

# Notatki na kolokwium z Matematyki dyskretnej – część 2

Michał Puchyr

## 1 Relacje równoważności

### Objaśnienia

Operacja rozbicia - rozdzielanie klas równoważności np.  $\{a, b, c\} \xrightarrow{op2} \{a, b\}\{c\}$

Operacja sklejenia - sklejanie klas równoważności np.  $\{a, b\}\{c\} \xrightarrow{op1} \{a, b, c\}$

Operacja odłączania - op3 przekształca podział w taki sposób, że usuwa część lub wszystkie obiekty z wybranej klasy podziału i lokuje je w specjalnym buforze pomocniczym. Jeśli bufor pomocniczy jest niepusty, obiekty są do niego dodawane.

Operacja dołączania - op4 przekształca podział w taki sposób, że usuwa część lub wszystkie obiekty z bufora pomocniczego i dołącza je do istniejącej klasy podziału albo tworzy z nich nową klasę podziału.

### Przykład

Niech dany będzie zbiór  $U = u_1, u_2, \dots, u_{10}$  oraz relacje równoważności  $P, Q \in \mathbb{E}(U)$  reprezentowane podziałami:

$$\mathbb{K}_P = \{\{u_1, u_2, u_3\}, \{u_4\}, \{u_5, u_6\}, \{u_7, u_8, u_9, u_{10}\}\}$$

$$\mathbb{K}_Q = \{\{u_1\}, \{u_2, u_3, u_4, u_5\}, \{u_6\}, \{u_7, u_8\}, \{u_9, u_{10}\}\}$$

Oblicz  $\delta^{1,2}(P, Q)$  oraz  $\delta^{3,4}(P, Q)$ .

Rozwiązanie

1.  $\delta^{1,2} = 4lub6$

2.  $\delta^{3,4} = 4$

## 2 Aproksymacje

### Przybliżenie zbioru

Mówimy, że dla zadanego uniwersum  $U$ , określona została przestrzeń aproksymacyjna  $\mathbb{K} = K_1, K_2, \dots, K_M$ , wtedy i tylko wtedy gdy  $\bigcup_{i=1}^M K_i = U$  oraz  $K_i \cap K_j = \emptyset$  dla  $i \neq j$ . Przybliżeniem zbioru  $D \subseteq U$  w przestrzeni aproksymacyjnej  $\mathbb{K}$  (określonym także terminem zbiór przybliżony) nazywamy parę:

$$\text{apr}(D) = (\text{apr}_{\mathbb{K}}, \text{apr}^{\mathbb{K}}(D))$$

gdzie

$$\text{apr}^{\mathbb{K}}(D) = \bigcup_{K \in \mathbb{K}, K \cap D \neq \emptyset} K \quad \text{apr}_{\mathbb{K}}(D) = \bigcup_{K \in \mathbb{K}, K \subseteq D} K$$

$\text{apr}^{\mathbb{K}}(D)$  – przybliżenie zbioru  $D$  z góry

$\text{apr}_{\mathbb{K}}(D)$  – przybliżenie zbioru  $D$  z dołu

### Miary jakości przybliżenia

$$\text{completeness}(\text{apr}(Z)) = \frac{|\text{apr}_{\mathbb{K}}(Z)|}{|Z|}$$

$$\text{precision}(\text{apr}(Z)) = \frac{|\text{apr}^{\mathbb{K}}(Z)|}{|Z|}$$

### Przykład

Niech  $U = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$  oraz niech  $\mathbb{K} = \{\{u_1, u_2\}, \{u_3, u_4, u_5\}, \{u_6, u_7\}\}$ . Wyznaczyć dolną i górną aproksymację dla  $Z_2 = \{u_1, u_2, u_3, u_4\}$ .

