1(a)

Assume that $X = Z - Y$, since $Y \subseteq Z$, we got $X \cap Y = \emptyset$, $Z = X \cup Y$

then $p(Z) = p(X \cup Y) = p(X) + p(Y) \geq p(Y)$

1(b)

$p(X \mid Z) = \frac{p(X \cap Z)}{p(Z)}$

Since $p(X \cap Z) \subseteq p(Z)$, then $p(x \cap Z) \leq p(Z)$, $\frac{p(X \cap Z)}{p(Z)} \leq 1$

then we got $p(X \mid Z) = \frac{p(X \cap Z)}{p(Z)} \leq 1$

On the other hand, Since $\emptyset \subseteq X \cap Z$, $0 = p(\emptyset) \leq p(X \cap Z)$;

Since $\emptyset \subseteq Z$, $0 = p(\emptyset) \leq p(Z)$

Therefore, $p(x \mid y) = \frac{p(X \cap Z)}{p(Z)} \geq 0$

To sum up, $p(X \mid Z)$ always fall in the range [0, 1]

1(c)

Since $E \cap \emptyset = \emptyset$, $p(\mathrm{E} \cup \emptyset) = p(\mathrm{E}) + p(\emptyset)$

$p(\mathrm{E}) = 1, p(\mathrm{E} \cup \emptyset) = p(E) = 1$

$p(\emptyset) = 0$

1(d)

Since $X \cap \bar{X} = \emptyset$,

We got $P(\bar{X} \cup X) = P(\bar{X}) + P(X)$

Then $P(X) = P(X \cup \bar{X}) - P(\bar{X})$

$= P(E) - P(\bar{X})$

$= 1 - P(\bar{X})$

1(e)

Assume $A = singing, B = rainy$

Then $p(singing\ AND\ rainy \mid rainy) = P(A \cap B \mid B)$

$= \frac{P((A \cap B) \cap B)}{P(B)}$

$= \frac{P(A \cap B \cap B)}{P(B)}$

$= \frac{P(A \cap B)}{P(B)}$

$= P(A \mid B)$

$= P(singing \mid rainy)$

1(f)

$$
\begin{aligned}
p(X \mid Y) + p(\bar{X} \mid Y) &= \frac{p(X \cap Y)}{p(Y)} + \frac{p(\bar{X} \cap Y)}{p(Y)} \\
&= \frac{p(X \cap Y) + p(\bar{X} + Y)}{p(Y)}
\end{aligned}
$$

Since

$$
(X \cap Y) \ \cap \ (\bar{X} \cap Y) = X \cap Y \cap X \cap \bar{Y} = \emptyset
$$

Then

$$
p((X \cap Y) \cup (\bar{X} \cap Y)) = p(X \cup Y) + p(\bar{X} \cap Y)
$$

Since

$$
\begin{aligned}
(X \cap Y) \ \cup \ (\bar{X} \cap Y) &= ((X \cap Y) \cup \bar{X}) \cap ((X \cap Y) \cup Y) \\
&= ((X \cup \bar{X}) \cap (Y \cup \bar{X})) \cap ((X \cup Y) \cap (Y \cup Y)) \\
&= (E \cap (Y \cup \bar{X})) \cap ((X \cup Y) \cap Y) \\
&= (Y \cup \bar{X}) \cap (Y \cup X) \\
&= (Y \cap (X \cup \bar{X})) \\
&= Y
\end{aligned}
$$

then the term above is equal to

$$
\begin{aligned}
p(X \mid Y) + p(\bar{X} \mid Y) &= \frac{p(X \cap Y) + p(\bar{X} + Y)}{p(Y)} \\
&= \frac{p((X \cap Y) \ \cup \ (\bar{X} \cap Y))}{p(Y)} \\
&= \frac{p(Y)}{p(Y)} = 1
\end{aligned}
$$

which means $p(X \mid Y) = 1 - p(\bar{X} \mid Y)$

1(g)

$$
\begin{aligned}
Original term &= (p(X \mid Y) \cdot p(Y) + p(X \mid \bar{Y}) \cdot p(\bar{Y})) \cdot p(\bar{Z}|X)/p(\bar{Z}) \\
&= (\frac{p(X \cap Y)}{p(Y)} \cdot p(Y) + \frac{p(X \cap \bar{Y})}{p(\bar{Y})} \cdot p(\bar{Y})) \cdot \frac{p(\bar{Z} \cap X)}{p(X) \cdot p(\bar{Z})} \\
&= (p(X \cap Y) + p(X \cap \bar{Y})) \cdot \frac{p(\bar{Z} \cap X)}{p(X) \cdot p(\bar{Z})}
\end{aligned}
$$

