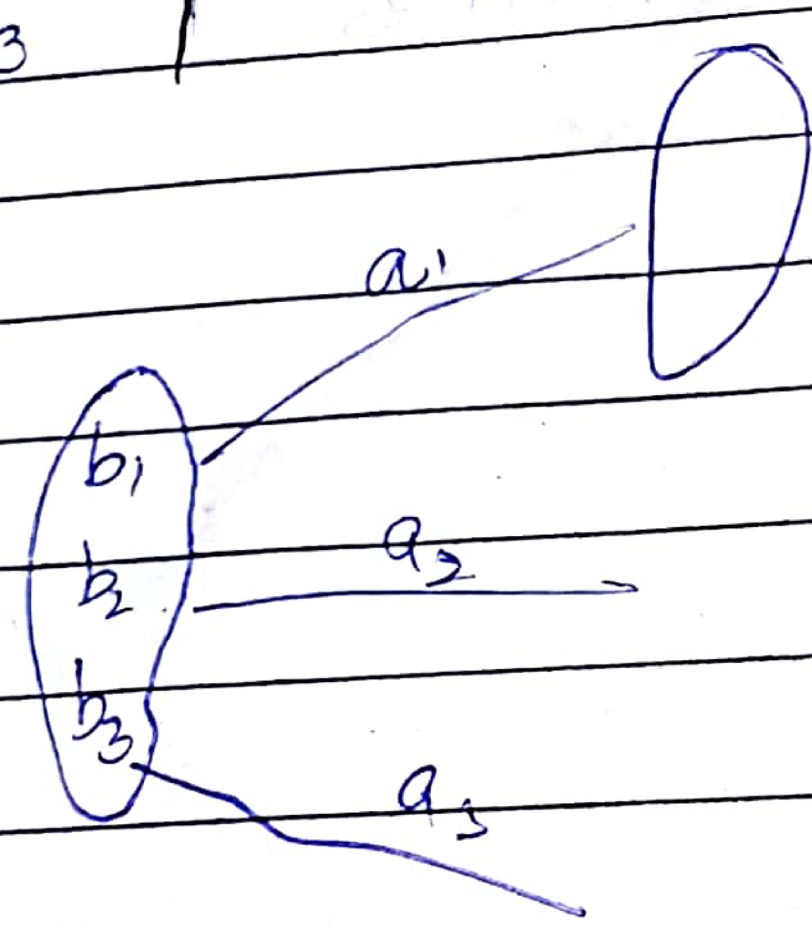
★ Rough Set Theory:

U	a_1	a_2	a_3	d
b_1				
b_2				
b_3				

$a : U \rightarrow V_a$

$a_1, a_2, a_3 \rightarrow$ condition attribute

$d \rightarrow$ decision attribute



a (Age, Height) — C (Play Volleyball?)

	Age	Height	Play Volleyball?
x_1	0–10	1–3	No
x_2	0–10	3–4.5	Yes
x_3	10–20	4–6	Yes
x_4	10–20	4–6	No
x_5	20–30	5–6.5	Yes
x_6	20–30	5–6.5	Yes
x_7	30–40	6–7	Yes
x_8	20–30	5–6.5	

DELTA Page No.
Date

→ Indiscernible objects :

Ind (B)

$[x]_{age}$ → equiv class of x over age

= $\{x_1, x_2\}$ $\{x_3, x_4\}$ $\{x_6\}$ $\{x_7\}$

→ Write relation :→

~~(x_1, x_2)~~

$$\left\{ (x, x') \in U^2 \mid \forall a \in b,\ V_a(x) = V_a(x') \right\}$$

IND(B) =

$$\left\{ (x, x') \in U^2 \mid \forall a \in b,\ V_a(x) = V_a(x') \right\}$$

IND(age) = ~~$\{ (x, x')$~~

$= \{ \{x_1, x_2\}, \{x_3, x_4\}, \{x_5\}, \{x_7\} \}$

IND (height) = $\{ x_1, x_2, \{x_3, x_4\}, \{x_5, x_6\}, x_7 \}$

IND (Age, Height)

