**ЗМІСТ**

**ЛЕКЦІЯ 12. ВИБІР ПІДХОДЯЧИХ (ОПТИМАЛЬНИХ) СИСТЕМ ДРУГОГО ЕПІСТЕМОЛОГІЧНОГО РІВНЯ ..... 1**
12.1 Поняття проекції функції поведінки ..... 1
12.2 Впорядкування підмасок за складністю та нечіткістю ..... 2

## Лекція 12. Вибір підходячих (оптимальних) систем другого епістемологічного рівня

**Мета лекції:** опанувати поняття проекції функції поведінки та набути вмінь з його практичного використання; набути знань та навичок визначення оптимальних систем з поведінкою за критеріями складності та нечіткості.

### План лекції

12.1 Поняття проекції функції поведінки.
12.2 Впорядкування підмасок за складністю та нечіткістю.

***Перелік ключових термінів і понять з теми:*** *проекція функції поведінки; побудова функцій поведінки за допомогою проекцій та через вибірки даних; розмір (потужність) маски; впорядкування за складністю, за підмасками та за нечіткістю; квазівпорядкування; об'єднане впорядкування; підходяща система з поведінкою; алгоритм побудови множини оптимальних породжуючих систем з поведінкою нейтрального типу*.

### 12.1 Поняття проекції функції поведінки

Дана система даних **D** з повністю впорядкованою параметричною множиною та з найбільшою допустимою маскою **M**, сумісної з **D**. Потрібно визначити всі системи з поведінкою, що задовольняють ***вимогам узгодженості, детермінованості і простоти***, причому вимога узгодженості є більш пріоритетною, ніж інші дві.

Будь-яка найбільш допустима маска **M** містить набір коректних масок, кожна з яких є підмножиною **M**. Для кожної маски може бути визначена функція поведінки, що добре узгоджується з даними, за допомогою розрідженої вибірки даних. Однак на практиці досить провести вибірку тільки для маски **M**. Функції поведінки для її підмасок можуть бути отримані обчисленням ***проекцій функції поведінки відповідної маски*** **M** [4].

Для заданої функції $f_B$, визначеної через повні стани деяких обраних змінних, кожна з її проекцій також є функцією поведінки, відповідної $f_B$, що заснована на певній підмножині обраних змінних. Нехай $s_k\left(k \in N_{|M|}\right)$ – вибіркові змінні, через які визначаються стани $f_B$; **M** – маска, через яку обираються значення обраних змінних. Нехай $\left[f_B \downarrow Z\right]$ – проекція $f_B$, де підмножина множини $\mathbf{N}_{[\mathbf{M}]}$ ідентифікаторів обраних змінних, тобто $Z \subset N$. Тоді

$$\left[f_B \downarrow Z\right]: \underset{k \in Z}{\times} S_k \to [0,1], \tag{12.1}$$

так що

$$\left[f_B \downarrow Z\right](\mathbf{x}) = a(\{f(\mathbf{c}) \mid \mathbf{c} \succ \mathbf{x}\}), \tag{12.2}$$

де $a$ – якась агрегуюча функція, обумовлена характером функції $f_B$ . Наприклад,

$$\left[f_B \downarrow Z\right](\mathbf{x}) = \sum_{\mathbf{c} \succ \mathbf{x}} f_B(\mathbf{c}), \qquad (12.3)$$

де $f_B$ – розподіл ймовірностей.

Будемо в контексті будь-якої конкретної задачі через ${}^{*}f_B$ позначати ***функцію поведінки для найбільш прийнятної маски*** **M**. Через ${}^{i}f_B\,(i = 2,3,...)$ будемо позначати ***функції поведінки для її різних осмислених підмасок*** ${}^{i}\mathbf{M}$, кожна з яких пов'язана з множиною ${}^{i}Z \subset N_{|\mathbf{M}|}$ ідентифікаторів вибіркових змінних.

За винятком дуже великих наборів даних, з точки зору обчислень простіше визначати функції поведінки за допомогою проекцій, а не через вибірки даних. Отже, краще робити вибірку тільки одного разу для найбільш прийнятної маски, а потім визначати функції поведінки для всіх змістовних підмасок як відповідні проекції.

