1.

Example	Input Attributes A	B	C	Class D	#
x_1	t	t	t	Yes	1
x_2	t	t	f	Yes	6
x_3	t	f	t	No	3
x_4	t	f	f	No	1
x_5	f	t	t	Yes	1
x_6	f	t	f	No	6
x_7	f	f	t	Yes	2
x_8	f	f	f	No	2

22

$Pr(D|A) = \frac{7}{11}$ $\quad Pr(D|\bar{A}) = \frac{3}{11}$ $\quad P(A) = 1/2$

$Pr(D|B) = \frac{8}{14}$ $\quad Pr(D|\bar{B}) = \frac{2}{8} = 1/4$ $\quad Pr(B) = 14/22$

$= 7/11$

$= 4/7$

$Pr(D|C) = \frac{4}{7}$ $\quad Pr(D|\bar{C}) = \frac{6}{15}$ $\quad Pr(C) = 7/22$

$$ENT(D|a) = -\left(Pr(d|a).\log_2(Pr(d|a)) + Pr(\bar{d}|a).\log_2(Pr(\bar{d}|a)\right)$$

$$= -\left(\frac{7}{11}.\log_2\left(\frac{7}{11}\right) + \frac{4}{11}\log_2\left(\frac{4}{11}\right)\right)$$

$$= \frac{7}{11} \times 0.6521 + \frac{4}{11} \times 1.4594$$

$$= 0.9456636$$

$$ENT(D|\bar{a}) = -\left(Pr(d|\bar{a}) \cdot \log_2 (Pr(d|\bar{a})) + Pr(\bar{d}|\bar{a}) \cdot \log_2 (Pr(\bar{d}|\bar{a}))\right)$$

$$= -\left(\frac{3}{11} \times \log_2 \frac{3}{11} + \frac{8}{11} \times \log_2 8/11\right)$$

$$= \frac{3}{11} \times 1.8745 + \frac{8}{11} \times 0.4594$$

$$= 0.84533$$

$$ENT(D|A) = \sum_{a} Pr(a) \cdot Pr(D|a) = \frac{1}{2}(0.9456636 + 0.84533)$$

$$= 0.895495$$

$$ENT(D|b) = -\frac{8}{14} \times \log_2 \frac{8}{14} - \frac{6}{14} \log_2 \frac{6}{14}$$

$$= \frac{+8}{14} \times 0.8074 + \frac{6}{14} \times 1.2224$$

$$= 0.985$$

$$ENT(D|\bar{b}) = -\frac{1}{4} \log 1/4 - \frac{3}{4} \log 3/4 = 0.81$$

$$ENT(D|B) = \frac{7}{11} \times 0.985 + \frac{4}{11} \times 0.81 = 0.9213$$

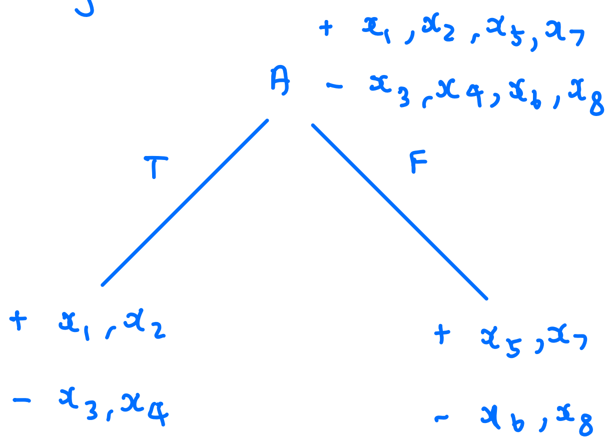
$$ENT(D|c) = -\frac{4}{7}\log\frac{4}{7} - \frac{3}{7}\log\frac{3}{7} = 0.985$$

$$ENT(D|\bar{c}) = \frac{-6}{15}\log\frac{6}{15} - \frac{9}{15}\log\frac{9}{15} = 0.971$$

$$ENT(D|C) = \frac{7}{22} \times 0.985 + \frac{15}{22} \times 0.971$$

$$= 0.9754$$

So splitting on A gives least conditional entropy.



Let us first check A = T

$\{x_1, x_2, x_3, x_4\}$

Example	Input Attributes			Class	#
	A	B	C	D	
x_1	t	t	t	Yes	1
x_2	t	t	f	Yes	6
x_3	t	f	t	No	3
x_4	t	f	f	No	1
					11

