CSE 422 ASSIGNMENT 2

NAME: ANIKA ISLAM

ID: 21101298

SECTION: 06

(1) (a) $P(\text{cheat} \wedge \text{pass} | \text{study}) \neq P(\text{cheat} | \text{study}) * P(\text{pass} | \text{study})$

$$\Rightarrow \frac{P(\text{cheat} \wedge \text{pass} \wedge \text{study})}{P(\text{study})} \neq \frac{P(\text{cheat} \wedge \text{study})}{P(\text{study})} * \frac{P(\text{pass} \wedge \text{study})}{P(\text{study})}$$

$$\Rightarrow \frac{0.25}{0.45} \neq \frac{0.27}{0.45} * \frac{0.40}{0.45}$$

$$\Rightarrow 0.556 \neq 0.533$$

$\therefore$ cheat and pass are not conditionally independent given study.

(b) $P(\text{Pass or Cheat}) = P(\text{Pass}) + P(\text{cheat}) - P(\text{Pass} \wedge \text{cheat})$

$= 0.63 + 0.59 - 0.35$

$= 0.87$

CamScanner

(2) (a) $P(\sim Cold) = 0.32 + 0.06 + 0.26 + 0.03$
$= 0.67$

(b) $$\frac{P(\sim Cloudy)}{P(\sim Rain \wedge \sim Cold)} = \frac{P(\sim Cloudy \wedge \sim Rain \wedge \sim Cold)}{P(\sim Rain \wedge \sim Cloud)}$$

$$= \frac{0.07}{0.04 + 0.07}$$

$$= 0.636$$

(c) $$\frac{P(\sim Rain\ )}{P(\sim Cloudy)} = \frac{P(\sim Rain \wedge \sim Cloudy)}{P(\sim Cloudy)}$$

$$= \frac{0.03 + 0.07}{0.26 + 0.03 + 0.10 + 0.07}$$

$$= 0.217$$

(d) $P(\sim Rain \text{ or } Cloudy) = P(\sim Rain) + P(Cloudy) - P(\sim Rain \wedge Cloudy)$
$= 0.2 + 0.54 - 0.1$
$= 0.64$

(3) (a) $P(\text{football} \wedge \text{left-handed person}) = 0.13$

(b) $\dfrac{P(\text{right-handed person})}{P(\text{cricket})} = \dfrac{P(\text{cricket} \wedge \text{right-handed person})}{P(\text{right-handed person})}$

$$= \frac{0.1}{0.1 + 0.1 + 0.26}$$

$$= 0.217$$

(c) $P(\text{football} \wedge \text{cricket}) = 0$

(d) $P(\text{right-handed or left handed})$

$= P(\text{right-handed}) + P(\text{left-handed}) - P(\text{right-handed} \wedge \text{left\_handed})$

$= 0.46 + 0.54 - 0$

$= 1$

(e) $P(\text{football} \wedge \text{Right-handed}) \neq P(\text{football}) * P(\text{Right-handed})$

$0.1 \neq (0.13 + 0.1) * (0.1 + 0.1 + 0.26)$

$0.1 \neq 0.115$

Football and right-handed depends on one another.

(4) Cov - Covid +ve

!Cov - Covid -ve

Pos - detected Covid +ve

!Pos — detected covid -ve

$P(Pos|Cov) = 0.92$

$P(!Pos|!Cov) = 1 - 0.92 = 0.08$

$P(Pos|!Cov) = 0.04$

$P(!Pos|!Cov) = 1 - 0.04 = 0.96$

$P(Cov) = 0.07$ $\quad P(!Cov) = 1 - 0.07 = 0.93$

$$P(Cov) = \frac{\cancel{P(Pos}}{P(Pos|Cov) * P(Cov) + P(Pos|!Cov) * P(!Cov)}$$

$$P(Cov) = \frac{\cancel{(0.92 * 0.07)}}{(0.92 * 0.07) + (0.04 * 0.93)}$$

~~= 0.634~~ 0

$P(Cov) = 0.1016$

(b) Learning phase

Outlook

	Yes	No
Overcast	2/5	0/3
Sunny	2/5	3/3
Rainy	1/5	0/3

Humidity

	Yes	No
Cool	2/5	0/3
mild	3/5	2/3
Hot	0/5	1/3

Temp

	Yes	No
Normal	4/5	0/3
High	1/5	3/3

Wind

	Yes	No
TRUE	3/5	1/3
FALSE	2/5	2/3

$P(\text{Play Tennis} = \text{Yes}) = \frac{5}{8}$ | $P(\text{Play Tennis} = \text{NO}) = \frac{3}{8}$