## 12.2 Впорядкування підмасок за складністю та нечіткістю

Для заданої системи даних **D** і найбільш допустимої маски **M** *вимога відповідності* призводить до обмеженої множини

$$Y_r = \left\{ {}^{i}\mathbf{F}_B = \left(\mathbf{I}, {}^{i}\mathbf{M}, {}^{i}f\right) \mid i = 1,2,...,\left(N(n, \Delta\mathbf{M})\right)\right\}, \qquad (12.4)$$

що містить по одній системі з поведінкою для кожної осмисленої підмаски ${}^{i}M \subseteq \mathbf{M}$; нехай для зручності ${}^{i}M = \mathbf{M}$.

Наступним кроком розв'язання розглянутої задачі має бути обчислення ступенів недетермінованості та складності для всіх систем з множини $Y_r$ [5].

Як відомо *ступінь детермінованості* задається відповідною ***мірою породжуючої нечіткості***, яка визначається для ймовірнісних систем *шеннонівською ентропією*.

Що стосується ***міри складності***, то тут можливо багато варіантів. Візьмемо для прикладу просту, але змістовну міру, яку часто використовують – ***розмір (потужність) маски***.

Нехай $q_u\ (i = 1,2,...)$ – значення відповідних породжуючих нечіткостей для систем з поведінкою ${}^{i}\mathbf{F}_B$ з обмеженої множини $Y_r$ .

Оскільки будь-яка система $\mathbf{F}_B$ однозначно ідентифікується своєю маскою **M**, ***потужність*** якої $\left|{}^{i}M\right|$ задає її складність, статус системи ${}^{i}\mathbf{F}_B$ в сенсі породжуючої нечіткості і складності зручно описувати парою $\left(\left|{}^{i}M\right|, {}^{i}q_u\right)$.

Чисельне впорядкування підмасок ${}^{i}M$ , ідентифікуючих системи з $Y_r$ за їх потужністю, задає ***впорядкування складності*** $\overset{c}{\leq}$ на множині $Y_r$ .

Чисельне упорядкування значень ${}^{i}q_u$ визначає ***впорядкування за нечіткістю*** $\overset{u}{\leq}$ на множні $Y_r$.

У той час, як впорядкування за складністю повністю визначається самими масками (підмасками), впорядкування за нечіткістю може бути визначене тільки після оцінки масок (підмасок).

Для будь-якої множини породжуючих масок (підмасок) можемо визначити часткове впорядкування

$${}^{i}M_G \leq {}^{j}M_G \text{ тоді і тільки тоді, коли } {}^{i}\mathbf{g}={}^{j}\mathbf{g} \text{ та } {}^{i}\overline{\mathbf{g}} \prec {}^{j}\overline{\mathbf{g}} \qquad (12.5)$$

(або ${}^{i}\mathbf{e} \prec {}^{j}\mathbf{e}$ для спрямованих систем), яке ми будемо називати ***впорядкуванням підмасок***.

З погляду упорядкованості за складністю будь-яка маска деякого рівня є безпосереднім наступником будь-якої маски найближчого більш високого рівня і безпосереднім попередником будь-якої маски найближчого нижчого рівня. Отже, ***впорядкування за складністю*** – це зв'язне квазівпорядкування (рефлексивне і транзитивне, визначене для будь-якої пари систем).

***Впорядкування за підмасками*** є частковим впорядкуванням, але решітки воно не утворює. Однак, воно являє собою набір решіток по одній для кожної множини породжуваних вибіркових змінних.

***Впорядкування за нечіткістю*** є зв'язним, але через те, що кілька різних систем можуть мати однакову породжуючу нечіткість, це відношення не є антисиметричним. Отже, в загальному випадку таке впорядкування є зв'язним квазівпорядкуванням, яке в деяких окремих випадках виявляється повним впорядкуванням.