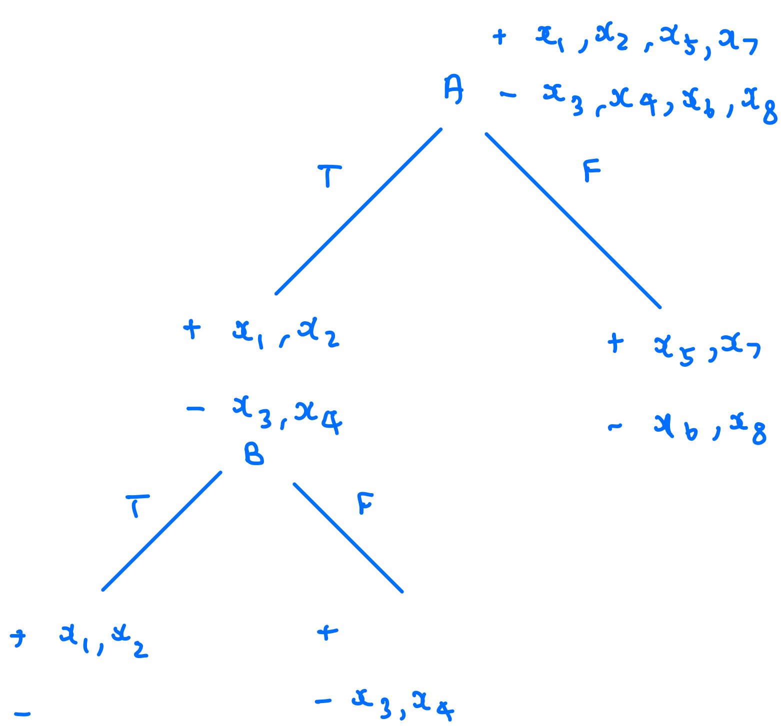
$Pr(D|b) = 1$ $\quad$ $Pr(D|\bar{b}) = 0$ $\quad$ $Pr(B) = 7/11$

$ENT(D|b) = 0$

$ENT(D|\bar{b}) = 0$

$ENT(D|B) = 0$

So no need to check C, we can split by B.



Let us check the A = F

		B	C	D	
x_5	f	t	t	Yes	1
x_6	f	t	f	No	6
x_7	f	f	t	Yes	2
x_8	f	f	f	No	2

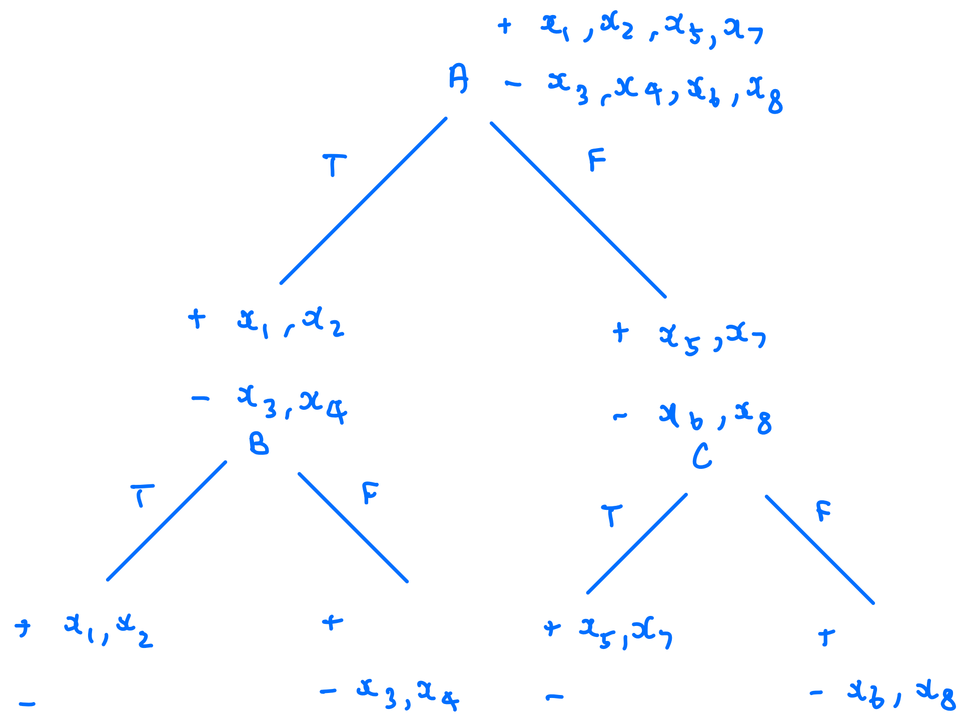
Let us try C

$Pr(D|c) = 1$ $\quad$ $Pr(D|\bar{c}) = 0$ $\quad$ $Pr(c) = 3/11$

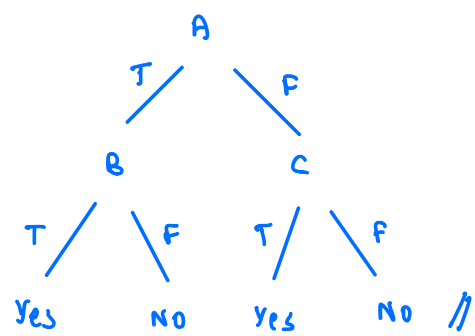
$ENT(D|c) = 0$ $\quad$ $ENT(D|\bar{c}) = 0$

$ENT(D|C) = 0$

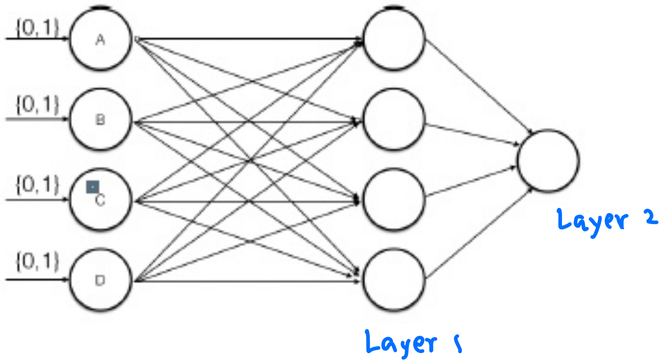
So no need to check B, we can split by C.