Since $(X \cap Y) \cap (X \cap \bar{Y}) = X \cap Y \cap X \cap \bar{Y} = \emptyset$

We get $p((X \cap Y) \cup (X \cap \bar{Y})) = p(X \cap Y) + p(X \cap \bar{Y})$

As proved above, $(X \cap Y) \cup (X \cap \bar{Y}) = X$

Then the term above

$$
\begin{aligned}
(p(X \cap Y) + p(X \cap \bar{Y})) \cdot \frac{p(\bar{Z} \cap X)}{p(X) \cdot p(\bar{Z})} &= p((X \cap Y) \cup (X \cap \bar{Y})) \cdot \frac{p(\bar{Z} \cap X)}{p(X) \cdot p(\bar{Z})} \\
&= p(X) \cdot \frac{p(\bar{Z} \cap X)}{p(X) \cdot p(\bar{Z})} \\
&= \frac{p(X \cap \bar{Z})}{p(\bar{Z})} = p(X \mid \bar{Z})
\end{aligned}
$$

1(h)

Under the condition that singing and rainy are mutually exclusive

It is because that $p(singing \cup rainy) = p(singing) + p(rainy) - p(singing \cap rainy)$

If you want $p(singing \cup rainy) = p(singing) + p(rainy)$, then $p(singing \cap rainy)$ must be 0.

Only when singing and rainy are mutually exclusive(which means $singing \cap rainy = \emptyset$), $p(singing \cap rainy) = 0$. Then $p(singing \cup rainy) = p(singing) + p(rainy)$

1(i)

Under the condition that singing and rainy are independant

It is because that $p(singing \cap rainy) = p(singing \mid rainy) \cdot p(rainy) = p(rainy \mid singing) \cdot p(singing)$

If you want $p(singing \cap rainy) = p(singing) \cdot p(rainy)$, then $p(singing) = p(singing \mid rainy)\ and\ p(rainy) = p(rainy \mid singsing)$,which means that singing and rainy are independent.

1(j)

Since $p(X \mid Y) = \frac{p(X \cap Y)}{p(Y)} = 0$

we get $p(X \cap Y) = 0$

Since $X \cap Y \cap Z \subseteq X \cap Y$, $p(X \cap Y \cap Z) \leq p(X \cap Y) = 0$, so $p(X \cap Y \cap Z) = 0$

Then $p(X \mid Y, Z) = \frac{p(X \cap Y \cap Z)}{p(Y \cap Z)} = 0$

1(k)

Since $p(W \mid Y) = \frac{p(W \cap Y)}{p(Y)} = 1$

we get $p(W \cap Y) = p(Y)$

Auusme $U = W \cap Y$, then $p(U \mid Y) = \frac{p(W \cap Y \cap Y)}{p(Y)} = p(W \mid Y) = 1$

Assume $V = Y - U$

Using 1(f), we have $p(V \mid Y) = 0$

Since $p(V \mid Y) = \frac{p(V \cap Y)}{p(Y)} = 0$,then $p(V \cap Y) = 0$

Owing to 1(a), $p(V \cap Y \cap Z) \leq p(V \cap Y) = 0$

Therefore $p(V \mid Y, Z) = \frac{p(V \cap Y \cap Z)}{p(y \cap Z)} = 0$

Using 1(f)$\Rightarrow$ $p(U \mid Y, Z) = 1$

Since $p(U \mid Y, Z) = \frac{p(U \cap U \cap Z)}{p(y \cap Z)} = \frac{p(W \cap Y \cap Y \cap Z)}{p(Y \cap Z)} = \frac{p(U \cap Y \cap Z)}{p(Y \cap Z)} = p(W \mid Y, Z)$

$p(W \mid Y, Z) = 1$

2(a)

Since

$$p(Actual = blue|Claimed = blue) = \frac{p(Actual = blur \cap Claimed = blue)}{p(Claimed = blue)}$$
$$p(Claimed = blue|Actual = blue) = \frac{p(Claimed = blue \cap Actual = blue)}{p(Actual = blue)}$$

We get:

$$\frac{p(Actual = bule \mid Claimed = blue)}{p(Claimed = blue \mid Actual = blue)} = \frac{p(Actual = blue)}{p(Claimed = blue)}$$

2(b)

prior probility is $p(Actual = blue)$

likelihood of the evidence is $p(Claimed = blue \mid Actual = blue)$

posterior probability is $p(Actual = blue \mid Claimed = blue)$

2(c)

The judge should care about the posterior probability, which is $p(Actual = blue \mid Claimed = blue)$

It is because that Bayes Theorem tells us

$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)}$

in which p(A | B) is the posterrior probablity, the final result of all calculation.