(a) P(Play Tennis = Yes | Outlook = Sunny, humidity = mild, Temp = ~~High~~ Normal, Wind = True)

= P(Outlook = Sunny | Play Tennis = Yes) *
P(humidity = mild | Play Tennis = Yes) *
P(Temp = ~~High~~ Normal | Play Tennis = Yes) *
P(Wind = True | Play Tennis = Yes) * P(Play Tennis = Yes)

$= (2/5)(3/5)(4/5)(3/5)(5/8)$

= [illegible]

= ~~0.00276~~ ~~0.018~~ 0.072

P(Play Tennis = No | Outlook = sunny, humidity = mild,
Temp = High, Wind = True)

= P(Outlook = Sunny | Play Tennis = No) *
P(humidity = mild | play Tennis = No) *
P(Temp = High | Play Tennis = No) *
P(~~Te~~ Wind = True | Play Tennis = No) * P(Play Tennis = No)

= (3/3)(2/3)(0/3)(1/3)(3/8)

= ~~0.0169~~ 0.~~0833~~ 0

probability of playing tennis ~~&~~ > probability of not playing tennis.

∴ Player is not going to play tennis.

(b) P(Play Tennis = Yes | outlook = overcast, humidity = hot)

= P(outlook = overcast | Play Tennis = Yes) * P(Humidity = hot | Play Tennis = Yes)
* P(Play Tennis = Yes)

= (2/2)(0/1)(5/8)

= 0

P(Play Tennis = No | outlook = overcast, humidity = hot)

= P(outlook = overcast | Play Tennis = No) * P(humidity = hot | Play Tennis = No)
* P(Play Tennis = No)

= (0/2)(1/1)(3/8)

= 0

No decision can be made

(b)(a) $E(\text{Edibible}) = -P(\text{Edible} = \text{Yes}) \log_2 P(\text{Edible} = \text{Yes})$

$$- P(\text{Edible} = \text{No}) \log_2 P(\text{Edible} = \text{No})$$

$$= -\left(\frac{9}{16} \log_2 \frac{9}{16}\right) - \left(\frac{7}{16} \log_2 \frac{7}{16}\right)$$

$$= 0.989$$

Using color:-

$E(\text{color} = \text{Yellow}) = -P(\text{Edible} = \text{Yes} \mid \text{Color} = \text{Yellow}) \log_2 P(\text{Edible} = \text{Yes} \mid \text{Color} = \text{Yellow})$

$$- P(\text{Edible} = \text{NO} \mid \text{Color} = \text{Yellow}) \log_2 P(\text{Edible} = \text{No} \mid \text{Color} = \text{Yellow})$$

$$= -\left(\frac{8}{13} \log_2 \frac{8}{13}\right) - \left(\frac{5}{13} \log_2 \frac{5}{13}\right)$$

$$= 0.961$$

$P(\text{color} = \text{Yellow}) = \frac{13}{16}$

$E(\text{color} = \text{Green}) = -P(\text{Edible} = \text{Yes} \mid \text{Color} = \text{Green}) \log_2 P(\text{Edible} = \text{Yes} \mid \text{Color} = \text{Green})$

$$- P(\text{Edible} = \text{No} \mid \text{Color} = \text{Green}) \log_2 P(\text{Edible} = \text{No} \mid \text{Color} = \text{Green})$$

$$= -\left(\frac{1}{3} \log_2 \frac{1}{3}\right) - \left(\frac{2}{3} \log_2 \frac{2}{3}\right)$$

$$= 0.918$$

$P(\text{color} = \text{Green}) = \frac{3}{16}$

Using size,

$$E(size = small) = -\frac{6}{8}\log_2\frac{6}{8} - \frac{2}{8}\log_2\frac{2}{8}$$

$$= 0.811$$

$$P(size = small) = \frac{8}{16}$$

$$E(size = large) = -\frac{3}{8}\log_2\frac{3}{8} - \frac{5}{8}\log_2\frac{5}{8}$$