$K_i$	$K_i \cap Z_2$	$K_i \subseteq Z_2$
$\{u_1, u_2\}$	$\{u_1, u_2\}$	$\{u_1, u_2\}$
$\{u_3, u_4, u_5\}$	$\{u_3, u_4\}$	$\emptyset$
$\{u_6, u_7\}$	$\emptyset$	$\emptyset$

$$\text{apr}_{\mathbb{K}}(Z_2) = \{u_1, u_2\}$$

$$\text{apr}^{\mathbb{K}}(Z_2) = \{u_1, u_2, u_3, u_4\}$$

Obliczyć kompletność i miarę jakości przybliżenia dla  $Z_2$ .

$$\text{completeness}(\text{apr}(Z_2)) = \frac{|\text{apr}_{\mathbb{K}}(Z_2)|}{|Z_2|} = \frac{2}{4} = 0.5$$

$$\text{precision}(\text{apr}(Z_2)) = \frac{|\text{apr}^{\mathbb{K}}(Z_2)|}{|Z_2|} = \frac{4}{4} = 1$$

### Aproksymacje - Na chłopski rozum

Przybliżenie dolne - jest to zbiór elementów, które w całości zawierają się w zbiorze przybliżanym.

Przybliżenie górne - zbiór wszystkich elementów, które mają chociaż jeden element wspólny z przybliżanym zbiorem.

Przybliżenie odgórnie równe jest wtedy kiedy przybliżenia górne zbiorów są takie same.

Zapis  $X =^{\mathbb{K}} Y$

Przybliżenie oddolnie równe jest wtedy kiedy przybliżenia dolne zbiorów są takie same.

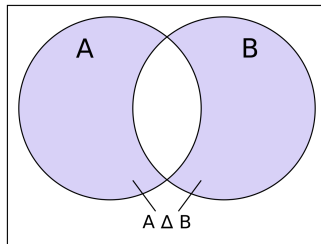
Zapis  $X =_{\mathbb{K}} Y$

Zbiory równe w przestrzeni aproksymacyjnej  $\mathbb{K}$  wtedy i tylko wtedy gdy  $X =^{\mathbb{K}} Y$  oraz

$X =_{\mathbb{K}} Y$ . Piszemy  $X = \frac{\mathbb{K}}{\mathbb{K}} Y$

## 3 Wyznaczanie reprezentanta

Różnica symetryczna zbiorów  $A$  i  $B$  to zbiór elementów należących do  $A$  lub  $B$ , ale nie należących do  $A \cap B$ .



Do wyznaczania podobieństwa można użyć wzór

$$\sigma = \frac{1}{1 + \delta}$$

## Przykład

Niech dane będzie uniwersum  $U = \{a, b, c, d, e, f\}$  z funkcją odległości  $\delta : U \times U \rightarrow R_+$  jako makrostrukturą:

$\delta$	a	b	c	d	e	f
a	0	2	3	1	8	6
b	2	0	2	1	10	9
c	3	2	0	2	11	8
d	1	1	2	0	10	6
e	8	10	11	10	0	1
f	6	9	8	6	1	0

a) Wyznacz medoid (medoidy) dla profilu  $P = \{2a, b, 3d, e, 2f\}$ .

		a	b	c	d	e	f
a	2	0	4	6	2	16	12
b	1	2	0	2	1	10	9
c	0	0	0	0	0	0	0
d	3	3	3	6	0	30	18
e	1	8	10	10	10	0	1
f	2	12	18	16	12	2	0
		25	35	40	25	58	40

Jest to  $a$  lub  $d$ .

b) Wyznacz centroid (centroidy) dla profilu  $P = \{2a, d\}$ .

		a	b	c	d	e	f
a	2	0	4	6	2	16	12
b	0	0	0	0	0	0	0
c	0	0	0	0	0	0	0
d	1	1	1	2	0	10	6
e	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0
		1	5	8	2	26	18

Jest to  $a$ .

c) Wyznacz reprezentanta profilu  $P = \{2a, b, e\}$  w zbiorze  $B = \{e, f\}$ .

		a	b	c	d	e	f
a	2	0	4	6	2	16	12
b	1	2	0	2	1	10	9
c	0	0	0	0	0	0	0
d	0	0	0	0	0	0	0
e	1	8	10	10	10	0	1
f	0	0	0	0	0	0	0
		10	14	18	13	26	22

Jest to  $f$ .

