SPARSH AGARWAL
CS 576
HW 6

A1 (a) P(A, B, C)

A	B	C	
T	T	T	0.024
T	T	F	0.096
T	F	T	0.196
T	F	F	0.084
F	T	T	0.084
F	T	F	0.336
F	F	T	0.126
F	F	F	0.054

(b) $P(A=T \mid C=F) = \frac{P(A=T, C=F)}{P(C=F)} = \frac{0.084+0.096}{0.84+0.96+0.336+0.054}$

$= \frac{0.18}{0.57}$

$= 0.3157894...$

(c) $P(A=T \mid B=F) = \frac{P(A=T, B=F)}{P(B=F)} = \frac{0.28}{0.46}$

$= 0.608695652$

(d) $P(A=T \mid B=T, C=F) = \frac{P(A=T, B=T, C=F)}{P(B=T, C=F)}$

$= \frac{0.096}{0.096 + 0.336}$

$= 0.22222\overline{2}$

(e) $P(C=T \mid A=T) = \frac{P(C=T, A=T)}{P(A=T)} = \frac{0.024 + 0.196}{0.024 + 0.196 + 0.096 + 0.084}$

$= \frac{0.22}{0.4} = 0.55$

$P(C=T \mid A=F) = \frac{P(C=T, A=F)}{P(A=F)} = \frac{0.084 + 0.126}{0.6} = 0.35$

C is not independent of A because probability of C = True depends on probability of A being True or false

Or we can check $P(C=T \mid A=T)$ & $P(C=T)$ (they will not be equal ∴ C is dependent on A)

~~P(C, A | B) = P(A|B) * P(C|B)~~

(1) $P(C, A \mid B) = P(A \mid B) * P(C \mid B)$

if this is true then A & C are independent given B.

$P(C=T, A=T \mid B=T)$

$= \frac{0.024}{0.54}$

$= \frac{0.024}{0.54}$

$P(A=T \mid B=T) * P(C=T \mid B=T)$

$= \frac{0.12}{0.54} * \frac{0.108}{0.54}$

$= \frac{0.0176}{0.54}$

$\therefore P(C, A \mid B) \neq P(A \mid B) * P(C \mid B)$

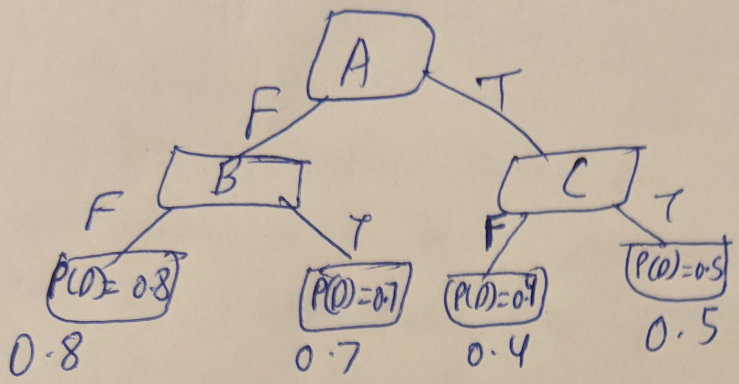
$\therefore$ C is dependent on A given B.

A2) (a)

A	B	C	D	Y=T	Y=F
T	T	T	T	0.2	0.8
T	T	T	F	0.2	0.8
T	T	F	T	0.9	0.1
T	T	F	F	0.15	0.85
T	F	T	T	0.25	0.75
T	F	T	F	0.25	0.75
T	F	F	T	0.25	0.75
T	F	F	F	0.25	0.75
F	T	T	T	0.2	0.8
F	T	T	F	0.2	0.8
F	T	F	T	0.9	0.1
F	T	F	F	0.15	0.85
F	F	T	T	0.75	0.25
F	F	T	F	0.75	0.25
F	F	F	T	0.75	0.25
F	F	F	F	0.75	0.25

