II-Probabilities
Probabilistic Reasoning

Logik für Erklärbare KI: Technische Einführung in das ENEXA Projekt

Maria Schnödt-Fuchs, Alex Goeßmann

Funded by the
European Union

15.+16. July, 2024

Example: Being at a dentist

We are reasoning about a factored system with three variables:

- Toothache $i \in [2]$, whether your tooth hurts
- Cavity $j \in [2]$, whether there is a cavity in your tooth
- Catch $k \in [2]$, whether the dentist catches in your tooth

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Formal Definition of Probability Tensors

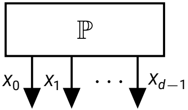
Definition (Probability Tensor)

Let there be a factored system $\mathcal{F}$ defined by variables X_k taking values in $[m_k]$. A probability distribution over the states of $\mathcal{F}$ is a map

$$\mathbb{P} : \bigtimes_{k\in[d]} [m_k] \to [0, \infty)$$

such that $\sum_{i_0,\dots,i_{d-1}} \mathbb{P}(i_0, \dots, i_{d-1}) = 1$.

We depict



Directed Tensors

We depict the condition, that coordinate sums are one, by directions on the legs.

Definition (Directed Tensor)

A tensor

$$T \in \bigotimes_{v \in \mathcal{V}} \mathbb{R}^{m_v}$$

is said to be directed with incoming variables $\mathcal{V}^{\mathrm{in}}$ and outgoing variables $\mathcal{V}^{\mathrm{out}}$, where $\mathcal{V} = \mathcal{V}^{\mathrm{in}} \dot{\cup} \mathcal{V}^{\mathrm{out}}$, when

$$\mathcal{C}\left(\{T\}, \mathcal{V}^{\mathrm{in}}\right) = \mathbb{I}^{\mathcal{V}^{\mathrm{in}}}$$

where $\mathbb{I}^{\mathcal{V}^{\mathrm{in}}}$ denoted the trivial tensor in $\bigotimes_{v \in \mathcal{V}^{\mathrm{in}}} \mathbb{R}^{m_v}$ which coordinates are all 1.

Example: Marginal Distributions

What is the probability that there is a cavity?

$$\mathbb{P}^{\text{Dentist,Cavity}}[\text{Cavity}] = \sum_{j\in[2]} \sum_{k\in[2]} \mathbb{P}^{\text{Dentist}}(:, j, k) \tag{1}$$

This is called a marginal distribution.

Exercise

Calculate the marginal probability of Cavity given the probability tensor $\mathbb{P}^{\text{Dentist}}$.

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Formal Definition of Marginal Distributions

Definition (Marginal Probability)

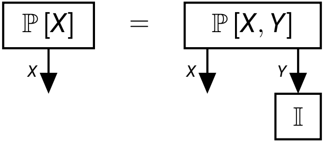
Given a distribution $\mathbb{P}$ of the categorical variables X and Y the marginal distribution of the categorical variable X is defined for each i_X as

$$\mathbb{P}\left[X = i_X\right] = \sum_{i_Y \in [m_Y]} \mathbb{P}\left[X = i_X, Y = i_Y\right] \,.$$

Marginal probabilities are contractions

$$\mathbb{P}\left[X\right] = \mathcal{C}\left(\{\mathbb{P}\}, \{X\}\right)$$

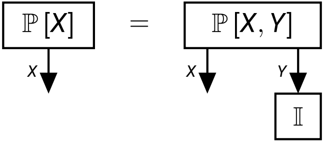
depicted by



Marginal Distributions in Contraction Formalism

Contractions are useful to

- Specify the Probability Tensor, or a decomposition of it
- Specify the variables to marginalize over as the ones left open



Directed notation preserved: Marginal probabilities are again probability distributions, since

$$\sum_{i\in[2]} \mathbb{P}^{\mathrm{Dentist},\mathrm{Cavity}}[i] = \sum_{i\in[2]}\sum_{j\in[2]}\sum_{k\in[2]} \mathbb{P}^{\mathrm{Dentist}}_{i,j,k} = 1\,. \tag{2}$$

Example: Conditional Distributions

What is the probability of having a cavity when having a toothache?

$$\mathbb{P}^{\text{Dentist}}[\text{Cavity}|\text{Toothache}] = \frac{\sum_{j\in[2]} \mathbb{P}^{\text{Dentist}}_{:,j,:}}{\sum_{j,k\in[2]} \mathbb{P}^{\text{Dentist}}_{:,j,k}} \quad (3)$$

This is called a conditional distribution.

Exercise

Calculate the probability of Cavity conditioned on Toothache.

