Roxanne Farhad – raf2192 – Contributed With Carrie Hay Kellar

1a)

$$P(A,B|d_k) = \frac{P(A)P(B)P(d_k|d_{k-1})}{d_k} = \sum_{d_{0,\dots,}d_{k-1}} P(A)P(B)P(d_0|A,B)\prod_{i=1}^{k} P(d_k\,|d_{k-1})$$

b)
The probability would have two values: 0 or 1. If all values of d were positive then the probability would be 1.0, otherwise it would be 0.

c)

$$P(A,B\,|d_k+)\alpha P(A,B\,|d_k+) = \sum_{d_0} P(A)P(B)P(d_0|A,B)(d_0|d_k\,+)$$

A	B	d_0	$P(A)P(B)P(d_0\|A,B)(d_0\|d_k\,+)$
+a	+b	+d_0	0.5*0.5*1*1 = 0.25
+a	+b	-d_0	0.5*0.5*0*0 = 0
+a	-b	+d_0	0.5*0.5*0.5*1 = 0.125
+a	-b	-d_0	0.5*0.5*0*0 = 0
-a	+b	+d_0	0.5*0.5*0.5*1 = 0.125
-a	+b	$-d_0$	0.5*0.5*0.5*0 = 0
-a	-b	+d_0	0.5 *0.5 * 0 * 0 = 0
-a	-b	-d_0	0.5*0.5*1*0 = 0

A	B	P(A,B,+d_k)
+a	+b	0.25
-a	+b	0.125
+a	-b	0.125
-a	-b	0

Normalised:

A	B	P(A,B,+d_k)
+a	+b	0.5
-a	+b	0.25
+a	-b	0.25
-a	-b	0

2a)

There are no variables that are guaranteed to be independent of Car Start.

Distributer OK, Voltage at Plug, Car Starts, Spark Quality, Spark Adequate are all not guaranteed independence.

2b)

Battery age and Alternator System OK, and Charging System OK are no longer guaranteed independence of each other when we observe Battery Voltage.

If we observe Main Fuse OK, all variables gain a new guarantee of independence.

2c)

$$P(D\ |ST = BAD) = \frac{\Pr(D) * \Pr\ (ST = BAD|D)}{\Pr\ (ST = BAD)}$$

D	ST	P(D | ST = BAD)
T	F	(0.02 * 0.99)/0.0228 = 0.868
T	F	(0.3 * 0.01)/0.0228 = 0.132

2d)

$$P(A, CS\ |BV = DEAD) = \frac{P(BV = DEAD\ |A, CS)P(A)P(CS\ |A)}{P(BV = DEAD)}$$

A	CS	P(A, CS |BV = DEAD)
T	F	(0.2088*0.9997*0.995)/0.0232 = 0.88898
T	T	0.4596*0.0003*0/0.0232 = 0
F	T	(0.2088*0.9997*0.005)/0.0232 = 0.09859
F	F	0.4596*0.0003*1/0.0232 = 0.00592

3a)

Y	P(Y)
+y	0.5
-y	0.5

y	$\mathbf{x_1}$	**P(y | $\mathbf{x_1}$)**
+y	+1	(1+1)/7 =2/7
	0	(2+1)/7 =3/7
	-1	(1+1)/7 = 2/7
-y	+1	(2+1)/7=3/7
	0	(2+1)/7=3/7
	-1	(0+1)/7=1/7

y	$\mathbf{x_2}$	**P(y | $\mathbf{x_2}$)**
+y	+1	(2+1)/7 =3/7

	0	(2+1)/7 =3/7
	-1	(0+1)/7 = 1/7
-y	+1	(0+1)/7=1/7
	0	(2+1)/7=3/7
	-1	(2+1)/7=3/7

y	**x3**	**P(y | x3)**
+y	+1	(3+1)/7 =4/7
	0	(1+1)/7 =2/7
	-1	(0+1)/7 = 1/7
-y	+1	(2+1)/7=3/7
	0	(0+1)/7=1/7
	-1	(2+1)/7=3/7

The output for +y:

$$Log\big(P(+y)\big) + \sum_i Log\left(P\big(fi(y)\big)\right)$$
$$= Log(0.5) + Log\big(P(x1 = 1|y = +y)\big) + Log\big(P(x2 = -1|y = \ +y)\big)$$
$$+ Log\big(P(x3 = \ +1 \mid y = \ +y)\big)$$
$$= Log(0.5) + Log\left(\frac{2}{7}\right) + Log\left(\frac{1}{7}\right) + Log\left(\frac{4}{7}\right) = -1.933$$

The output for -y:

$$Log\big(P(+y)\big) + \sum_i Log\left(P\big(fi(y)\big)\right)$$
$$= Log(0.5) + Log\big(P(x1 = 1|y = -y)\big) + Log\big(P(x2 = -1|y = \ -y)\big)$$
$$+ Log\big(P(x3 = \ +1 \mid y = \ -y)\big)$$
$$= Log(0.5) + Log\left(\frac{1}{7}\right) + Log\left(\frac{3}{7}\right) + Log\left(\frac{3}{7}\right) = -1.882$$

Therefore, the likely output class is -y

3c)

The output class that maximises the probability of over-selling is: $X_1 = 0$, $X_2 = +1$, $X_3 = +1$

4a)

The values for the average acceptance rate and average weights are extremely similar. The average acceptance rate is based on consistency between the evidence and samples, and weights are calculated based on proportion of that sample among all samples. If more samples are rejected then lower is the acceptance rate. If a sample is very common, then it will have a lower weight of the individual sample.

4b)
The accuracy of Gibbs sampling depends on

4c)

Gibbs sampling has a quicker computation time than the other samples, but seems to have greater differences between the values of the other two samples.

4d)
Gibbs sampling compared more poorly, and with a longer computation time. This makes sense because if there is no evidence, or little evidence, provided then the algorithm will generate random samples to condition that node on.
4e)
As we observe more followers with the same value, the marginal probabilities converge to that value and the samples become more consistent.