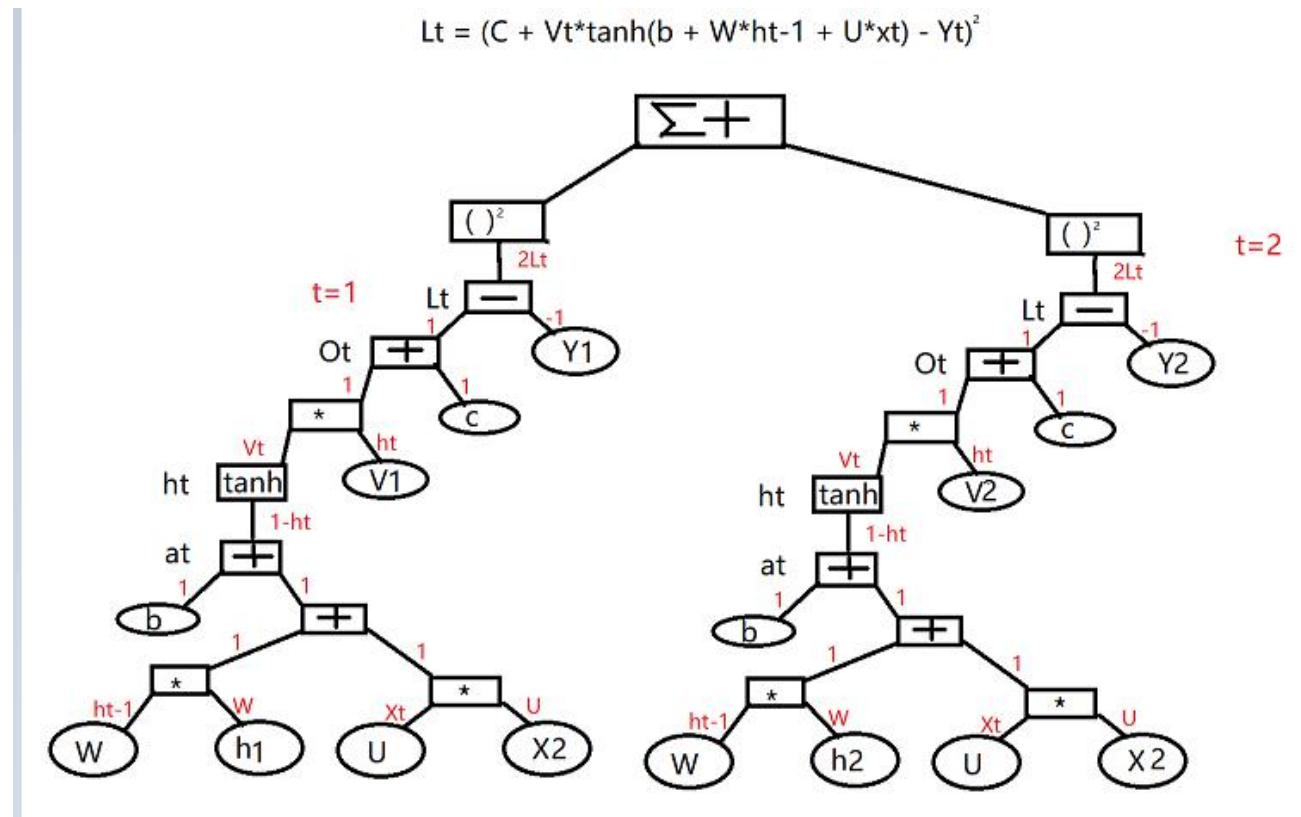


## Problem 1: Computational Graphs and Backpropagation

Q1. Construct the computational graph.



Q2. What are the model parameters to be learned?

The parameters of  $b$ ,  $W$ ,  $U$ ,  $c$  and  $V$  will be learned.

Q3.