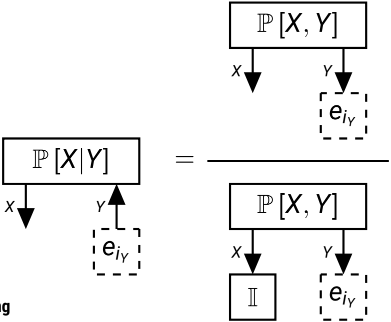
	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Formal Definition of Conditional Distributions

Definition (Conditional Probability)

Given a distribution $\mathbb{P}$ of the categorical variables X and Y, the conditioned distribution of X is defined by

$$\mathbb{P}\left[X = i_X | Y = i_Y\right] = \frac{\mathbb{P}\left[X = i_X, Y = i_Y\right]}{\mathbb{P}\left[Y = i_Y\right]} \,.$$



Definition (Normation of Tensor Networks)

A tensor network $\mathcal{T}$ on variables $\mathcal{V}$ can be normed on $\tilde{\mathcal{V}}$, if the coordinates of no slice with respect to $\tilde{\mathcal{V}}$ sum to 0. Then we define the normed tensor

$$\mathcal{N}\left(\mathcal{T}, \mathcal{V}^{\text{out}}, \mathcal{V}^{\text{in}}\right) \in \left(\bigotimes_{v \in \mathcal{V}^{\text{in}}} \mathbb{R}^{m_v}\right) \otimes \left(\bigotimes_{v \in \mathcal{V}^{\text{out}}} \mathbb{R}^{m_v}\right)$$

by

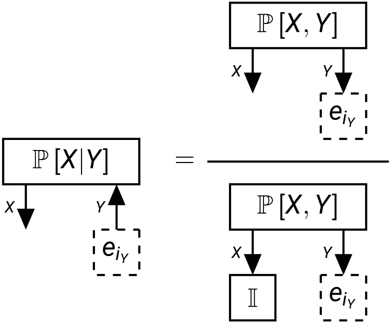
$$\mathcal{N}\left(\mathcal{T}, \mathcal{V}^{\text{out}}, \mathcal{V}^{\text{in}}\right) = \sum_{i_{\mathcal{V}^{\text{in}}} \in \times_{v \in \mathcal{V}^{\text{in}}} m_v} e_{i_{\mathcal{V}^{\text{in}}}} \otimes \frac{\mathcal{C}\left(\mathcal{T} \cup \{e_{i_{\mathcal{V}^{\text{in}}}}\}, \mathcal{V}^{\text{out}}\right)}{\mathcal{C}\left(\mathcal{T} \cup \{e_{i_{\mathcal{V}^{\text{in}}}}\}, \varnothing\right)} .$$

Conditional Distributions by Normations

Conditioning is the normation

$$\mathbb{P}[X|Y] = \mathcal{N}(\{\mathbb{P}\}, \{X\}, \{Y\})$$

depicted by



The directed notation highlights Conditions by incoming legs and Distributions by outgoing legs.

The Bayes Theorem

The Bayes Theorem relates conditional probabilities:

Theorem (Bayes Theorem)

For any joint distribution of two categorical variables X and Y it holds that

$$\mathbb{P}[X|Y] = \frac{\mathbb{P}[X,Y]}{\mathbb{P}[Y]} = \frac{\mathbb{P}[Y|X]\,\mathbb{P}[X]}{\mathbb{P}[Y]}\,.$$

Bayes Theorem in the Dentist Example

Directions of Reasoning

- Causal direction: Toothache is caused by cavity
- Diagnostic direction: Cavity is probable because of toothache

Bayes Theorem allows us to reason in diagnostic direction, given an underlying causal influence:

$$\begin{aligned}
&\mathbb{P}^{\text{Dentist}}[\text{Cavity}|\text{Toothache}] \\
&\qquad = \mathbb{P}^{\text{Dentist}}[\text{Toothache}|\text{Cavity}] \frac{\mathbb{P}^{\text{Dentist}}[\text{Cavity}]}{\mathbb{P}^{\text{Dentist}}[\text{Toothache}]}
\end{aligned}$$

Contraction Calculus for Probability Tensors

Probabilistic queries can be answered by contractions

- Marginal probabilities

$$\mathbb{P}[X] = \mathcal{C}(\{\mathbb{P}\}, \{X\})$$

- Conditional probabilities

$$\mathbb{P}[X|Y] = \mathcal{N}(\{\mathbb{P}\}, \{X\}, \{Y\})$$

Outlook: Tensor network decompositions of $\mathbb{P}$ increase the execution efficiency!