

Лабораторна робота 7

Тема. Побудова множини оптимальних породжуючих систем з поведінкою нейтрального та спрямованого типів

Мета роботи – вивчити основні поняття та визначення, що стосуються побудови множини оптимальних систем з поведінкою, розкрити основні методологічні підходи до побудови множини оптимальних систем з поведінкою нейтрального та спрямованих типів, визначити основні етапи побудови оптимальних систем з поведінкою зазначених типів, навчитися тлумачити отримані результати.

Стислі теоретичні відомості

Нехай задана система даних D з повністю впорядкованою параметричною множиною W і найбільшою допустимою маскою $M = V \times R$, сумісною з системою даних D .

Потрібно визначити множину систем з поведінкою Y_D , що задовольняє вимогам **узгодженості, детермінованості і простоти**.

Для цього розіб'ємо процес визначення зазначеної множини на етапи.

1. Першим етапом є визначення множини систем з поведінкою Y_r .

З огляду на те, що будь-яка найбільша допустима маска містить множину коректних підмасок iM ($i = 1, \dots, N(n, \Delta M)$), кожна з яких є підмножиною M , число яких визначається наступним чином

$$N(n, \Delta M) = (2^{\Delta M} - 1)^n - (2^{\Delta M - 1} - 1)^n, \quad (1)$$

де n – число базових змінних, ΔM – глибина найбільшої допустимої маски ($\Delta M \leq 10$), то для кожної підмаски iM може бути визначена функція поведінки ${}^i f_B$, за допомогою обчислення проєкції функції поведінки відповідної підмаски iM .

Для заданої системи даних D і найбільшої допустимої маски M , вимога **узгодженості** призводить до визначення обмеженої множини

$$Y_r = \left\{ {}^i F_B = \left(I, {}^i M, {}^i f_B \right), \left(i = 1, \dots, N(n, \Delta M) \right) \right\} \quad (2)$$

яка містить по одній системі з поведінкою для кожної осмисленої підмаски ${}^iM \subseteq M$ (для зручності позначимо ${}^1M = M$).

Для заданої функції поведінки f_B , визначеної через повні стани вибіркового змінних S_k ($k \in N_{|M|}$), кожна з її проєкцій також є функцією

поведінки, що відповідає функції поведінки f_B в сенсі субстанів, заснованих на визначенні підмножин вибірових змінних S_k .

Нехай $S_k (k \in N_{|M|})$ – вибірові змінні, через які визначаються стани функції поведінки f_B ; M – маска, через яку обираються значення вибірових змінних.

Нехай $[f_B \downarrow Z]$ – проекція функції поведінки f_B , де Z – підмножина множини $N_{|M|}$ ідентифікаторів вибірових змінних $S_k (k \in N_{|M|})$, тобто $Z \subset N$.

Тоді

$$[f_B \downarrow Z]: \times_{k \in Z} S \rightarrow [0,1].$$

Якщо функція f_B – розподіл ймовірностей, тоді проекція визначається наступним чином

$$[f_B \downarrow Z](x) = \sum_{c > x} f_B(c). \quad (3)$$

Слід зазначити, що визначення функцій поведінки ${}^i f_B$ за допомогою проєкцій простіше з точки зору обчислень, ніж через вибірки даних. Чим більшим є обсяг даних, тим більше часу економиться, якщо замість вибірок даних використовувати проєкції на обчислювальні операції.

Отже, згідно зазначеному підходу, можна робити вибірку станів вибірових змінних тільки один раз для найбільшої допустимої маски M (${}^1 M = M$), а потім визначити функції поведінки ${}^i f_B (i=2, \dots, N(n, \Delta M))$, для всіх змістовних подмасок ${}^i M (i=2, \dots, N(n, \Delta M))$, як відповідні проєкції визначеної функції поведінки.

2. Наступним (другим) етапом при визначенні множини оптимальних систем з поведінкою має бути визначення ступенів **недетермінованості і складності** для всіх систем з множини Y_r та виділено множини систем з поведінкою Y_Q , які або еквівалентні, або непорівнянні щодо об'єднаного впорядкування по складності і нечіткості.