See below

Q3													
Forward Pass	b=[0.1, 1]T		w=	0.5	-0.4		U=	0.8	-0.3	0.5			
				-0.3	0.4			-1	0.1	0.9			
	c=0.7		V=[0.5, 0.8]T										
	x1=[1, -1, 1]T	y1=1.5											
	x2=[1, -1, -1]T	y2=0.8											
	a1=	b+Who+Ux0	=	[0.1, 1]T	+		0.8	-0.3	0.5	*	[1, -1, 1]T	=	1.7
							-1	0.1	0.9				-0.2
	h1=	tanh(a1)	=	0.9345									
				-1974									
	O1=	c+Vth1	=	0.7	+	[0.5, 0.8]	*	0.9345	=	1.0096			
								-1974					
	L1=	(O1 - y1) <sup>2</sup>	=	(1.0098-1.5)	=	0.2403							
	as same as the above												
	a2=	b+Wh1+Ux2	=	1.2407									
				-2.4188									
	h2=	tanh(a2)	=	0.8473									
				-0.9842									
	O2=	c+Vth2	=	0.3383									
	L2=	(O2 - y2) <sup>2</sup>	=	0.21501769									
Backward Pass													
	$\partial TL / \partial c$	=	2(O1-y1)+2(O2-y2)	=	-1.9078								
	$\partial TL / \partial v$	=	2(O1-y1)h1+2(O2-y2)h2	=	-1.703								
					1.1064								
	$\partial TL / \partial b$	=	2(O1-y1)diag(1-h1 <sup>2</sup> )	+	-0.1921								
			2(O2-y2)diag(1-h2 <sup>2</sup> )	=	-0.7769								
	$\partial TL / \partial U$	=	2(O1-y1)diag(1-h1 <sup>2</sup> )vX1	+	-0.1921	0.1921	0.0695						
			2(O1-y1)diag(1-h1 <sup>2</sup> )vX2	=	-0.7769	0.7769	-0.7306						
	$\partial TL / \partial W$	=	2(O1-y1)diag(1-h1 <sup>2</sup> )vh0	+	-0.1224	0.0258							
			2(O1-y1)diag(1-h1 <sup>2</sup> )vh1	=	-0.0217	0.0046							

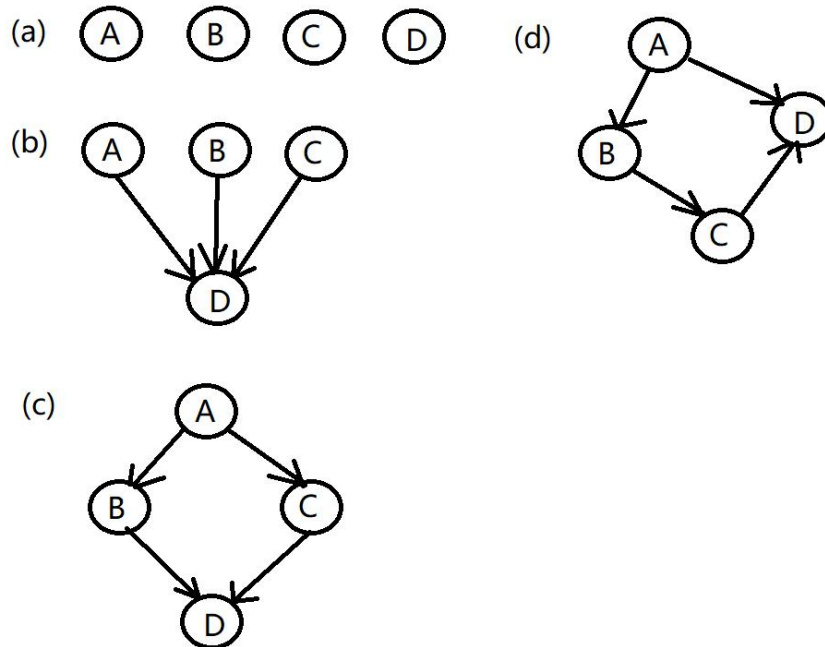
## Problem 2: Naïve Bayes Classifier

P(ripe=yes) = 8 / 17	0.470588235				
P(ripe=no) = 9 / 17	0.529411765				
P(color=green   ripe=yes) = 3 / 8	0.375				
P(color=green   ripe=no) = 3 / 9	0.333333333				
P(root=slightly curly   ripe=yes) = 3 / 8	0.375				
P(root=slightly curly   ripe=no) = 4 / 9	0.444444444				
P(texture=clear   ripe=yes) = 7 / 8	0.875				
P(texture=clear   ripe=no) = 2 / 9	0.222222222				
P(texture=hard   ripe=yes) = 6 / 8	0.75				
P(texture=hard   ripe=no) = 6 / 9	0.666666667				
color=green, root=slightly curly, texture=clear, and surface=hard					
P(ripe=yes) = 8 / 17	P(color=green   ripe=yes) = 3 / 8	P(root=slightly curly   ripe=yes) = 3 / 8	P(texture=clear   ripe=yes) = 7 / 8	P(texture=hard   ripe=yes) = 6 / 8	0.043428309
P(ripe=no) = 9 / 17	P(color=green   ripe=no) = 3 / 9	P(root=slightly curly   ripe=no) = 4 / 9	P(texture=clear   ripe=no) = 2 / 9	P(texture=hard   ripe=no) = 6 / 9	0.011619463
					the result ripe label will be yes

The prediction label will be Yes

### Problem 3: Bayesian Networks

Q1.