2(d)

$$\begin{aligned}\frac{p(B \mid A,Y)\cdot p(A \mid Y)}{p(B \mid Y)} &= \frac{p(B\cap A\cap Y)\cdot \frac{1}{p(A\cap Y)}\cdot p(A\cap Y)\cdot \frac{1}{p(Y)}}{\frac{p(B\cap Y)}{p(Y)}} \\ &= \frac{p(B\cap A\cap Y)}{p(B\cap Y)} \\ &= p(A \mid B,Y)\end{aligned}$$

2(e)

$$\begin{aligned}&p(B \mid A,Y)\cdot p(A \mid Y) + p(B \mid \bar{A},Y)\cdot p(\bar{A} \mid Y) \\ &= \frac{p(B\cap A\cap Y)}{p(A\cap Y)}\cdot\frac{p(A\cap Y)}{p(Y)} + \frac{p(B\cap \bar{A}\cap Y)}{p(\bar{A}\cap Y)}\cdot\frac{p(\bar{A}\cap Y)}{p(Y)} \\ &= \frac{p(B\cap A\cap Y)}{p(Y)} + \frac{p(B\cap \bar{A}\cap Y)}{p(Y)} \\ &= \frac{p(B\cap Y)}{p(Y)} \\ &= p(B \mid Y)\end{aligned}$$

So,

$$\begin{aligned}p(A \mid B,Y) &= \frac{p(B \mid A,Y)\cdot p(A \mid Y)}{p(B \mid Y)} \\ &= \frac{p(B \mid A,Y)\cdot p(A \mid Y)}{p(B \mid A,Y)\cdot p(A \mid Y) + p(B \mid \bar{A},Y)\cdot p(\bar{A} \mid Y)}\end{aligned}$$

2(f)

A = "Actual = Blue", $p(A) = 0.1$

B= "Claimed = Blue", $p(B \mid A) = 0.8$

Y = "city = Baltimore", p(Y) = 1

Then

$$p(A \mid B,Y) = \frac{0.8\times 0.1}{0.8\times 0.1 + (1-0.8)*(1-0.1)} = 0.3077$$

3(a)

$$\begin{aligned}\sum_{i=1}^{3} p(cry_i \mid situation = Predator!) = 1 \\ \sum_{i=1}^{3} p(cry_i \mid situation = Timber!) = 1 \\ \sum_{i=1}^{3} p(cry_i \mid situation = Ineedhelp!) = 1\end{aligned}$$

3(b)

p(cry, situation)	Predator!	Timber!	I need help!	Total
bwa	0	0	0.64	0.64
bwee	0	0	0.08	0.08
Kiki	0.2	0	0.08	0.28
Total	0.2	0	0.8	1

3(c)

i. $p(situation = Predator! \mid cry = kiki)$

i. $\frac{p(situation=Predator! \cap cry=kiki)}{p(cry=kiki)}$

iii. $\frac{0.2}{0.28} = 0.7143$

iv. $\frac{p(Predator!) \cdot p(kiki)}{p(Predator!) \cdot p(kiki) + p(Timber!) \cdot p(kiki) + p(Ineedhelp!) \cdot p(kiki))}$

v. $\frac{0.2}{0.2+0+0.08} = 0.7243$

4(a)