## Przykład

Stosując miarę 2Dice wyznaczyć reprezentację zdefiniowanego poniżej profilu zbiorów

Zbiór obiektów:  $U = \{a, b, c\}$

Zbiory profilu:  $P_1 = \{a\}, P_2 = \{b, c\}, P_3 = \{a, b\}$

Profil:  $P = \{2P_1, 3P_2, P_3\}$

$$P_{2Dic} = \frac{2|X \cap Y|}{|X| + |Y|}$$

	$P_1$	$P_2$	$P_3$
$P_1$	1	0	$2/3$
$P_2$	0	1	$1/2$
$P_3$	$2/3$	$1/2$	1

$$\Sigma \text{ dla } P_1 = \frac{2}{2} \cdot 2 + 0 + \frac{2}{3} = 2\frac{2}{3}$$

$$\Sigma \text{ dla } P_2 = \frac{2 \cdot 0}{2} \cdot 2 + 3 \cdot \frac{2 \cdot 2}{2} + \frac{1}{2} = 6\frac{1}{2} \quad \text{MAX}$$

$$\Sigma \text{ dla } P_3 = \frac{2}{3} \cdot 2 + 3 \cdot \frac{1}{2} + 1 = \frac{4}{3} + \frac{3}{2} + 1 = \frac{23}{6}$$

Najlepszym reprezentantem jest  $P_2$ , bo dla niego uzyskano najwyższe podobieństwo dla całego profilu.

## 4 Przydatne

Operator min - iloczyn

Operator max - suma

4. Niech dana będzie tabela:

U	Cena	Rok	Stan
$u_1$	wysoka	1966	dobry
$u_2$	niska	1971	zły
$u_3$	średnia	1971	dobry
$u_4$	wysoka	1968	dobry
$u_5$	niska	1968	zły

gdzie  $u_1, u_2, u_3, u_4, u_5$  są obiektami, Cena, Rok, Stan są tzw. atrybutami (cechami obiektów), a w komórkach wpisane są wartości atrybutów wymienionych w kolumnie dla obiektów podanych w wierszach.

Mówimy, że obiekty  $x \in U$  i  $y \in U$  są nierozróżnialne ze względu na atrybut  $a \in \{\text{Cena, Rok, Stan}\}$  wtedy i tylko wtedy, gdy wartości atrybutu  $a$  dla obiektów  $x$  i  $y$  są równe (symbolicznie piszemy  $x \equiv^a y$ ). Pokaż, że relacja nierozróżnialności ze względu na atrybut jest relacją równoważności. Zwróć uwagę, że dla dowolnego atrybutu  $a$  generuje ona podział  $\text{Klasy}(\{a\})$  znany z listy piętej.

$$\delta^{RS}(\equiv^{Cena}, \equiv^{Rok}) = 8$$

$$\delta^{1,2}(\equiv^{Cena}, \equiv^{Rok}) = 4$$

Operacja Rozbicia – op1

Operacja Sklejenia – op2

$$\begin{aligned} \mathcal{K}_{\equiv^{Cena}} &= \{\{u_1, u_4\}, \{u_2, u_5\}, \{u_3\}\} \\ \text{op1} \downarrow & \{ \{u_1, u_4\}, \{u_2, u_5\}, \{u_3\} \} \\ \text{op2} \downarrow & \{ \{u_1, u_4\}, \{u_2, u_5\}, \{u_3\} \} \\ \text{op1} \downarrow & \{ \{u_1, u_4\}, \{u_2, u_5\}, \{u_3\} \} \\ \text{op2} \downarrow & \{ \{u_1, u_4\}, \{u_2, u_5\}, \{u_3\} \} \\ \text{op1} \downarrow & \{ \{u_1, u_4\}, \{u_2, u_5\}, \{u_3\} \} = \mathcal{K}_{\equiv^{Rok}} \end{aligned}$$

$\equiv^{Cena}$   

1	0	0	1	0
0	1	0	0	1
0	0	1	0	0
1	0	0	1	0
0	1	0	0	1

$\equiv^{Rok}$   

1	0	0	0	0
0	1	1	0	0
0	1	1	0	0
0	0	0	1	1
0	0	0	1	1