Q2.

```

graph TD
    A((A)) -- "P(A=1)=0.4" --> B((B))
    B -- "P(B=1|A=1)=0.8" --> C((C))
    B -- "P(B=1|A=0)=0.5" --> D((D))
    C -- "P(C=1|B=1)=0.2" --> D
    C -- "P(C=1|B=0)=0.3" --> D
    D -- "P(D=1|B=1)=0.7" --> E((E))
    D -- "P(D=1|B=0)=0.2" --> E
    E -- "P(E=1|D=1)=0.9" --> End1[ ]
    E -- "P(E=1|D=0)=0.2" --> End2[ ]
  
```

$P(B=1, A=1) =$	$P(B=1   D=1) =$	$P(D=1   B=1) =$	$P(B=1   A=1) =$	$P(A=1) =$	+
$P(B=1   D=0) =$	$P(D=0   B=1) =$	$P(B=1   A=1) =$	$P(A=1) =$	+	
$P(B=1   D=1) =$	$P(D=1   B=0) =$	$P(B=0   A=1) =$	$P(A=1) =$	+	
$P(B=1   D=0) =$	$P(D=0   B=0) =$	$P(B=0   A=1) =$	$P(A=1) =$	+	
$0.9 \times 0.7 \times 0.8 \times 0.4 =$	+				
$0.2 \times 0.3 \times 0.8 \times 0.4 =$	+				
$0.9 \times 0.2 \times 0.2 \times 0.4 =$	+				
$0.2 \times 0.8 \times 0.2 \times 0.4 =$					
	0.2016	0.0192	0.0144	0.0128	
	0.248				

$P(A=1)$  0.4  
 $P(A=0)$  0.6

$P(B=1 | A=1)$  0.8  
 $P(B=1 | A=0)$  0.5

$P(C=1 | B=1)$  0.2  
 $P(C=1 | B=0)$  0.3

$P(D=1 | B=1)$  0.7  
 $P(D=1 | B=0)$  0.2

$P(E=1 | D=1)$  0.9  
 $P(E=1 | D=0)$  0.2

$P(B=1)$  0.62  
 $P(B=0)$  0.38

$P(D=1)$  0.51  
 $P(D=0)$  0.49

$P(E=1)$  0.557  
 $P(E=0)$  0.443

		A			
		1	0		
B	1	0.32	0.3	$P(B=1)$	0.62
	0	0.08	0.3	$P(B=0)$	0.38

$P(B=1) = 0.4$   
 $P(A=0) = 0.6$   
 $P(B=0 | A=1) = \frac{P(B=0 | A=1)}{P(A=1)} = \frac{0.08}{0.4} = 0.2$

		B			
		1	0		
D	1	0.434	0.076	$P(D=1)$	0.51
	0	0.186	0.304	$P(D=0)$	0.49

$P(B=1) = 0.62$   
 $P(B=0) = 0.38$   
 $P(D=0 | B=1) = \frac{P(D=0 | B=1)}{P(B=1)} = \frac{0.186}{0.62} = 0.3$

$P(D=0 | B=0) = \frac{P(D=0 | B=0)}{P(B=0)} = \frac{0.304}{0.38} = 0.8$

		D			
		1	0		
E	1	0.459	0.098	$P(E=1)$	0.557
	0	0.051	0.392	$P(E=0)$	0.443

$P(D=1) = 0.51$   
 $P(D=0) = 0.49$   
 $P(E=0 | D=1) = \frac{P(E=0 | D=1)}{P(D=1)} = \frac{0.051}{0.51} = 0.1$

$$P(E=1, A=1) = 0.248$$

### Problem 4: Reinforcement Learning

The agent starts from Square 1 and makes the following actions:  
MoveSouth, MoveEast, MoveNorth, MoveWest.

reward=10				
step1: Initialize the Q table				
	east	west	north	south
1	0			0
2			0	0
3	0		0	
4			0	0
step2: perform update				
$Q(s,a) \leftarrow Q(s,a) + \eta \left( r + \left[ \gamma \max_{a'} Q(s',a') \right] - Q(s,a) \right)$				
Action 1 (MoveSouth): s=1, a=south, s1 = 3, r=10				
Q(1,south)	<-	0+0.5*(10+0-0)	=	5
	east	west	north	south
1	0			5
2			0	0
3	0		0	
4			0	0
Action 2 (MoveEast): s=3, a=east, s1 = 4, r=0				
Q(1,east)	<-	0+0.5*(0+0-0)	=	0
	east	west	north	south
1	0			5
2			0	0
3	0		0	
4			0	0
Action 3 (MoveNorth): s=4, a=north, s1 = 2, r=0				
Q(1,north)	<-	0+0.5*(0+0-0)	=	0
	east	west	north	south
1	0			5
2			0	0
3	0		0	
4			0	0
Action 4 (MoveWest): s=2, a=west, s1 = 1, r=5				
Q(1,west)	<-	0+0.5*(5+0-0)	=	2.25
	east	west	north	south
1	0			5
2		2.25		0
3	0		0	
4			0	0

The final state will be

	East	west	north	south
1	0			5
2	0	2.25		0
3	0			0
4		0	0	