$$= 0.954$$

$$P(size = large) = \frac{8}{16}$$

Using shape,

$$E(shape = Round) = -\frac{6}{12}\log_2\frac{6}{12} - \frac{6}{12}\log_2\frac{6}{12}$$

$$= 1$$

$$P(shape = Round) = \frac{12}{16}$$

$$E(shape = Irregular) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4}$$

$$= 0.811$$

$$P(shape = Irregular) = \frac{4}{16}$$

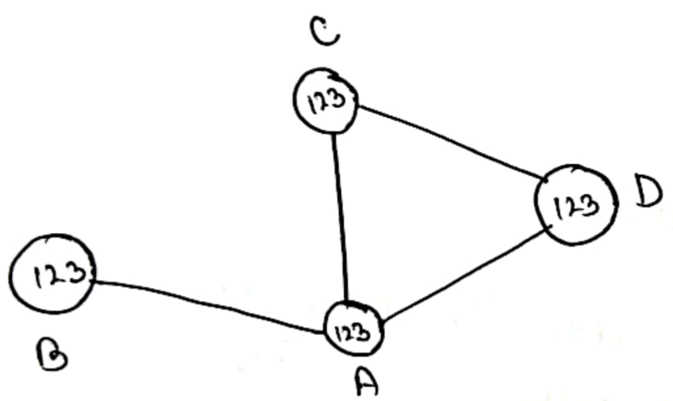
(7) (a) variable = {A, B, C, D}

domain = {1, 2, 3}

constraint = no two adjacent regions can have same digit

goal = All regions have number

(b)



Since all regions have same domain, most constrained variable cannot be used.

We can use most constraining variable.

C can reduce A and D

D can reduce A and C

A can reduce C, B and D

B can reduce A

Here, A can reduce maximum regions.
So, A is choosen to give a number.

(b) $IG(color) = E(Decision) - E(color = Yellow)\,P(color = Yellow) - E(color = Green)\,P(color = Green)$

$= 0.989 - (0.961 * 13/16) - (0.918 * 3/16)$

$= 0.0361$

$IG(size) = 0.989 - (0.811 * 8/16) - (0.954 * 8/16)$

$= 0.1065$

$IG(shape) = 0.989 - (1 * 12/16) - (0.811 * 4/16)$

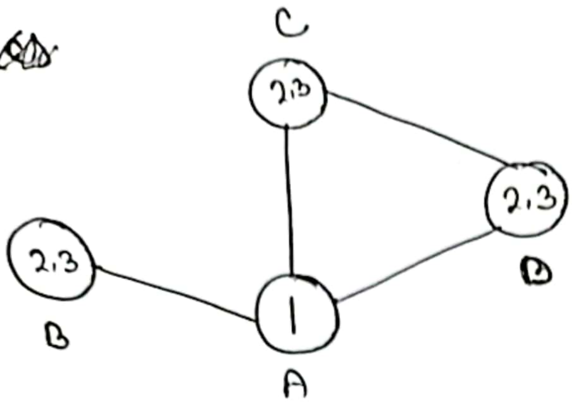
$= 0.03625$

$IG(size) > IG(shape) > IG(color)$

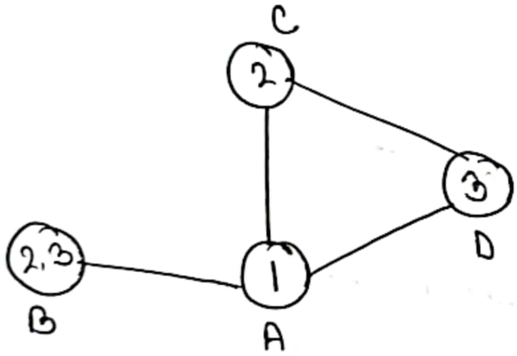
size is a better feature as it has the highest IG than the other features so it has more information and no information conflicts.

CamScanner

Ans.