$= \{ x_1, x_2, \{x_3, x_4\}, \{x_5, x_6\}, x_7 \}$

DELTA Page No.
Date

$$\cancel{X_{c_j} = \{ x \mid c(x) = V_{c_j} \}}$$

$$X_{Yes} = \{ x_2, x_3, x_5, x_6, x_7 \}$$

$$X_{No} = \{ x_1, x_4 \}$$

→ Lower Boundary over a subset of attributes B,

$$\underline{B}_Y(X) = \{ x \mid [x]_B \subseteq X_Y \}$$

Note: $[x]_B$ → equivalence class of x over attribute B

$$B_Y(\text{Age, Height})(X) = \{ x_2, x_5, x_6, x_7 \}$$

→ Upper approximation:

$$\overline{B(X)}_Y = \{ x \mid [x]_B \cap X_Y \neq \phi \}$$

$$\overline{AH}(X)_Y = \{ x_2, x_3, x_4, x_5, x_6, x_7 \}$$

↓

Age, Height

→ Boundary Region = Upper appx − Lower appx

$$\underline{BX}_Y = \overline{B}(x)_Y - \underline{B}(x)_Y$$

$$= \{ x_3, x_4 \}$$

Outside B

NOW

IN

→ Outside Region:

$B^{\circ}(X)_y = U - \overline{B(x)}_y$

$= x_1$

→ Now for Age:

$IND(age) = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, x_7\}$

$\underline{B}(x)_y = \{x_5, x_6, x_7\}$

$\overline{B}(X)_y = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$

$\underline{\overline{B(X)}}_y = \{x_1, x_2, x_3, x_4\}$

$B^{\circ}(X)_y = U - \overline{B(x)}_y$

$= \{\phi\}$, means not externally defined in terms of age

→ Now for Age, Height, Play:

$IND(AHP) = \{x_1, x_2, x_3, x_4, \{x_5, x_6\}, x_7\}$

$\underline{B}(x)_y = \{x_2, x_3, x_5, x_6, x_7\}$

$\overline{B}(x)_y = \{x_2, x_3, x_5, x_6, x_7\}$

$\underline{B}(x)_y = \overline{B(x)}_y$

$\underline{\overline{B}}(x)_y = \phi$, Boundary region empty

$\underline{B}(x)_y = \{x_1, x_4\}$

Classes of Approximation:

1. **Roughly Definable:**

$\underline{B}(x)_{c_j} \neq \phi$ and $\overline{B(X)}_{c_j} \neq U$

2. **Internally UnDefinable:**

$\underline{B}(X)_{c_j} = \phi$ and $\overline{B(X)}_{c_j} \neq U$

→ No one is in for sure, but atleast some are excluded.

3. **Externally UnDefinable:**

$\underline{B}(x) \neq \phi$ and $\overline{B(x)} = U$

→ Some definetely yes, but we are not sure that anyone of is NO.

4. **Totally UnDefinable:**

$\underline{B}(X) = \phi$ and $\overline{B(X)} = U$

5. **Crisp:**

$\underline{B}(X) = \overline{B(X)} = U$

→ ~~Ideal~~

→ All deterministic.

→ No one in boundary.

Accuracy

DELTA Page No.
Date

→ Accuracy of Approximation for given destination class X wrt target attributes B towards X

$$\alpha_B(x) = \frac{\underline{B}(x)}{\overline{B}(x)} = \frac{\underline{B}(X)}{\overline{B}(X)}$$

→ 1 crisp → may result in overfitting
→ < 1 rough

U	diploma	Exp	French	Reference	Decision
x_1	MBA	Med	Y	Excellent	Accept
x_2	MBA	Low	Y	Neutral	Reject
x_3	~~MBA~~ MCE	Low	Y	Good	R
x_4	MSc	High	Y	N	A
x_5	MSc	Med	Y	N	R
x_6	MSc	High	Y	E	A
x_7	MBA	High	N	G	A
x_8	MCE	Low	N	E	R

