CS331 WRITING ASSIGNMENT #3

1.

a) $P(Toothache = false) = 0.576 + 0.144 + 0.008 + 0.072 = 0.8$
$P(Toothache = true) = 1 - P(Toothache = false) = 1 - 0.8 = 0.2$

b) $P(Cavity = true) = 0.008 + 0.072 + 0.012 + 0.108 = 0.2$
$P(Cavity = false) = 1 - P(Cavity = true) = 1 - 0.2 = 0.8$

c)

Toothache	Cavity	P(Toothache / Cavity)
F	F	$(0.576 + 0.144)/0.8 = 0.9$
F	T	$(0.008 + 0.072)/0.2 = 0.4$
T	F	$(0.064 + 0.016)/0.8 = 0.1$
T	T	$(0.012 + 0.108)/0.2 = 0.6$

2. $P(X,Y) = P(X)P(Y)$ by (1), but $P(X,Y) = P(X|Y)P(Y)$ by definition of conditional probability. Thus,
$P(X)P(Y) = P(X|Y)P(Y) \rightarrow (P(X)P(Y))/P(Y) = P(X|Y)$
$\rightarrow P(X) = P(X|Y) \rightarrow P(X|Y) = P(X)$. Thus, (1) and (2) are equivalent. By noting that X and Y are arbitrary designations in this case (since ~~s~~wapping the two doesn't change the result of (1) since multiplication is communitative), we have:
$P(Y,X) = P(Y)P(X) = P(Y|X)\,P(X) \rightarrow P(Y) = P(Y|X)$
$\rightarrow P(Y|X) = P(Y)$. Thus, they are all equivalent.

3.

a) If we know X, then successive flips of the coin are independent of each other because ~~they~~ the chance of heads (and tails) remains constant regardless of what happened before.

b) If we don't know X, then the flips are <u>not</u> independent because each flip will update our best estimate of X, thus altering the expected probability of the next flip.

4. Let d be the probability that I have the disease and let t be the probability that I tested positive for the disease. By Bayes' rule, the odds that I actually have the disease given a positive test is:

$$P(d|t) = \frac{P(t|d)P(d)}{P(t)} = \frac{(0.99)(1/10{,}000)}{P(t)}$$

We can ~~x~~ see that the probability of us having the disease is directly proportional to the prior probability of having the disease. Since the disease is more rare (1 in 10,000) than the accuracy of the test (1 in 100), this is good news for us.

To determine the actual probability, we need $P(t)$:

T	D	$P(T\|D)$	$P(T, D)$	
true	false	0.01	0.009999	since $P(T,D) = P(T\|D)P(D)$
true	true	0.99	0.000099	and we know $P(D)$

Thus, $P(t) = 0.009999 + 0.000099 = 0.010098$

and $P(d|t) = \dfrac{(0.99)(0.0001)}{(0.010098)} \approx \boxed{0.0098}$; meaning there is less than a 1% chance of having the disease with a positive test result.

5.

a) It is possible to calculate the most likely color of the taxi if we have an estimate for the probability that any given taxi is blue (equivalent for green, just swap x with $1-x$). We'll call this probability x. Let A be the appeared color of the taxi and T be the actual color.

A	T	$P(A\|T)$	$P(A, T)$
b	g	0.25	$0.25(1-x)$
b	b	0.75	$0.75x$

$P(T=b) = x$, $P(A=b) = 0.25(1-x) + 0.75x$

$= 0.25 - 0.25x + 0.75x$

$= 0.25 + 0.5x$

By Bayes' Rule:

$$P(T=b|A=b) = \frac{P(A=b|T=b)P(T=b)}{P(A=b)} = \frac{(0.75)(x)}{(0.25+0.5x)} = \frac{0.75x}{0.25+0.5x}$$

If we assume a 50/50 split of blue and green taxis:

$x = 0.5$, $P(T=b|A=b) = 0.75(0.5)/(0.25+0.5(0.5)) = 0.375/0.5 = 0.75$

In this case, the actual color of the taxi is more likely to be blue

b) Here, $x = 0.1$ since $1-x = 0.9$. Thus,

$$P(T=b|A=b) = \frac{0.75(0.1)}{0.25+0.5(0.1)} = \frac{0.075}{0.25+0.05} = 0.25,$$

meaning the actual color of the taxi is more likely to be green with $1-0.25 = 0.75$ probability.