Системи з множини Y_Q , що визначаються як

$$Y_Q = \left\{ {}^i F_B \in Y_r \mid (\forall {}^j F_B \in Y_r) \left({}^j F_B \stackrel{*}{\leq} {}^i F_B \Rightarrow {}^i F_B \stackrel{*}{\leq} {}^j F_B \right) \right\}, \quad (4)$$

прийнято називати підходящими системами з поведінкою.

Об'єднане впорядкування \leq^* визначається наступним чином:

$${}^i F_B \stackrel{*}{\leq} {}^j F_B \quad \text{тоді і тільки тоді, коли} \quad |{}^i M| \stackrel{c}{\leq} |{}^j M| \quad \text{и} \quad {}^i q_u \stackrel{u}{\leq} {}^j q_u, \quad (5)$$

де ${}^i F_B, {}^j F_B \in Y_r, \leq^c$ – чисельне впорядкування подмасок ${}^i M$ і ${}^j M$ за складністю (складність визначається потужністю порівнюваних підмасок) на множині Y_r (рис. 2), \leq^u – чисельне впорядкування по нечіткості на множині Y_r .

Ступінь недетермінованості визначається відповідною мірою породжуючої нечіткості.

Нехай ${}^i q_u (i=1,2,\dots)$ це значення відповідних породжуваних нечіткостей для систем з поведінкою, що визначаються при вибраному порядку породження даних і відповідним розбиттям маски ${}^i M$ на ${}^i M_g$ і ${}^i M_{\bar{g}}$, тобто введенням. ${}^i M_{G=}$ (${}^i M; {}^i M_g, {}^i M_{\bar{g}}$)

Породжувана нечіткість нейтральної породжувальної системи з поведінкою визначається наступним чином

$${}^i q_u = {}^i H({}^i G | {}^i \bar{G}) = {}^i H({}^i C) - {}^i H({}^i \bar{G}), (i=1, \dots, N(n, \Delta M)). \quad (6)$$

Об'єднане впорядкування $\left(\leq^*\right)$ може бути представлено графічно.

Розглянемо **приклад**.

Для заданої системи даних D , наведеної на рис. 1, визначимо, всі системи з поведінкою з множини Y_Q , що використовуються для прогнозу.

$t =$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
V_1	1	1	0	3	3	3	3	3	4	3	3	0	2	1	1
V_2	4	3	4	3	3	3	3	3	4	4	3	3	4	2	1
$t =$	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
V_1	1	1	4	4	4	4	4	0	2	2	4	4	4	4	4
V_2	1	1	4	4	4	4	3	1	4	2	2	4	4	4	4
$t =$	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
V_1	4	4	2	2	2	2	2	0	2	1	0	2	1	1	4
V_2	4	4	4	4	1	1	1	3	3	1	3	3	3	3	4

Рисунок.1. Система даних D

Визначимо, всі системи з поведінкою з множини Y_Q , що використовуються для прогнозу.

Нехай $\Delta M = 2$. Тоді згідно (1) є 8 змістовних подмасок ${}^i M (i=1,2,\dots,8)$, які можуть бути впорядковані за складністю $\left(\leq^c\right)$ таким чином:

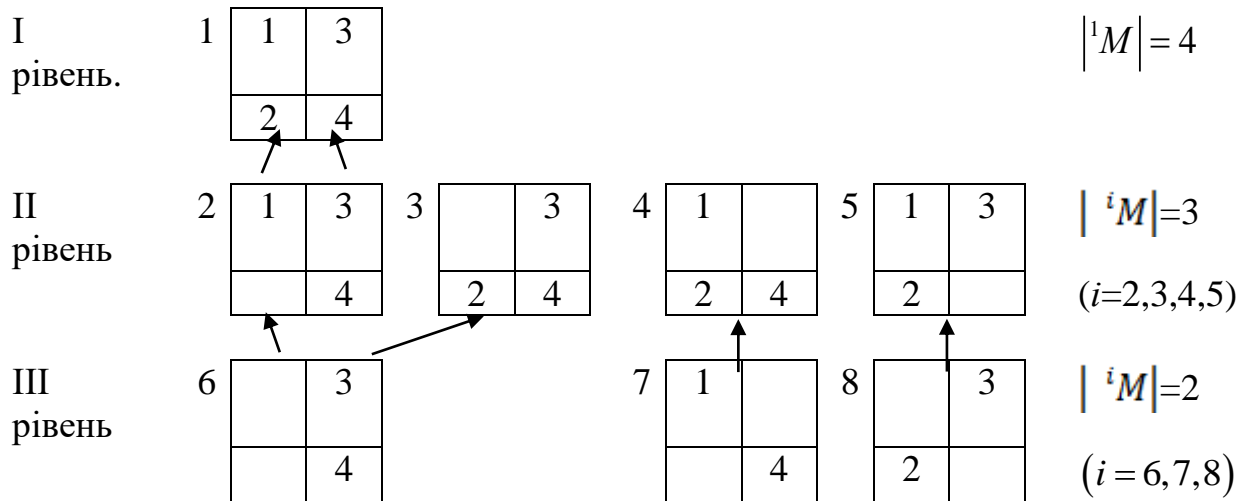


Рисунок. 2. Упорядкування підмасок iM ($i=1,2,\dots,8$) за складністю

Отже, маємо три рівня складності, які відображені на рисунку 2.

Після визначення повної вибірки для найбільшої допустимої маски ${}^1M=M$ по частотам $N(c)$ визначаємо ймовірність появи повних станів вибірових змінних ${}^1f_B(c)$ при $N(\alpha)=44$ та представляємо в стандартній формі – таблиці виду

	S_1	S_2	S_3	S_4	$N(c)$	${}^1f_B(c)$
$c=$	1	4	1	3	1	0,0227
	1	3	0	4	1	0,0227
	0	4	3	3	1	0,0227
	3	3	3	3	4	0,0909
	3	3	4	4	1	0,0227
	4	4	3	4	1	0,0227
	3	4	3	3	1	0,0227
	3	3	0	3	1	0,0227
	0	3	2	4	1	0,0227
	2	4	1	2	1	0,0227
	1	2	1	1	1	0,0227
	1	1	1	1	2	0,04545
	1	1	4	4	1	0,0227
	4	4	4	4	8	0,1818
	4	4	4	3	1	0,0227
	4	3	0	1	1	0,0227
	0	1	2	4	1	0,0227

	2	4	2	2	1	0,0227
	2	2	4	2	1	0,0227
	4	2	4	4	1	0,0227
	4	4	2	4	1	0,0227
	2	4	2	4	1	0,0227
	2	4	2	1	1	0,04545
	2	1	2	1	2	0,0227
	2	1	0	3	1	0,0227
	0	3	2	3	2	0,04545
	2	3	1	1	1	0,0227
	1	1	0	3	1	0,0227
	2	3	1	3	1	0,0227
	1	3	1	3	1	0,0227
	1	3	4	4	1	0,0227

Визначимо шенновську ентропію за наступним виразом

$${}^1H(C) = -\sum_{c \in C} f_B(c) \log_2 f_B(c).$$

Визначимо породжуючу нечіткість для ${}^iM_G = ({}^iM, {}^iM_g, {}^iM_{\bar{g}})$

$\overbrace{S_1 \quad S_2}^{{}^1\bar{g}}$		${}^1f(\bar{g})$
1	4	0,0227
1	3	$(0,0227 * 3) = 0,06815$
0	4	0,0227
3	3	$(0,0909 + 0,0227 * 2) = 0,1363$
0	3	$(0,0227 * 3) = 0,06815$
2	4	$(0,0227 * 4) = 0,0908$
1	2	0,0227
1	1	$(0,0227 * 4) = 0,0908$
4	3	0,0227
0	1	0,0227
2	2	0,0227
4	2	0,0227
4	4	$(0,0227 * 3 + 0,1818) = 0,2497$
2	1	$(0,04545 + 0,0227) = 0,06815$
2	3	$(0,0227 * 2) = 0,04545$

$${}^1H(\bar{G}) = -7 * 0,0227 \log_2 0,0227 - 3 * 0,0615 \log_2 0,0615 - 0,1363 \log_2 0,1363 - 2 * 0,0908 \log_2 0,0908 - 0,2497 \log_2 0,2497 - 0,04545 \log_2 0,04545.$$

Породжуча нечіткість для першої системи з поведінкою дорівнює

$${}^1H(G|\bar{G}) = {}^1H(C) - {}^1H(\bar{G}) = 1,11.$$

Визначимо для інших семи подмасок з множини ${}^iM (i=1,2,\dots,8)$, породжуючі нечіткості, як проєкції на ${}^1M = M$.