$$\begin{aligned} p(w_1 w_2 w_3 w_4) &= \prod_{i=1}^{5} p(w_i \mid w_{i-2}, w_{i-1}) \\ &= p(w_5 \mid w_4 w_3) \cdot p(w_4 \mid w_3 w_2) \cdot p(w_3 \mid w_2 w_1) \cdot p(w_2 \mid w_1 w_0) \cdot p(w_1 \mid w_0 w_{-}1) \\ &= \frac{c(w_3 w_4 w_5)}{c(w_3 w_4)} \cdot \frac{c(w_2 w_3 w_4)}{c(w_2 w_3)} \cdot \frac{c(w_1 w_2 w_3)}{c(w_1 w_2)} \cdot \frac{c(w_0 w_1 w_2)}{c(w_0 w_1)} \cdot \frac{c(w_{-1} w_0 w_1)}{c(w_{-1} w_0)} \\ &= \frac{c(soy\ lint\ EOS)}{c(soy\ lint)} \cdot \frac{c(loves\ soy\ lint)}{c(loves\ soy)} \cdot \frac{c(arya\ loves\ soy)}{c(arya\ loves)} \cdot \frac{BOS\ arya\ loves}{c(BOS\ arya)} \cdot \frac{c(BOS\ BOS\ arya)}{c(BOS\ BOS)} \end{aligned}$$

c(BOS BOS) is the count of times that all sentences that appear in the training corpus , since "BOS BOS" is the start of every sentence

C(BOS BOS arya) is the count of times of all sentences that starts with "arya" in the training corpus

c(soy lint EOS) means the count of times that sentence ends with "soy lint" in the training corpus

4(b)

Becuase <s> do you think the < /s> is not a complete sentence. It doesn't conform to the grammar. Any good language model will not generate a sentence ends with "the </s>". $p(</s>|\ think\ the) = \frac{c(think\ the\ </s>)}{c(think\ the)}$ (in which $c(think\ the\ </s>)$ is almost zero) is the parameter that responsible for making the probablity low.

4(c)

(1) : B

(2): A

(3): C

(1):B contains both "<s> "(the start sign) and "</s>"(the end sign), which presents that do you think in B is a complete sentence;

(2):A only contains "do you think", there is no start sign or end sign, shows that they are only 3 words in order;

(3) C contains only "<s>" , the start sign. It means that "do you think" serves as the start of a sentence

---

5. See "Log-Linear Model for Distinguishing Different Levels of Programmers " on piazza

---

6.

$$\begin{aligned}&p(\neg fortune, \neg race, \neg horse, \neg shoe \mid \neg nail) \\ &= p(\neg fortune \mid \neg race, \neg horse, \neg shoe, \neg nail) \cdot p(\neg race | \neg horse, \neg shoe, \neg nail) \cdot p(\neg horse \mid \neg shoe, \neg nail) \cdot p(\neg shoe \mid \neg nail)\end{aligned}$$

Using the axiom from 1(k), we have

$$\begin{gathered}p(\neg fortune \mid \neg race) = 1 \Rightarrow p(\neg fortune \mid \neg race, \neg horse, \neg shoe, \neg nail) = 1 \\ p(\neg race \mid \neg horse) = 1 \Rightarrow p(\neg race \mid \neg horse, \neg shoe, \neg nail) = 1 \\ p(\neg horse \mid \neg shoe) = 1 \Rightarrow p(\neg horse \mid \neg shoe, \neg nail) = 1\end{gathered}$$

So we get

$$\begin{aligned}&p(\neg fortune, \neg race, \neg horse, \neg shoe \mid \neg nail) \\ &= p(\neg fortune \mid \neg race, \neg horse, \neg shoe, \neg nail) \cdot p(\neg race | \neg horse, \neg shoe, \neg nail) \cdot p(\neg horse \mid \neg shoe, \neg nail) \cdot p(\neg shoe \mid \neg nail) \\ &= 1 \times 1 \times 1 \times 1 = 1\end{aligned}$$

Using the axiom from 1(a)

Since $(\neg fortune \cap \neg race \cap \neg horse \cap \neg shoe \cap \neg nail) \subseteq (\neg fortuen \cap \neg nail)$

we have $p(\neg fortune \cap \neg race \cap \neg horse \cap \neg shoe \cap \neg nail) \leq p(\neg fortune \cap \neg nail)$

Therefore,

$$\begin{aligned}p(\neg fortune \mid \neg nail) &= \frac{p(\neg fortune \cap \neg nail)}{p(\neg nail)}) \\ &\geq \frac{p(\neg fortune \cap \neg race \cap \neg horse \cap \neg shoe \cap \neg nail)}{p(\neg nail)} \\ &= p(\neg fortune, \neg race, \neg horse, \neg shoe \mid \neg nail) = 1\end{aligned}$$

Owing to a(b), $p(\neg fortune \mid \neg nail)$ falls in the range [0, 1],so $p(\neg fortune \mid \neg nail) \leq 1$

Therefore, $p(\neg fortune \mid \neg nail) = 1$