AI homework 11

13.8

a. $P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$

b. $P(Cavity) =< 0.2, 0.8 >$

c.$P(Toothache|cavity) =< (0.108 + 0.012)/0.2, (0.072 + 0.008)/0.2 >=< 0.6, 0.4 >$

d.
$P(Cavity|toothacheVcatch) =< (0.108 + 0.012 + 0.072)/0.416, (0.016 + 0.064 + 0.144)/0.416 >=< 0.4615, 0.5384 >$

13.11

The correct message is received if either zero or one of the n+ 1 bits are corrupted.
Since corruption occurs independently with probability c, the probability that zero bits are corrupted is $(1 - \epsilon)^{n+1}$.There are $n + l$ mutually exclusive ways that exactly one bit can be corrupted,one for each bit in the message. Each has probability $\epsilon(1 - \epsilon)^n$, so the overall probability that exactly one bit is corrupted is $n\epsilon(1 - \epsilon)^n$. Thus,the probability that the correct message is received is $(1 - \epsilon)^{n+1} + n\epsilon(1 - \epsilon)^n$.
The maximum feasible value of n, therefore, is the largest n satisfying the inequality
$(1 - \epsilon)^{n+1} + n\epsilon(1 - \epsilon)^n > 1 - \delta$
Numerically solving this for c = 0.001,8 = 0.01, we find n= 147.

13.13

Let V be the statement that the patient has the virus,and A and B the statements that the medical tests A and B returned positive, respectively. The problem statement gives:

$P(V) = 0.01$

$P(A|V) = 0.95$

$P(A|\neg V) = 0.10$

$P(B|V)0.90$

$P(B|\neg V) = 0.05$

The test whose positive result is more indicative of the virus being present is the one whose posterior probability,$P(V|A)$ or $P(V|B)$ is largest. One can compute these probabilities directly from the information given, finding that $P(V|A) = 0.0876$ and $P(V|B) = 0.1538$, so B is more indicative.

13.22

a. The model consists of the prior probability $P(Category)$ and the conditional probabilities $P(Word|Category)$ .For each category c,$P(Category = c)$is estimated as the fraction of all documents that are of category c. Similarly,$P(Word_i = true|Category = c)$ is estimated as the fraction of documents of category e that contain word i.

b. Every evidence variable is observed, since we can tell if any given word appears in a given document or not.

c. The independence assumption is clearly violated in practice. For example, the word pair “artificial intelligence”occurs more frequently in any given document category than would be suggested by multiplying the probabilities of “artificial”and "intelligence".