A NOTE ON THE DECISION THEORY SLIDES

During the lecture, I said there must be an error on slide 12 of the Decision Theory slides, which would carry over to slide 13. These are the second and third of the three slides headed "Inference". There is actually no error, just difficult-to-read notation that confused me during the lecture.

The passages in question are:

- Note that for any $\omega_i \in \Omega$, we can write $P[\omega_i] = P[\omega_i \mid X = x] \cdot P[\{\omega_j \mid X(\omega_j) = x\}]$. This follows directly from the definition of conditional probability.
-

$$\begin{aligned}
E_P[\phi(\delta_0, \delta_1)] &= \sum_{i=1}^{n} \phi\big(\omega_i, \delta_0, \delta_1(X(\omega_i))\big) P[\omega_i] \\
&= \sum_{x \in S} \sum_{\{\omega_i \mid X(\omega_i) = x\}} \phi\big(\omega_i, \delta_0, \delta_1(x)\big) P[\omega_i \mid X = x] P[\{\omega_j \mid X(\omega_j) = x\}] \\
&= E_P\big[E_P[\phi(\delta_0, \delta_1) \mid X = x]\big] .
\end{aligned}$$

- This is a special case of the **law of iterated expectations**. If there are no constraints that link our choice of $\delta_1(x)$ to our choice of $\delta_1(y)$ for $x \neq y$ or to our choice of δ_0, then we can choose δ_1 separately for each possible value of $X(\omega)$. And it is clear from the formula that to minimize expected loss overall, we should always choose $\delta_1(x)$ to minimize $E[\phi(\delta_0, \delta_1(x)) \mid X = x]$.

The notation that confused me is in red above. It is perhaps clearer if we write it in two lines:

$$\text{Let } A = \{\omega_j \mid X(\omega_j) = x\} .$$
$$P[\omega_i] = P[\omega_i \mid A] \cdot P[A] .$$

The first line, in words, says A is the set of ω values such that $X(\omega_j)$ takes on the value x. The second line is then indeed a clear consequence of the definition of conditional probability, and what appears on the following slide is therefore OK as is.