DECISION TREE:



2. $(A \vee \neg B) \oplus (\neg C \vee D)$



Let $\Delta = (A \vee \neg B) \oplus (\neg C \vee D)$

$x \oplus y$ is $(x \vee y) \wedge (\neg x \vee \neg y)$

So converting to CNF

$$\Delta = [(A \vee \neg B \vee \neg C \vee D) \wedge (\neg(A \vee \neg B) \vee \neg(\neg C \vee D))]$$

$$= [(A \vee \neg B \vee \neg C \vee D) \wedge ((\neg A \wedge B) \vee (C \wedge \neg D))]$$

$$= [(A \vee \neg B \vee \neg C \vee D) \wedge (\neg A \vee (C \wedge \neg D)) \wedge (B \vee (C \wedge \neg D))]$$

$$= (A \vee \neg B \vee \neg C \vee D) \wedge (\neg A \vee C) \wedge (\neg A \vee \neg D) \wedge (B \vee C) \wedge (B \vee \neg D)$$

So layer 1 neurons implement the clauses.

Layer 2 applies the AND gate across clauses.

Clause 1: $A \vee \neg B \vee \neg C \vee D$

Only false case is

$A = \text{False},\ B = \text{True},\ C = \text{True},\ D = \text{False}$

If $w_A = 1$

$w_B = -1$

$w_C = -1$

$w_D = 1$

False case is

$0 - 1 - 1 + 0 = -2$

It cannot be less then -2 in any other case. This is the least possible.

So as long as the computed value > -2, we can return true.

So $t = -1.5$ will do.

Clause 2: $\neg A \vee C$

$w_A = -1$

$w_C = 1$

False when $A = \text{true},\ C = \text{False}$

Out $= -1 + 0 = -1$

So $t = -0.5$

Clause 3: $\neg A \vee \neg D$

$w_A = -1$

$w_D = -1$

False when $A =$ False, $D =$ False

out $= -2$

$t = -1.5$

Clause 4: $B \vee C$

$w_B = 1$

$w_C = 1$

False when $B =$ Fave, $C =$ False

out $= 0$

$t = 0.5$

Clause 5: $B \vee \neg D$

$w_B = 1$

$w_D = -1$

False when $B =$ False, $D =$ True

out $= -1$

$t = -0.5$

2^{ND} LAYER :

