J. Košecká CS 687 Homework 1, Due Feb 28th

1. On a day when an assignment is due (A=a), the newsgroup tends to be busy (B=b), and the computer lab tends to be full (C=c). Consider the following conditional probability tables for the domain, where A = a, ¬a, B = b, ¬b, C = c, ¬c.

$P(A)$

A	P
a	0.2
¬ a	0.8

$P(B|A)$

B	A	P
b	a	0.9
¬ b	a	0.1
b	¬ a	0.4
¬ b	¬ a	0.6

$P(C|A)$

C	A	P
c	a	0.7
¬ c	a	0.3
c	¬ a	0.5
¬ c	¬ a	0.5

Construct the joint distribution out of these conditional probabilities tables assuming B and C are independent given A. What is the marginal distribution P(B,C)? Justify your answer using the actual probabilities, not your intuitions. What is the posterior distribution over A given that $B = b, P(A|B = b)$? What is the posterior distribution over A given that C=c, $P(A|C = c)$? What about $P(A|B = b, C = c)$? Explain the pattern among these posteriors and why it holds.

2. Often we need to carry out reasoning over some pair of variables X, Y conditioned on the value of other variable E. Using the definitions of conditional probabilities, prove the conditionalized version of the product rule:

$$P(x, y|e) = P(x|y, e)P(y|e)$$

Prove the conditionalized version of Bayes' rule:

$$P(y|x, e) = P(x|y, e)P(y|e)/P(x|e)$$

State whether this is true or give a counterexample

$$P(a|b, c) = P(b|a, c) \text{ then } P(a|c) = P(b|c)$$

3. 13.8.

4. Naive Bayes. Given the following titles of movies and songs and using the bag of words representation determine whether the word is more likely to be from a title of the movie or of a song and estimate the probabilities below:

Table 1: default

movie	song
A perfect world	A perfect day
My perfect woman	Electric Storm
Pretty woman	Another rainy day

(a) $P(movie)$ and $P(song)$

(b) $P(\text{"}Perfect\text{"}|movie), P(\text{"}Perfect\text{"}|song), P(\text{"}Storm\text{"}|movie)$ and $P(\text{"}Storm\text{"}|song)$.

(c) Using Laplace Smoothing[1] with k=1 compute $P(movie|\text{"}Perfect\ Storm\text{"})$

(d) Compute $P(movie|\text{"}Perfect\ Storm\text{"})$ with no Laplacian Smoothing.

[1]Consider Laplace Smoothing where $P(w_i|c)$ is estimated as $\frac{n_i+k}{n+ck}$ where n_i is the number of occurrences of w_i in c and C is the total number of classes.