$X_{accept} = x_1, x_4, x_6, x_7$

$X_{reject} = x_2, x_3, x_5, x_8$

Let B = DEFR

~~B(Xacc) 7~~

$IND(B) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}\}$

$\underline{B(X)}_Y = \{x_1, x_4, x_6, x_7\}$

$\cancel{\overline{B(X)} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}}$

$\overline{B(X)}_Y = \{x_1, x_4, x_6, x_7\}$

$B^{\circ}(X)_Y = U - \overline{B(X)}_Y$

$= x_2, x_3, x_5, x_8$

$\overline{\underline{B(X)}}_Y = \overline{B(X)}_Y - \underline{B(X)}_Y$

$= \phi$

$X_{reject} = \{x_2, x_3, x_5, x_8\}$

$\underline{B(X)}_N = x_2, x_3, x_5, x_8$

$\overline{B(X)}_N = x_2, x_3, x_5, x_8$

$B^{\circ}(X)_N = U - \overline{B(X)}_N$

$= x_2, x_3, x_5, x_8$

$\overline{\underline{B(X)}}_N = \overline{B(X)}_N - \underline{B(X)}_N$

$= \cancel{x_2} \ \phi$

→ **Now $B = ER$:**

$IND(B) = \{x_1, x_2, \dots, x_8\}$

Same results as $B = DEFR$

So, it is better to reduce the attributes to ER as the approximation remains the same.

★ Indiscernable Matrix :

x_1 and x_2 → diff in E and R.

$\therefore E \vee R$

~~Difference many one sign make them indiscernable~~

	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	$E \vee R$						
x_2	–						
x_3	$d \vee R$	–					
x_4	$d \vee E$		–				
x_5	$d \vee E$			–			
x_6					–		
x_7						–	
x_8							–

$$C_{ij} = \{a \in A \mid a(x_i) \neq a(x_j)\} \quad \forall\, i,j = 1 \dots N$$

e.g., $C_{12} = \{e, r\}$

$C_{13} = \{d, e, r\}$

→ Discernability f^n:

$$C_{ij}^* = \{ a^* / a \in C_{ij} \}$$

e.g., $C_{13} = \{d, e, r\}$ $\quad C_{13}^* = d \vee e \vee r$

Dis f^n,

$$\wedge \{ \vee C_{ij}^* \}$$

$1 \le i, j \le N$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	—	—	—	—	—	—	—	—
x_2	$e \vee r$	—	—	—	—	—	—	—
x_3	$d \vee e \vee r$	$d \vee r$	—	—	—	—	—	—
x_4	$d \vee e \vee r$	$d \vee e$	$d \vee f \vee r$	—	—	—	—	—
x_5	$d \vee r$	$d \vee e$	$d \vee e \vee r$	e ✓	—	—	—	—
x_6	$d \vee e$	$d \vee e \vee r$	$d \vee e \vee r$	r ✓	$e \vee r$	—	—	—
x_7	$e \vee f \vee r$	$e \vee f \vee r$	$d \vee e \vee f$	$d \vee f \vee r$	$d \vee e \vee f \vee r$	$d \vee f \vee r$	—	—
x_8	$d \vee e \vee f$	$d \vee f \vee r$	$f \vee r$	$d \vee e \vee f \vee r$	$d \vee e \vee f \vee r$	$d \vee e \vee f$	$d \vee e \vee r$	—

de \ fr	00	01	11	10
00	0	0	0	0
01	0	1	1	0
11	0	1	1	0
10	0	0	0	0

DELTA Page No.
Date

$$fr = \frac{(e \vee R)(d \vee e \vee R)(d \vee R)(d \vee e)}{(e \vee f \vee R)(d \vee e \vee f)(d \vee f \vee R)}$$

$e(d + r)$ (NI)

$ed + dr$

$\mathcal{E}q_v \left(IND_{R_i}(X_k) \right)$