Отже, на множині $Y_r$ визначені ***два зв'язних квазівпорядкування – за складністю*** **та** ***за нечіткістю.***

Оскільки для розглянутого типу задач потрібно, щоб і складність, і породжуюча нечіткість систем у множині рішень $Y_Q$ були мінімізовані, відповідне ***об'єднане впорядкування*** $\overset{*}{\leq}$ визначається наступним чином:

$${}^{i}\mathbf{F}_B \overset{*}{\leq} {}^{j}\mathbf{F}_B \text{ тоді і тільки тоді, коли } \left|{}^{i}M\right| \overset{c}{\leq} \left|{}^{j}M\right| \text{ та } {}^{i}q_u={}^{j}q_u \ , \qquad (12.6)$$

де ${}^{i}\mathbf{F}_B, {}^{j}\mathbf{F}_B \in Y_r$.

Це впорядкування не є зв'язним, оскільки пари ${}^{i}\mathbf{F}_B, {}^{j}\mathbf{F}_B$, для яких $\left|{}^{i}M\right| < \left|{}^{j}M\right|$ та ${}^{i}q_u > {}^{j}q_u$ або $\left|{}^{i}M\right| > \left|{}^{j}M\right|$ та ${}^{i}q_u < {}^{j}q_u$ (подібні пари, зрозуміло, можуть існувати), є непорівнювальними. Воно також є неантисиметричним, оскільки не виключеною є можливість того, що

$$\left|{}^{i}M\right| = \left|{}^{j}M\right| \text{ та } {}^{i}q_u={}^{j}q_u \qquad (12.7)$$

для деяких $i \neq j$.

Отже, об'єднане впорядкування – це загального вигляду квазівпорядкування (рефлексивне і транзитивне відношення) на $Y_r$.

Тепер множину рішень $Y_Q$ можна визначити як множину всіх систем з $Y_r$, які або еквівалентні, або непорівнювальні щодо об'єднаного впорядкування (12.6).

Дві системи з $Y_r$, скажімо системи ${}^iF_B$ та ${}^jF_B$, непорівнювальні в сенсі об'єднаного впорядкування, якщо виконано одну з наступних умов:

а) ${}^iF_B$ є більш складною і більш детермінованою, ніж ${}^jF_B$;

б) ${}^iF_B$ менш складною і менш детермінованою, ніж ${}^jF_B$.

Формально

$$Y_Q = \left\{ {}^i\mathbf{F}_B \in Y_r \mid \left(\forall {}^j\mathbf{F}_B \in Y_r\right)\left( {}^j\mathbf{F}_B \overset{*}{\leq} {}^i\mathbf{F}_B \Rightarrow {}^i\mathbf{F}_B \overset{*}{\leq} {}^j\mathbf{F}_B \right)\right\}. \qquad (12.8)$$

Системи з множини рішень $Y_Q$ будемо називати ***підходящими системами з поведінкою*** для розглядуваного типу задач.

Описаний в даному підрозділі пошук відповідних систем з поведінкою може бути реалізований самими різними способами.

*Основний принцип полягає в тому, що змістовні маски отримуються за допомогою деякого алгоритму з найбільшої прийнятної маски в порядку зменшуваною складності*. Серед масок однакової складності обираються тільки маски з мінімальною породжуючою нечіткістю. При цьому, якщо значення цієї мінімальної нечіткості менше або дорівнює значенню нечіткості для попереднього рівня складності, то всі раніше прийняті системи відкидаються. В результаті застосування цієї процедури у нас залишаються тільки підходящі (оптимальні) системи.

## Питання для самоконтролю

1. Як визначається кількість коректних підмасок при обраній найбільш припустимій масці?
2. Як визначається проекція функції поведінки $f_B$?
3. Вкажіть, що є простішим з точки зору обчислень, побудова функцій поведінки за допомогою проекцій або через вибірки даних.
4. Як визначається розмір (потужність) маски?
5. Охарактеризуйте впорядкування за складністю, за підмасками та за нечіткістю.
6. Які з впорядкувань (за складністю, за підмасками, за нечіткістю) є квазівпорядкуваннями?
7. Що називається об'єднаним впорядкуванням?
8. Як визначається об'єднане впорядкування?
9. Що називається підходящою системою з поведінкою?
10. Наведіть алгоритм побудови множини оптимальних породжуючих систем з поведінкою нейтрального типу.