення. Необхідність умов теореми очевидна. Справді, за означенням множина S невиконувана тоді й тільки тоді, коли ця множина фальшива при будь-якій інтерпретації, у тому числі й при будь-якій H -інтерпретації.

Достатність випливає з леми 3.1, бо якщо S виконувана, то існує хоча б одна інтерпретація I , при якій усі диз'юнкти з S істинні. Але тоді всі диз'юнкти з S будуть істинні й при H -інтерпретації I^*H , яка відповідає інтерпретації I . А це суперечить тому, що S невиконувана при всіх H -інтерпретаціях множини S .

Теорему про H -інтерпретацію доведено. Ця теорема дуже важлива, бо дає змогу розглядати тільки інтерпретації над ербранівським універсумом (і навіть обмежитись H -інтерпретаціями) для перевірки невиконаності множини диз'юнктив.

3.6. Важливі твердження щодо H -інтерпретації

Звернемо увагу на такі очевидні твердження.

1. Основний приклад D' диз'юнкту D **виконується** в H -інтерпретації \mathbf{IH} тоді і тільки тоді, коли **існує** основний літерал L' в D' , такий, що L' є також в \mathbf{IH} , $D' \cap \mathbf{IH} \neq \emptyset$. (В останній формулі вважаємо, що D' подано як множину літералів.)

2. Диз'юнкт D **виконується** в H -інтерпретації \mathbf{IH} тоді і тільки тоді, коли **кожний** основний приклад диз'юнкту D **виконується** в \mathbf{IH} .

3. Диз'юнкт D **спростовується** H -інтерпретацією \mathbf{IH} тоді і тільки тоді, коли **існує** такий основний приклад D' для D , що D' **не виконується** в \mathbf{IH} .

4. Множина диз'юнктив S **невиконувана** (в жодній інтерпретації) тоді і тільки тоді, коли для **кожної** H -інтерпретації \mathbf{IH} існує основний приклад D' якогось диз'юнкту D із S , що D' **не виконується** в \mathbf{IH} .

Приклад 3.8. Розглянемо диз'юнкт $D = \neg P(x) \vee Q(f(x))$ і такі H -інтерпретації:

$$\mathbf{IH}_1 = \{\neg P(a), \neg Q(a), \neg P(f(a)), \neg Q(f(a)), \neg P(f(f(a))), \neg Q(f(f(a))), \dots\};$$

$$\mathbf{IH}_2 = \{P(a), Q(a), P(f(a)), Q(f(a)), P(f(f(a))), Q(f(f(a))), \dots\};$$

$$\mathbf{IH}_3 = \{P(a), \neg Q(a), P(f(a)), \neg Q(f(a)), P(f(f(a))), \neg Q(f(f(a))), \dots\}.$$

Тут кожен H -інтерпретацію подано як множину літералів $\{L_1, L_2, \dots, L_n, \dots\}$, де $L_i \in B_i$, якщо $\mathbf{I}(B_i) = \mathbf{T}$, і $L_i \in \neg B_i$, якщо $\mathbf{I}(B_i) = \mathbf{F}$.

Диз'юнкт D виконується в H -інтерпретаціях \mathbf{IH}_1 і \mathbf{IH}_2 , бо кожний його основний приклад виконується в цих інтерпретаціях – це означає, що кожний основний приклад цього диз'юнкту містить літерал, який також є в цих інтерпретаціях, представлених у вигляді множин літералів. Зокрема, розглянемо основний приклад $D' = \neg P(f(a)) \vee Q(f(f(a)))$ цього диз'юнкту. Літерал $\neg P(f(a))$ є в інтерпретації \mathbf{IH}_1 , отже, D' виконується в цій інтерпретації. Літерал $Q(f(f(a)))$ є в інтерпретації \mathbf{IH}_2 , отже, D' виконується також і в інтерпретації \mathbf{IH}_2 . Легко побачити, що аналогічні твердження справджуються і для будь-яких основних прикладів цього диз'юнкту.

Але диз'юнкт $D = \neg P(x) \vee Q(f(x))$ спростовується H -інтерпретацією \mathbf{IH}_3 . Справді, основний приклад $D' = \neg P(a) \vee Q(f(a))$ не виконується в \mathbf{IH}_3 , бо ні літералу $\neg P(a)$, ні літералу $Q(f(a))$ у цій інтерпретації немає. Іншими словами, $D' \cap \mathbf{IH}_3 = \emptyset$; тут диз'юнкт D' подано як множину літералів $\{\neg P(a), Q(f(a))\}$.

Приклад 3.9. Розглянемо множину диз'юнктив $S = \{P(x), \neg P(a)\}$. Існує тільки дві H -інтерпретації: $\mathbf{IH}_1 = \{P(a)\}$ і $\mathbf{IH}_2 = \{\neg P(a)\}$. Для H -інтерпретації \mathbf{IH}_1 існує основний приклад другого диз'юнкту: $\neg P(a)$, який не виконується в цій інтерпретації, а в H -інтерпретації \mathbf{IH}_2 не виконується основний приклад $P(a)$ першого диз'юнкту. Отже, множина S невиконувана, бо спростовується цими двома H -інтерпретаціями (а їх усього дві).