Assume input clauses as c_1, c_2, c_3, c_4, c_5

$$c_1 \wedge c_2 \wedge c_3 \wedge c_4 \wedge c_5$$

Assume $w_{c_1} = w_{c_2} = w_{c_3} = w_{c_4} = w_{c_5} = 1$

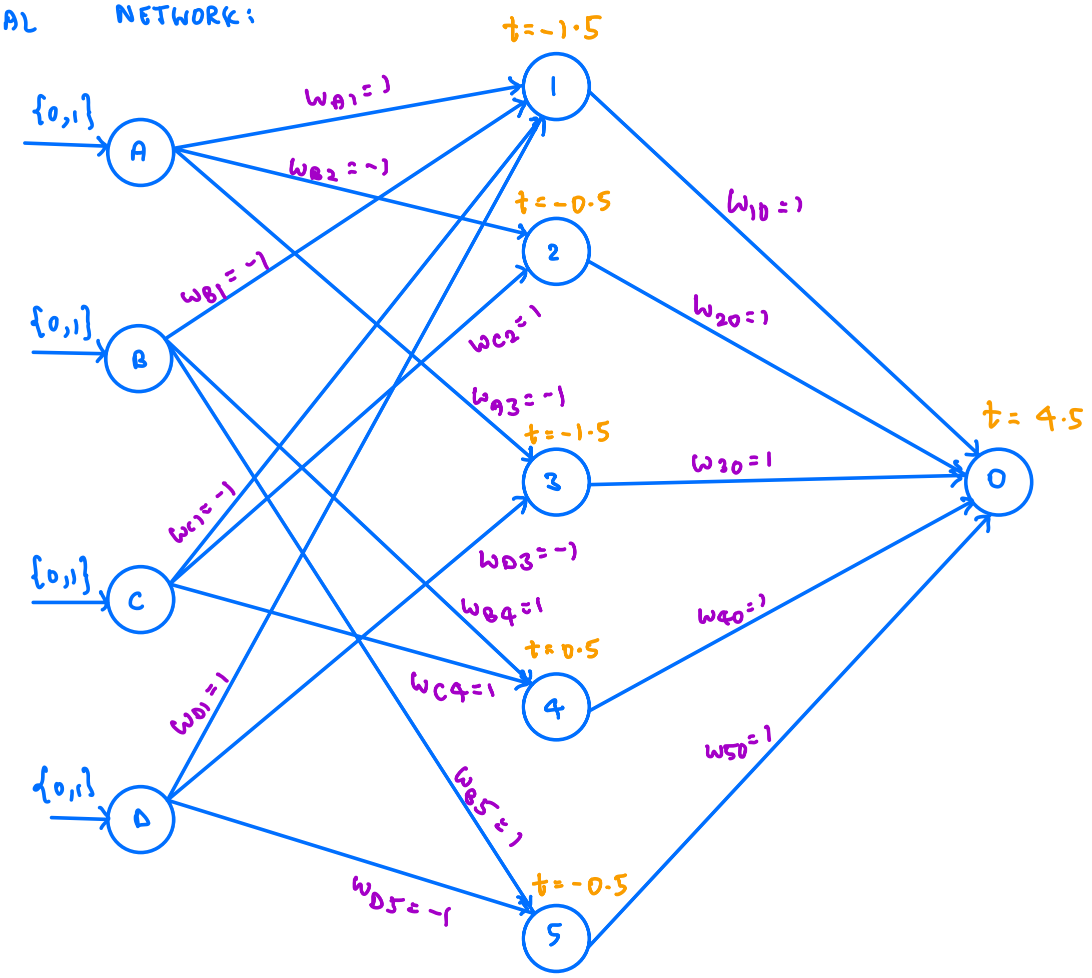
It is true only when

$$c_1 = c_2 = c_3 = c_4 = c_5 = \text{True}$$

$$\text{out} = 5$$

$$t = 4.5$$

NEURAL NETWORK:



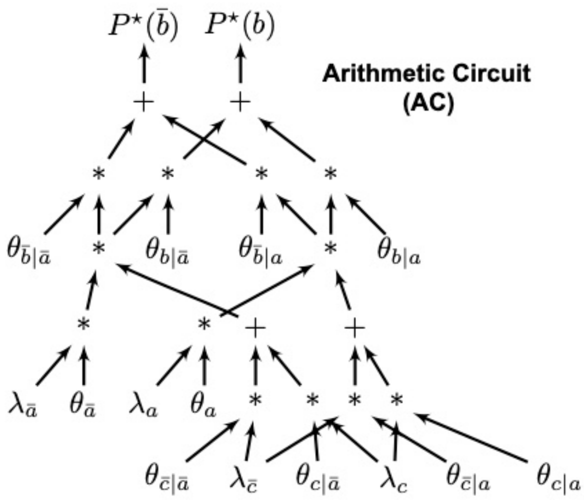
3. $\theta_a = 0.9$

$\theta_{b|a} = 0.2$

$\theta_{b|\bar{a}} = 0.5$

$\theta_{c|a} = 0.7$

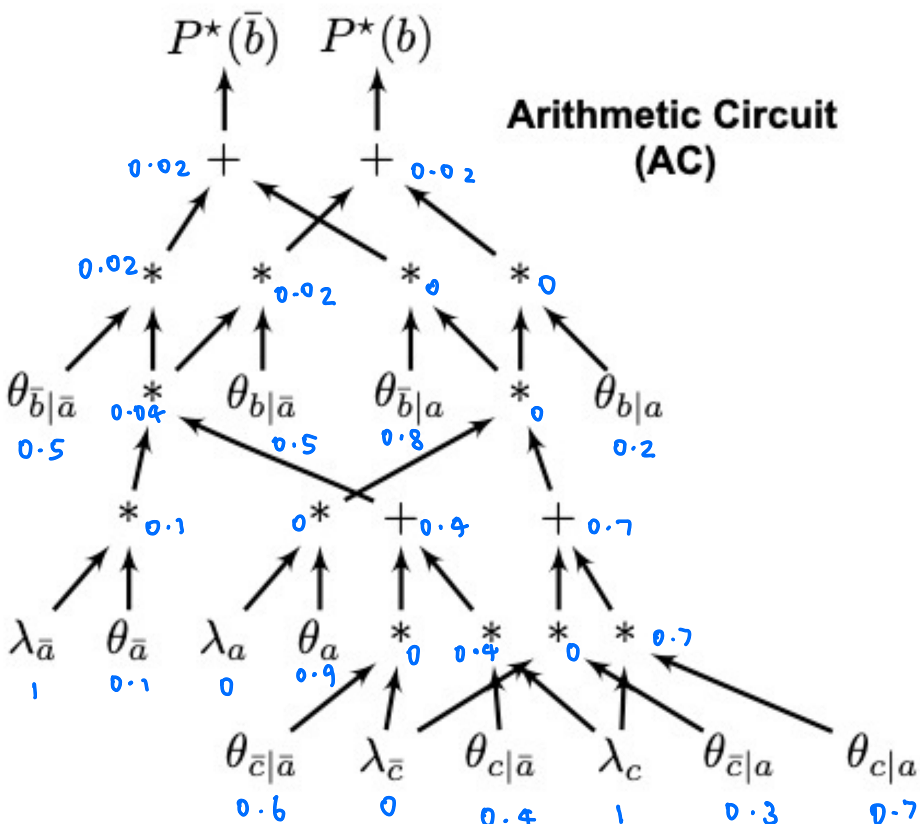
$\theta_{c|\bar{a}} = 0.4$



a. $e_1 = \bar{a}, c$

$\lambda_a = 0$ $\lambda_c = 1$

$\lambda_{\bar{a}} = 1$ $\lambda_{\bar{c}} = 0$

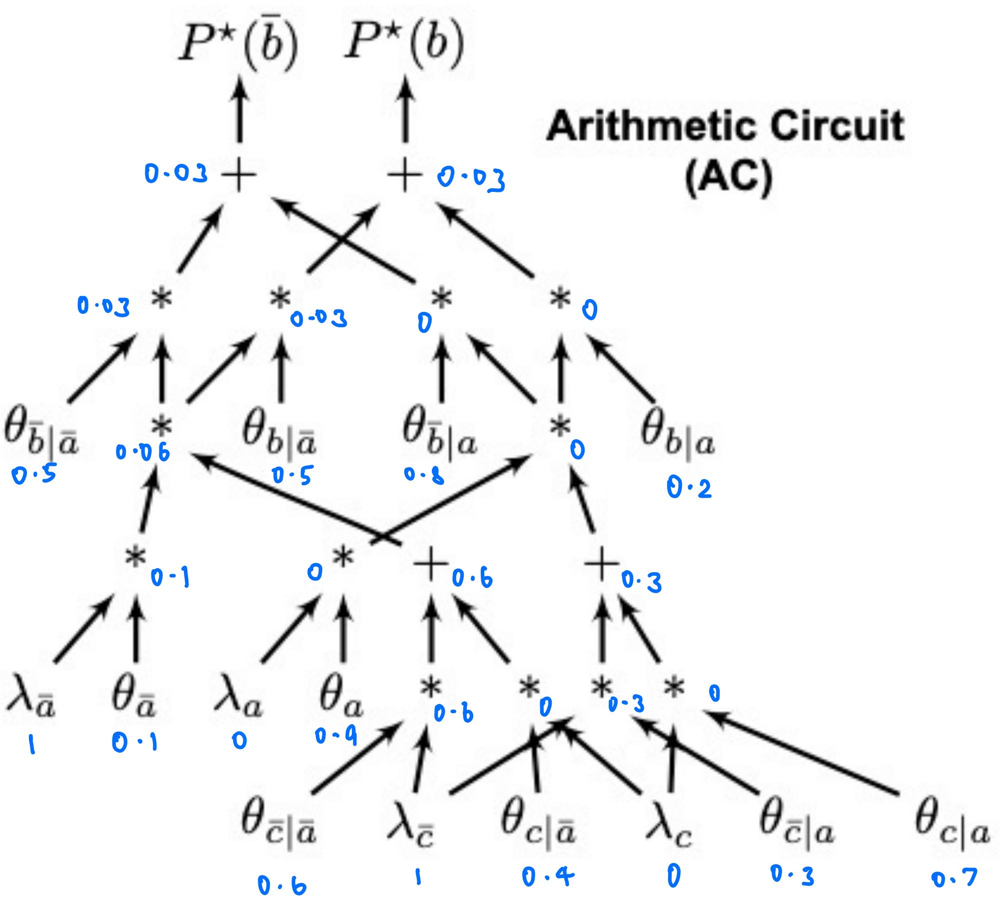


$p^*(\bar{b}) = p^*(b) = 0.02$

$e_2 = \bar{a}, \bar{c}$

$\lambda_a = 0 \qquad \lambda_c = 0$

$\lambda_{\bar{a}} = 1 \qquad \lambda_{\bar{c}} = 1$



$P^*(\bar{b}) = P^*(b) = 0.03$

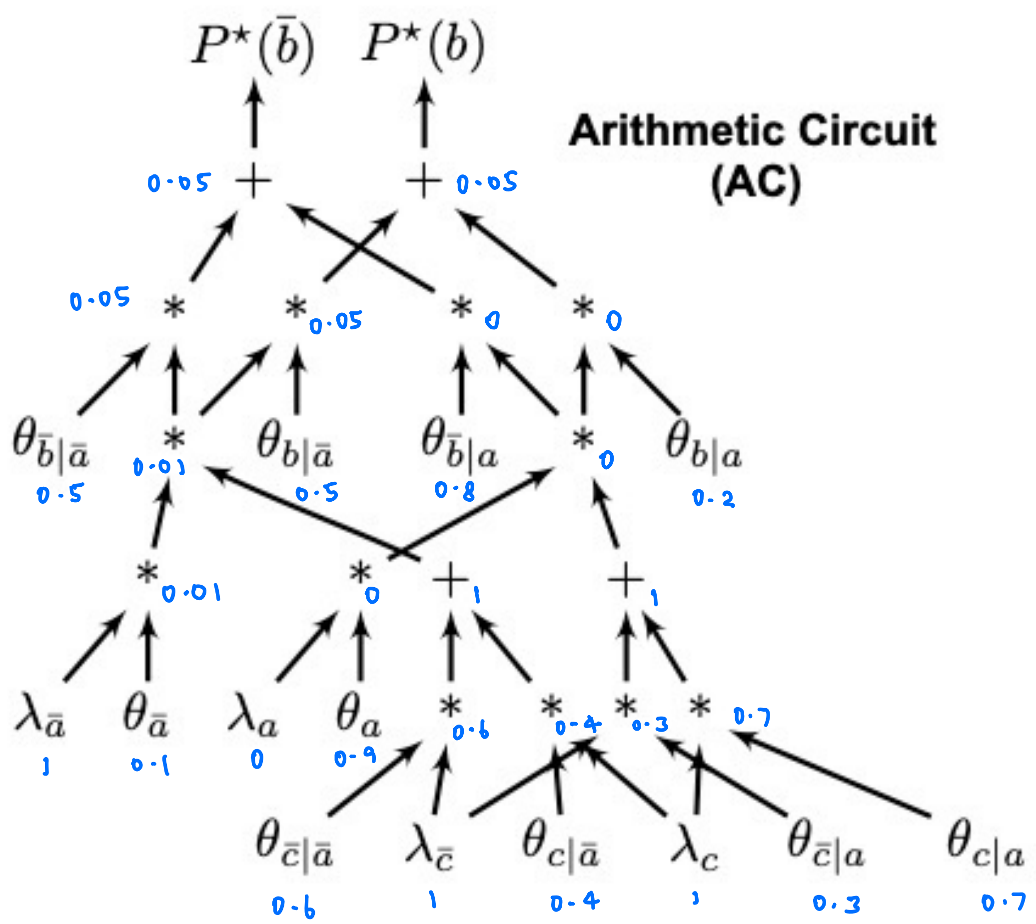
$P^*(\bar{b}) = \theta_{\bar{b}|\bar{a}} \cdot \theta_{\bar{a}} \cdot (\theta_{\bar{c}|\bar{a}})$

$P^*(b) = \theta_{\bar{a}} \cdot \theta_{b|\bar{a}} \cdot \theta_{\bar{c}|\bar{a}}$

$e_3 = \bar{a}$

$\lambda_a = 0 \qquad \lambda_c = 1$

$\lambda_{\bar{a}} = 1 \qquad \lambda_{\bar{c}} = 1$



$P^\star(\bar{b}) = P^\star(b) = 0.05$

$P^\star(\bar{b}) = \theta_{\bar{b}|\bar{a}} \cdot \theta_{\bar{a}} \cdot (\theta_{\bar{c}|\bar{a}} + \theta_{c|\bar{a}})$

$= \theta_{\bar{b},\bar{a}}$

$P^\star(b) = \theta_{b|\bar{a}} \cdot \theta_{\bar{a}} (\theta_{\bar{c}|\bar{a}} + \theta_{c|\bar{a}})$