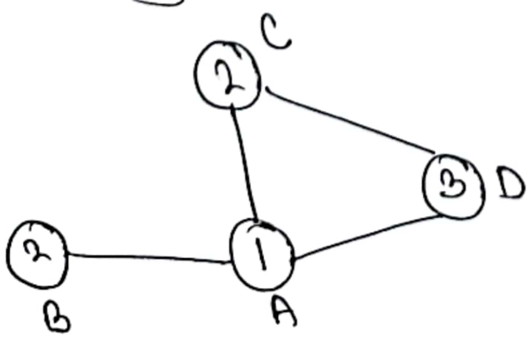


Again, domains are the same and most constraining variable needs to be used.
Now, B, C and D all can reduce maximum regions.
So, C is choosen randomly and assigned with 2



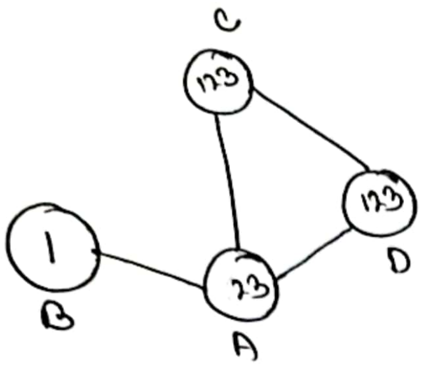
Domain of D < Domain of B.
So, most constrained variable is used and D has been assigned with 3.

Now, the remaining region B has been assigned with 2



order $\Rightarrow$ A $\rightarrow$ C $\rightarrow$ D $\rightarrow$ B

(c)



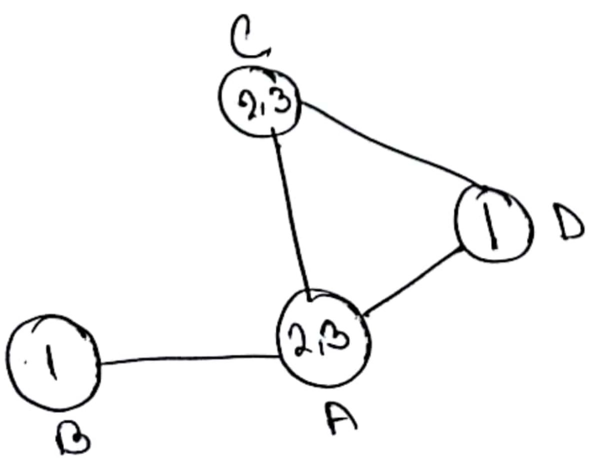
If D is assigned with 2, it can reduce A to {3} and C to {1,3}

If D is assigned with 3, it can reduce A to {2} and C to {1,2}

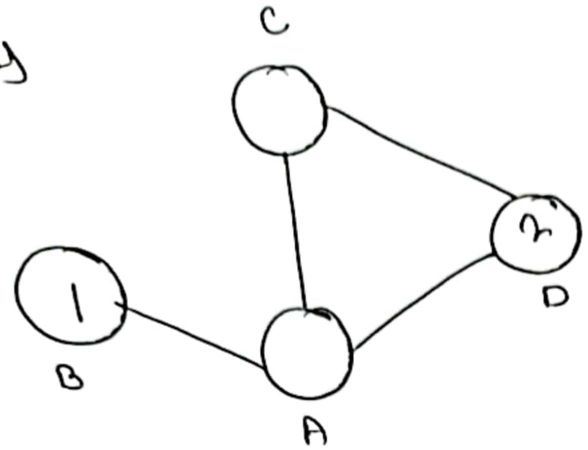
If D is assigned with 1, it can reduce C to {2,3} and cannot reduce A.

Thus, least number of regions have been reduced using D with 1.

D is assigned with 1.



(d)

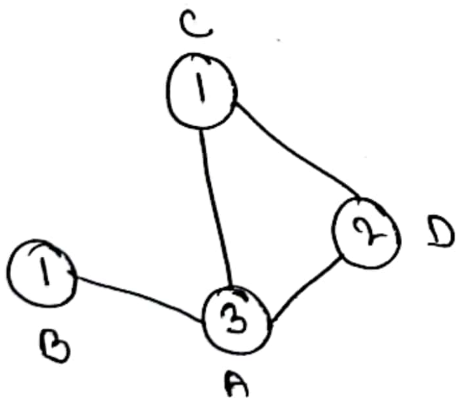


3	1	1,3	2
A	B	C	D

A is assigned with 3

3	1	1	2
A	B	C	D

C is assigned with 1



No region is empty after assigning and reducing every regions.
So, CSP remain arc consistent.