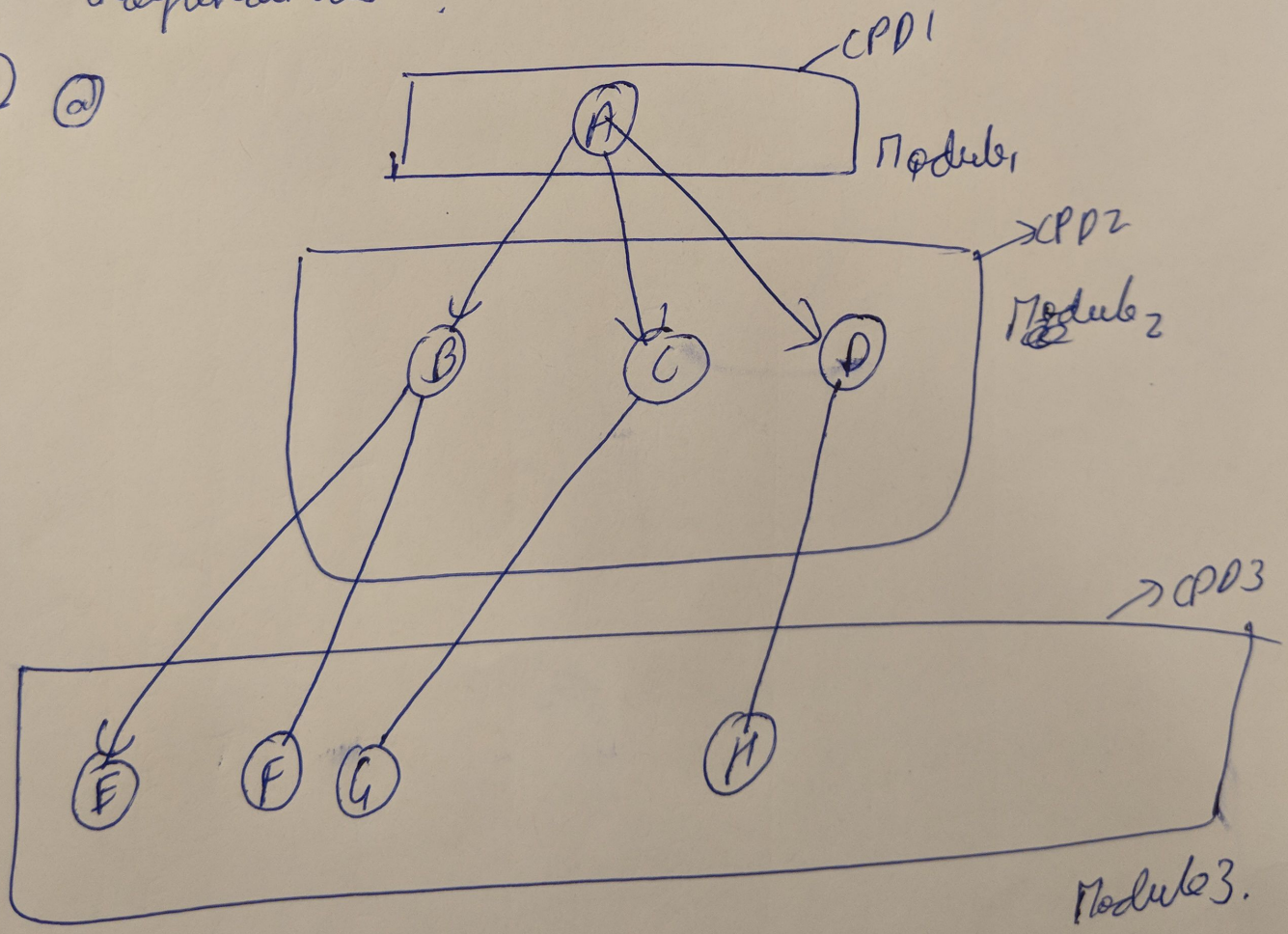
SPARSH AGARWAL

(b)



(c) We can get the independence relations of parameters directly through tree representation.
∴ using tree we can get additional properties of independencies.

Ans (a)



(b) A learned module network is better model than a learned ~~bay~~ bayesian network. ~~Parameter~~ Parameter sharing in the same module allows each parameter to be estimated based on much larger sample. Moreover, this allows us to learn dependencies that are considered ~~to~~ too weak based on statistics of single variable