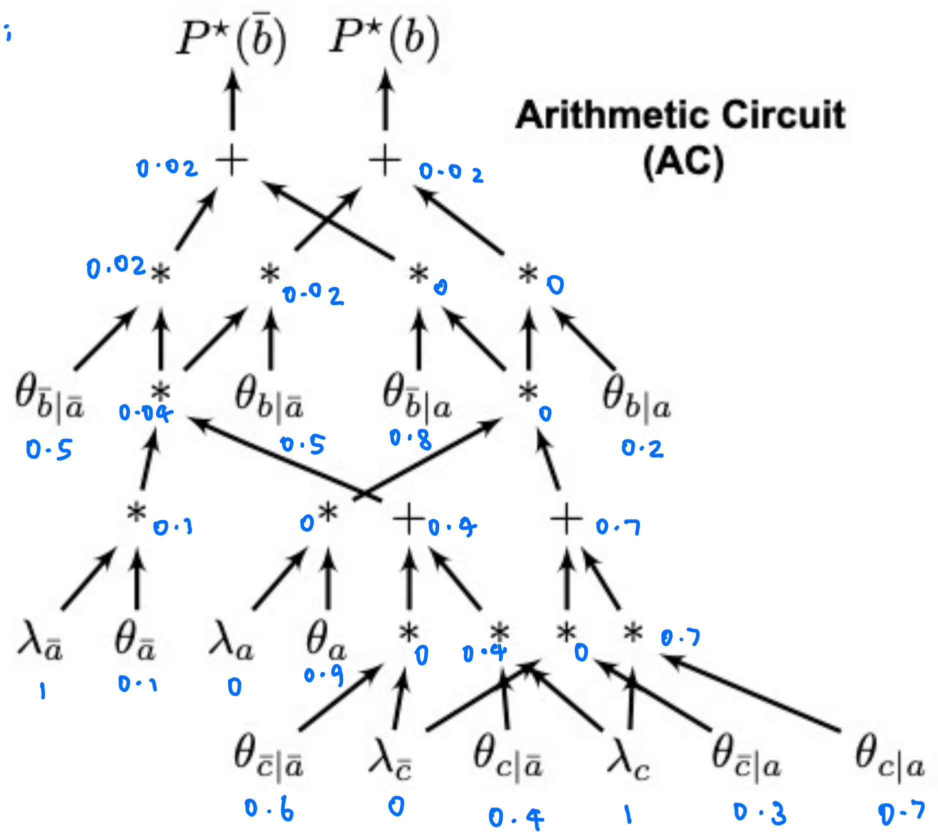
$= \theta_{b,\bar{a}}$

b. $P(b, \bar{a}, c) = P(\bar{b}, \bar{a}, c) = 0.02$

$P(b, \bar{a}, \bar{c}) = P(\bar{b}, \bar{a}, \bar{c}) = 0.03$

$P(b, \bar{a}) = P(\bar{b}, \bar{a}) = 0.05$

e_1:



$P^*(\bar{b}) = \theta_{\bar{b}|\bar{a}} \cdot \theta_{\bar{a}} \cdot \theta_{c|\bar{a}} = \theta_{\bar{a}} \cdot \theta_{\bar{b},c|\bar{a}}$

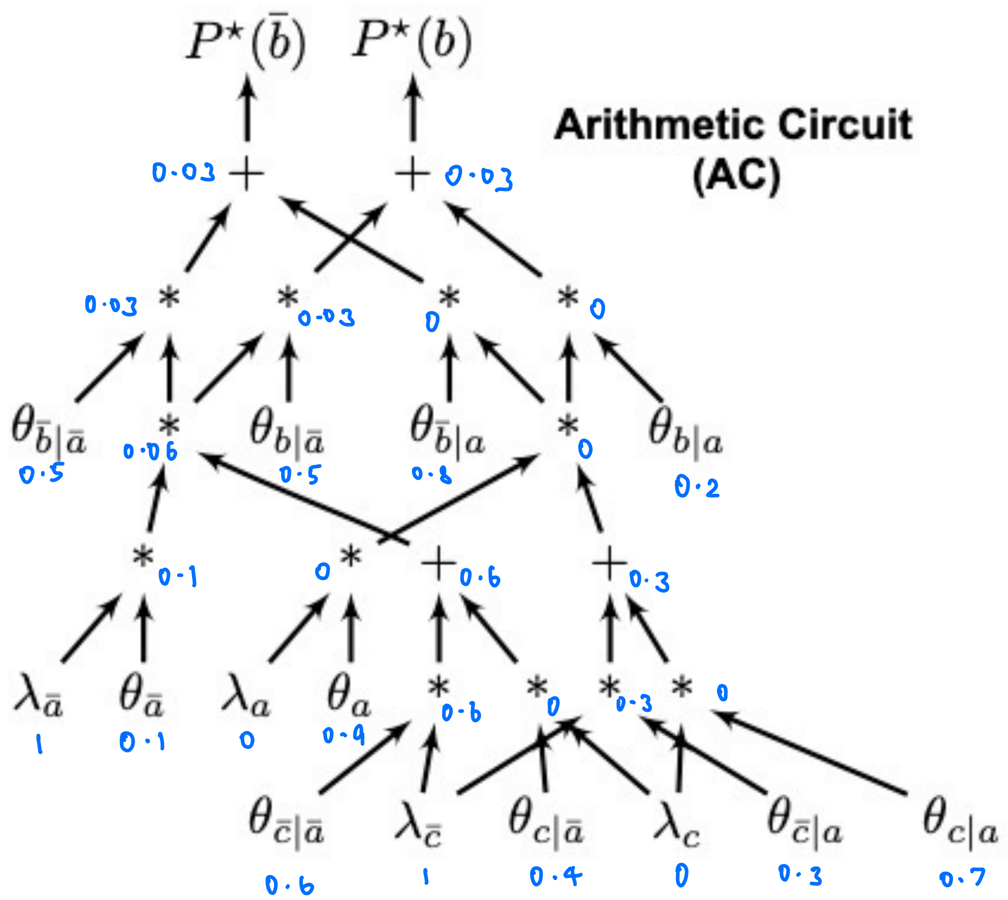
[b, c → conditional independent given a]

$= \theta_{\bar{b},\bar{a},c} = 0.02$

$P^*(b) = \theta_{b|\bar{a}} \cdot \theta_{\bar{a}} \cdot \theta_{c|\bar{a}}$

$= \theta_{b,\bar{a},c} = 0.02$

e_2 :



$$p^*(b) = \theta_{\bar{a}} \cdot \theta_{b|\bar{a}} \cdot \theta_{\bar{c}|\bar{a}}$$

