1 Bayes’ Nets Representation and Probability

Suppose that a patient can have a symptom ($S$) that can be caused by two different diseases ($A$ and $B$). It is known that the variation of gene $G$ plays a big role in the manifestation of disease $A$. The Bayes’ Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

(a) Compute the following entry from the joint distribution:

$$P(+g, +a, +b, +s) = P(+g)P(+a|+g)P(+b)P(+s|+b, +a) = (0.1)(1.0)(0.4)(1.0) = 0.04$$

(b) What is the probability that a patient has disease $A$?

$$P(+a) = P(+a|+g)P(+g) + P(+a|-g)P(-g) = (1.0)(0.1) + (0.1)(0.9) = 0.19$$

(c) What is the probability that a patient has disease $A$ given that they have disease $B$?

$$P(+a|+b) = P(+a) = 0.19$$ The first equality holds true since $A \perp \perp B$, which can be inferred from the graph of the Bayes’ net.

(d) What is the probability that a patient has disease $A$ given that they have symptom $S$ and disease $B$?

$$P(+a|+s, +b) = \frac{P(+a,+b,+s)}{P(+a,+b,+s) + P(-a,+b,+s)} = \frac{P(+a)P(+b)P(+s|+a,+b)}{P(+a)P(+b)P(+s|+a,+b) + P(-a)P(+b)P(+s|-a,+b)}$$

$$= \frac{(0.19)(0.4)(1.0)}{(0.19)(0.4)(1.0) + (0.81)(0.4)(0.8)} = \frac{0.076}{0.076 + 0.2592} \approx 0.2267$$

(e) What is the probability that a patient has the disease carrying gene variation $G$ given that they have disease $A$?

$$P(+g|+a) = \frac{P(+g)P(+a|+g)}{P(+g)P(+a|+g) + P(-g)P(+a|-g)} = \frac{(0.1)(1.0)}{(0.1)(1.0) + (0.9)(0.1)} = \frac{0.1}{0.109} = 0.5263$$
Q2. [Optional] Encrypted Knowledge Base

We have a propositional logic knowledge base as shown below, and we are trying to find a satisfying assignment for the variables $A, B, C, D,$ and $E$. Each line corresponds to a valid propositional logic sentence:

\[
\begin{align*}
\neg A \\
B &\Rightarrow A \\
D \\
C &\lor B \\
D &\lor E
\end{align*}
\]

(a) Your buddy Albert runs his solver, and hands you the model $M = \{A = \text{False}, B = \text{False}, C = \text{True}, D = \text{True}, E = \text{True}\}$ that causes all of the knowledge base sentences to be true. We have a query sentence $\alpha$ specified as $(A \lor C) \Rightarrow E$. Our model $M$ also causes $\alpha$ to be true. Can we say that the knowledge base entails $\alpha$? Explain briefly (in one sentence) why or why not.

No, the knowledge base does not entail $\alpha$. There are other models for which the knowledge base could be true and the query be false. Specifically $\{A = \text{False}, B = \text{False}, C = \text{True}, D = \text{True}, E = \text{False}\}$ satisfies the knowledge base but causes the query $\alpha$ to be false.

(b) Now we attempt to use theorem-proving methods to see whether our knowledge base entails a query sentence. To use these methods, it is useful to convert our knowledge base to conjunctive normal form (CNF), which satisfies:

- The sentence is a conjunction of (one or more) clauses.
- Each clause is a disjunction of literals.
- Each literal is a symbol or a negated symbol.

(i) Which sentences in the knowledge base are not already in conjunctive normal form? Convert them to CNF.

$B \Rightarrow A$ is converted to $\neg B \lor A$

(ii) Write the entire knowledge base as a single sentence in CNF.

After taking the conjunction of all sentences, we get:

\[
\neg A \land (\neg B \lor A) \land D \land (C \lor B) \land (D \lor E)
\]

(iii) Describe the steps necessary for converting $(A \land B) \lor (C \land D)$ to CNF.

We can distribute $\lor$ over $\land$ in the following way: $(\alpha \lor (\beta \land \gamma)) \equiv ((\alpha \lor \beta) \land (\alpha \lor \gamma))$. So from $(A \land B) \lor (C \land D)$, we can take $\alpha = (A \land B), \beta = C, \gamma = D$. This gives us:

\[
((A \land B) \lor C) \land ((A \land B) \lor D)
\]

Applying the same distributive property above a second time, we get:

\[
(((C \lor A) \land (C \lor B)) \land ((D \lor A) \land (D \lor B)))
\]

Since $\land$ is associative, we can rewrite:

\[
(C \lor A) \land (C \lor B) \land (D \lor A) \land (D \lor B)
\]

Finally by commutativity of $\lor$ we have:

\[
(A \lor C) \land (B \lor C) \land (A \lor D) \land (B \lor D)
\]

(You could have stopped at the previous step without applying commutativity, and that would have also been a perfectly valid CNF form.)