PMR lecture notes

February 27, 2014

Chapter 1

Prerequisites

1.1 Joint probability distribution

A joint probability distribution is the PD of two random variables. This must sum to 1 as the probability of some combination must be 1.

If there are only two random variables, this is known as a bivariate distribution. If there are more than two variables this is known as a multivariate distribution.

1.1.1 Conditioning

We can use conditioning as a means of reducing a joint probability distribution. However, the reduced version results in an normalised distribution which no longer sums to 1.

Chapter 2

Belief networks

A Bayes(or belief) network is a directed acyclic graph (DAG) whose nodes represent random variables $X_1 ... X_n$. For each node, X_i has a CPD $P(X_i Par_G(x_i))$ which denotes dependency on it's parent.

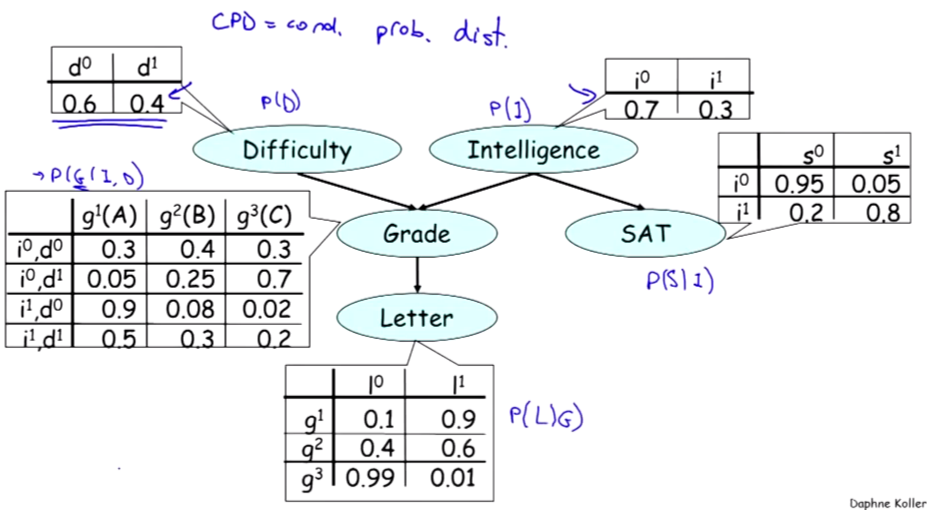

Figure 2.1: Bayes network with CPD details

This is an example of a bayesian network. The probabilities are dependent on each other as indicated by that arrows.

Definition: Chain rule

The chain rule takes all of the CPD's and multiplies them together

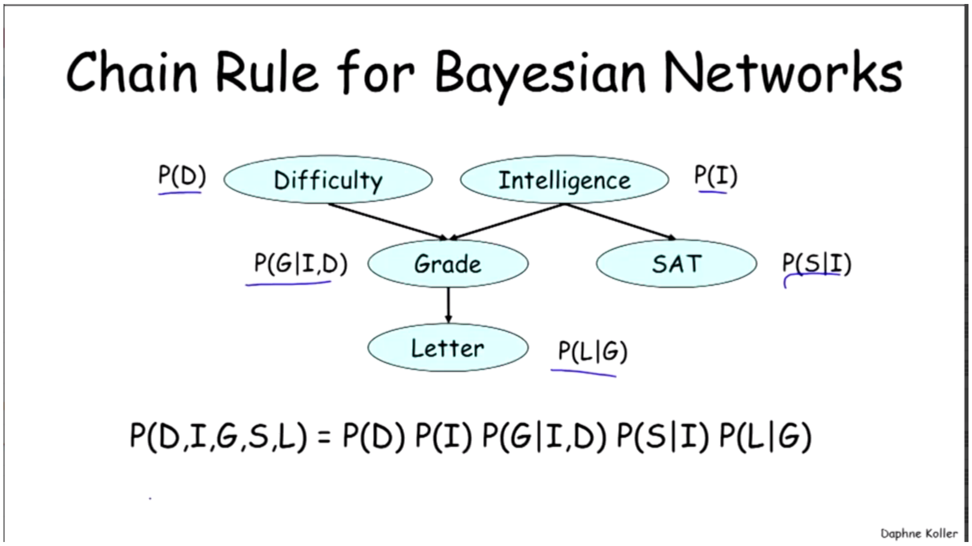

Figure 2.2: Bayes network with higher level details of CPDs

2.1 Questions

2.1.1 Question 1

Given the Bayesian network in **??**, what do you think is an appropriate factorization of the joint distribution P(D, I, G, S, L)?

Answer: P(D)P(I)P(G—I,D)P(S—I)P(L—G). This factorisation is also known as the chain rule.

Question 2

What is $P(d^0, i^1, g^3, s^1, l^1)$ in the Bayesian network **??**?

Using the two images for information:

$i^1 = 0.3$
$d^0 = 0.6$
$l^1 = 0.01$
$g^3 = 0.02$
$s^1 = 0.8$

And we can multiply these together to get the answer.

Chapter 3

Glossary

Definition: CPD

Conditional probability distribution

Definition: Conditioning

When we set one of the probabilities values (we condition it) as a means of reducing the network. See Coursera(Week 1, Distributions)