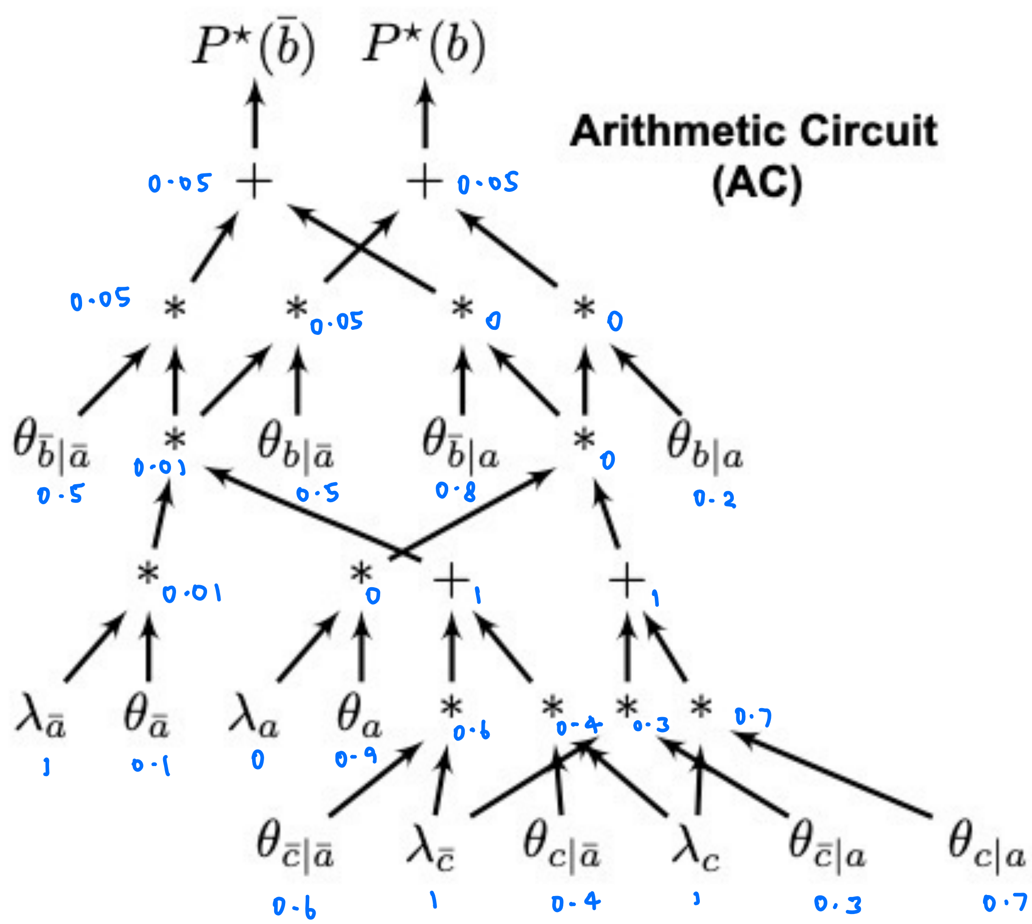
$$= \theta_{\bar{a}} \cdot \theta_{b,\bar{c}|\bar{a}}$$ [b, c → conditional independent given a]

$$= \theta_{b,\bar{c},\bar{a}} = 0.03$$

$$p^*(\bar{b}) = \theta_{\bar{b}|\bar{a}} \cdot \theta_{\bar{a}} \cdot (\theta_{\bar{c}|\bar{a}})$$

$$= \theta_{\bar{b},\bar{c},\bar{a}} = 0.03$$

e_3:



$$P^*(\bar{b}) = \theta_{\bar{b}|\bar{a}} \cdot \theta_{\bar{a}} \cdot (\theta_{\bar{c}|\bar{a}} + \theta_{c|\bar{a}})$$

$$= \theta_{\bar{b},\bar{a}} = 0.05$$

$$P^*(b) = \theta_{b|\bar{a}} \cdot \theta_{\bar{a}} (\theta_{\bar{c}|\bar{a}} + \theta_{c|\bar{a}})$$

$$= \theta_{b,\bar{a}} = 0.05$$

So $P^*(b)$ is $P(b,e)$ and $P^*(\bar{b})$ is $P(\bar{b},e)$ //

c. $Pr(\bar{b}|e_1) = Pr(\bar{b}|\bar{a}, c) = \theta_{\bar{b}|\bar{a},c} = \dfrac{\theta_{\bar{b},\bar{a},c}}{\theta_{\bar{a},c}}$

$$= \frac{0.02}{\theta_{c|\bar{a}} \cdot \theta_{\bar{a}}}$$

$$= \frac{0.02}{0.4 \times 0.1}$$

$$= \frac{0.02}{0.04}$$

$$= 0.5 \,//$$

$Pr(\bar{b}|e_2) = Pr(\bar{b}|\bar{a}, \bar{c}) = \theta_{\bar{b}|\bar{a},\bar{c}} = \dfrac{\theta_{\bar{b},\bar{a},\bar{c}}}{\theta_{\bar{a},\bar{c}}}$

$$= \frac{0.03}{\theta_{\bar{c}|\bar{a}} \cdot \theta_{\bar{a}}}$$

$$= \frac{0.03}{0.6 \times 0.1}$$

$$= 0.5 \,//$$

$Pr(\bar{b}|e_3) = Pr(\bar{b}|\bar{a}) = 1 - Pr(b|\bar{a}) = 0.5 \,//$