MPE proof

April 5, 2019

Theorem 0.1. *Given the definition of Most Probable Explanation (MPE):*

$$MPE(e) = \arg\max_{q} \mathbb{P}(Q = q | E = e) \tag{1}$$

the following algorithm constructs the branch corresponding to the MPE in a Bayesian Network composed of the set of nodes $V = \{Y_1 \dots Y_N\}$*:*

Result: Sequence of nodes corresponding to MPE
$E := Y_1$;
branch $= \emptyset$;
while $E \neq V$ **do**
  append $\arg\max_{Y_k} \mathbb{P}(Y_k|E)$ to branch;
  $E = E \cup \arg\max_{Y_k} \mathbb{P}(Y_k|E)$;
end

Proof. By backwards analysis, we can reconstruct the steps followed by the algorithm.

In the last step we are calculating the probability:

$$\max_{y_N} \mathbb{P}(Y_N | Y_{N-1} \dots Y_1) \tag{2}$$

To arrive in that state, in the previous step we calculated:

$$\max_{y_{N-1}} \mathbb{P}(Y_{N-1} | Y_{N-2} \dots Y_1) \tag{3}$$

and so on to arrive to the initial step:

$$\max_{y_2} \mathbb{P}(Y_2 | Y_1) \tag{4}$$

The total probability of the branch is:

$$
\begin{aligned}
&\max_{y_N} \mathbb{P}(Y_N|Y_{N-1}\dots Y_1) \\
&\quad \times \max_{y_{N-1}} \mathbb{P}(Y_{N-1}|Y_{N-2}\dots Y_1) \\
&\quad \times \dots \\
&\quad \times \max_{y_2} \mathbb{P}(Y_2|Y_1) \\
&= \max_{y_N} \frac{\mathbb{P}(Y_N, Y_{N-1}, \dots, Y_1)}{\mathbb{P}(Y_{N-1}, \dots, Y_1)} \\
&\quad \times \max_{y_{N-1}} \frac{\mathbb{P}(Y_{N-1}, \dots, Y_1)}{\mathbb{P}(Y_{N-2}, \dots, Y_1)} \\
&\quad \times \dots \\
&\quad \times \max_{y_2} \frac{\mathbb{P}(Y_2, Y_1)}{\mathbb{P}(Y_1)}
\end{aligned} \tag{5}
$$

And by a simple cyclic reordering of the denominators where each factor i receives the denominator of term $i+1$:

$$
\begin{aligned}
&= \max_{y_N} \frac{\mathbb{P}(Y_N, Y_{N-1}, \dots, Y_1)}{\mathbb{P}(Y_1)} \\
&\quad \times \max_{y_{N-1}} \frac{\mathbb{P}(Y_{N-1}, \dots, Y_1)}{\mathbb{P}(Y_{N-1}, \dots, Y_1)} \\
&\quad \times \dots \\
&\quad \times \max_{y_2} \frac{\mathbb{P}(Y_2, Y_1)}{\mathbb{P}(Y_2, Y_1)} \\
&= \max_{y_N} \frac{\mathbb{P}(Y_N, Y_{N-1}, \dots, Y_1)}{\mathbb{P}(Y_1)} \\
&\quad \times 1 \\
&\quad \times \dots \\
&\quad \times 1 \\
&= \max_{y_N} \frac{\mathbb{P}(Y_N, Y_{N-1}, \dots, Y_1)}{\mathbb{P}(Y_1)} \\
&= \max_{y_N} \mathbb{P}(Y_N, \dots, Y_{N-1}|Y_1)
\end{aligned} \tag{6}
$$

which is exactly the probability of the MPE we were looking for.

□