CS262a Bayesian Networks

© Tingfeng Xia @ UCLA

Winter Quarter, 2022

Preface

I shall organize everything based on the chapters later.

0.1 Principle Logical Forms

Inconsistent. Something that never holds; $\text{Mods}(\cdot) =$; $\Pr(\alpha) = 0$

Valid. Somethng that always holds; $\text{Mods}(\cdot) = \Omega$; $\Pr(\alpha) = 1$

Equivalent. $\text{Mods}(\alpha) = \text{Mods}(\beta)$

Mutual Exclusive. $\text{Mods}(\alpha) \cap \text{Mods}(\beta) = \emptyset$

Exhaustive. $\text{Mods}(\alpha) \cup \text{Mods}(\beta) = \Omega$

Entailment / Implication. $\alpha \vDash \beta \triangleq \text{Mods}(\alpha) \subseteq \text{Mods}(\beta)$

0.2 Equivalent Forms

- $\text{Mods}(\alpha \wedge \beta) = \text{Mods}(\alpha) \cap \text{Mods}(\beta)$
- $\text{Mods}(\alpha \vee \beta) = \text{Mods}(\alpha) \cup \text{Mods}(\beta)$
- $\text{Mods}(\neg\alpha) = \overline{\text{Mods}(\alpha)}$

0.3 Instantiation Agreement

Two instantiation, each of which can cover a subset of different varaibles, are said to be compatible with each other if they argree on all common variables. Denoted as $\mathbf{x} \sim \mathbf{y}$.

0.4 Information Theory

Entropy.

$$\mathrm{ENT}(X) = -\sum_{x} \Pr(x) \log \Pr(x) \tag{0.1}$$

where $0 \log 0 = 0$ by convention. With a higher entropy, we say that it is more chaotic.

Conditional Entropy.

$$\mathrm{ENT}(X|Y) = \sum_{y} \Pr(y)\mathrm{ENT}(X|y) \quad \text{where} \quad \mathrm{ENT}(X|y) = -\sum_{x} \Pr(x|y) \log \Pr(x|y) \tag{0.2}$$

Conditioning never increases the entropy, i.e.

$$\mathrm{ENT}(X|Y) \leq \mathrm{ENT}(X) \tag{0.3}$$

Mutual Information

$$\begin{align}
\mathrm{MI}(X;Y) &= \sum_{x,y} \Pr(x,y) \log \frac{\Pr(x,y)}{\Pr(x)\Pr(y)} \tag{0.4}\\
&= \mathrm{ENT}(X) - \mathrm{ENT}(X|Y) \tag{0.5}\\
&= \mathrm{ENT}(Y) - \mathrm{ENT}(Y|X) \tag{0.6}
\end{align}$$

Conditional Mutual Information

$$\begin{align}
\mathrm{MI}(X;Y|Z) &= \sum_{x,y,z} \Pr(x,y,z) \log \frac{\Pr(x,y|z)}{\Pr(x|z)\Pr(y|z)} \tag{0.7}\\
&= \mathrm{ENT}(X|Z) - \mathrm{ENT}(X|Y,Z) \tag{0.8}\\
&= \mathrm{ENT}(Y|Z) - \mathrm{ENT}(Y|X,Z) \tag{0.9}
\end{align}$$

0.5 Bayesian Conditioning

Bayesian Condition is specified by the formula

$$\Pr(\alpha|\beta) = \frac{\Pr(\alpha \wedge \beta)}{\Pr(\beta)} \tag{0.10}$$

In particular, not to be confused with Bayesian inference (to be added later).

0.6 Independence and Notations

Independence.

$$\alpha \perp\!\!\!\perp \beta \iff \Pr(\alpha|\beta) = \Pr(\alpha) \vee \Pr(\beta) = 0 \tag{0.11}$$

$$\iff \Pr(\alpha \wedge \beta) = \Pr(\alpha)\Pr(\beta) \tag{0.12}$$

Conditional Independence.

$$(\alpha \perp\!\!\!\perp \beta)|\gamma \iff \Pr(\alpha|\beta \wedge \gamma) = \Pr(\alpha|\gamma) \vee \Pr(\beta \wedge \gamma) = 0 \tag{0.13}$$

$$\iff \Pr(\alpha \wedge \beta|\gamma) = \Pr(\alpha|\gamma)\Pr(\beta|\gamma) \vee \Pr(\gamma) = 1 \tag{0.14}$$

Set Independence

$$I_{\Pr}(X, Z, Y) \iff (x \perp\!\!\!\perp y)|z, \quad \forall x, y, z \in X, Y, Z \tag{0.15}$$