Нехай для прикладу визначимо проєкцію ймовірнісної функції поведінки для 2M (для $Z = \{1,3,4\}$).

	S_1	S_3	S_4	$[f \downarrow \{1,3,4\}](x)$
c=	1	1	3	$(0,0227 + 0,0227) = 0,04545$
	1	0	4	0,0227
	0	3	3	0,0227

	3	3	3	$(0,0909 + 0,0227) = 0,1136$
	3	4	4	0,0227
	4	3	4	0,0227
	3	3	3	0,0227
	3	0	3	0,0227
	0	2	4	$(0,0227 + 0,0227) = 0,04545$
	2	1	2	0,0227
	1	1	1	$(0,0227 + 0,04545) = 0,06815$
	1	4	4	$(0,0227 + 0,0227) = 0,04545$
	4	4	4	$(0,1818 + 0,0227) = 0,2045$
	4	4	3	0,0227
	4	0	1	0,0227
	2	2	2	0,0227
	2	4	2	0,0227
	4	2	4	0,0227
	2	2	4	0,0227
	2	2	1	$(0,0227 + 0,04545) = 0,06815$
	2	0	3	0,0227
	0	2	3	$(0,0227 + 0,0227 + 0,0227) = 0,06815$
	2	1	1	0,0227
	1	0	3	0,0227
	0	2	3	0,0227
	2	1	3	0,0227
	1	1	3	0,0227
	1	4	4	0,0227

$${}^2H(C) = -16 * 0,0227 \log_2 0,0227 - 3 * 0,04545 \log_2 0,04545 -$$

$$- 0,1136 \log_2 0,1136 - 3 * 0,06815 \log_2 0,06815 - 0,2045 \log_2 0,2045.$$

$${}^2\bar{g} = \{1\}.$$

$${}^2H(G|\bar{G}) = {}^2H(C) - {}^2H(\bar{G}) = 1,88.$$

Інші породжуючі нечіткості визначаються аналогічно та остаточні дані наводяться на рис. 3.

Об'єднане впорядкування по складності і нечіткості для ${}^iM (i = 1, 2, \dots, 8) \binom{*}{\leq}$ наведено на рис 3.

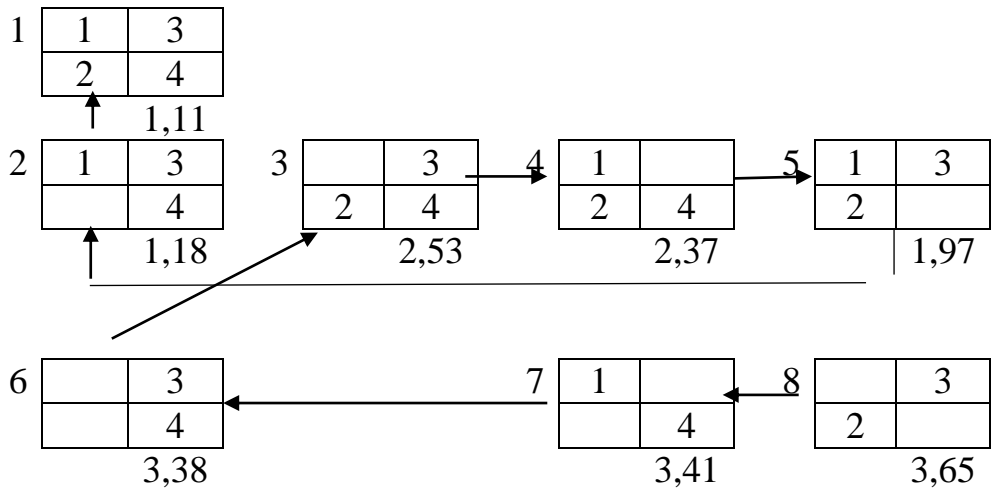


Рисунок. 3 Об'єднане впорядкування по складності і нечіткості

для ${}^iM (i = 1, 2, \dots, 8) \binom{*}{\leq}$

Індекси вгорі схем вказують номер підмаски, а внизу – значення породжуючої нечіткості.

Отже, оптимальними системами з поведінкою будуть наступні системи $Y_Q = \{ {}^1F_B, {}^2F_B, {}^6F_B \}$.

Алгоритм отримання оптимальних породжуючих спрямованих систем з поведінкою є аналогічним.