

Reasoning under Uncertainty

A Note on Notation

关于命题变量 A

- 之前, A 是一个形成公式的命题变量:
 - $P(A)$ 是公式 A 为真的概率
 - $P(\neg A) = 1 - P(A)$ 是公式 $\neg A$ 为真的概率
- 从现在起, A 是一个取值在 $\{0, 1\}$ 的命题变量, 即 $P(A)$ 是两个可能输入值 $A = 1$ 和 $A = 0$ 的函数:
 - $P(A = 1)$ = 公式 A 为真的概率
 - $P(A = 0) = 1 - P(A = 1)$ = 公式 A 为假的概率

关于 $P(A, B) = P(A) \cdot P(B)$

- 声明 $P(A, B) = P(A) \cdot P(B)$ 意味着:
 - $P(A = 1, B = 1) = P(A = 1) \cdot P(B = 1)$
 - $P(A = 0, B = 1) = P(A = 0) \cdot P(B = 1)$
 - $P(A = 1, B = 0) = P(A = 1) \cdot P(B = 0)$
 - $P(A = 0, B = 0) = P(A = 0) \cdot P(B = 0)$

独立性 (Independence)

(主要是一个计算概念)

定义 (独立性)

- 两个变量 A 和 B 是独立的, 当且仅当它们的联合分布可以分解为所谓的边缘分布 (marginal distributions), 即:

$$P(A, B) = P(A) \cdot P(B) \quad (1)$$

- **解释:** 这表示 A 和 B 的联合概率可以表示为它们各自概率的乘积, 说明两个变量之间没有相互影响。
- 在这种情况下, $P(A|B) = P(A)$ 。记作: $A \perp B$ 。
 - **解释:** 这意味着知道 B 的信息不会提供关于 A 的额外信息, 反之亦然, 两个变量是统计独立的。

条件独立性 (Conditional Independence)

(主要是一个计算概念)

定义 (条件独立性)

- 两个变量 A 和 B 在给定变量 C 的条件下是独立的, 当且仅当它们的条件分布可以分解, 即:

$$P(A, B|C) = P(A|C) \cdot P(B|C) \quad (2)$$

- **解释：**这表示在已知 C 的情况下， A 和 B 的联合条件概率可以表示为它们各自条件概率的乘积，说明 C 提供了足够信息，使得 A 和 B 之间没有额外相关性。
- 在这种情况下，我们有 $P(A|B, C) = P(A|C)$ ，即在 C 的信息下， B 不会提供关于 A 的进一步信息。记作： $A \perp B|C$ 。
 - **解释：**这意味着在已知 C 的情况下， B 的信息不会影响对 A 的预测， A 和 B 在 C 条件下的分布是独立的。

有向图模型（Directed Graphical Models）

（又称贝叶斯网络、贝叶斯网、信念网络等）

[Judea Pearl, *Probabilistic Reasoning in Intelligent Systems*, 1988]

定义（贝叶斯网络，初步定义——后续讲座会深入）

- 有向图模型（DGM，简称贝叶斯网络）是一个关于变量 $\{X_1, \dots, X_D\}$ 的概率分布，可以表示为：

$$P(X_1, X_2, \dots, X_D) = \prod_{i=1}^D P(X_i | \text{pa}(X_i)) \quad (3)$$

其中， $\text{pa}(X_i)$ 是变量 X_i 的父变量，即 $X_i \notin \text{pa}(X_j) \forall X_j \in \text{pa}(X_i)$ 。

解释：这个公式表示一个联合概率分布可以通过多个条件概率的乘积来表示。每个变量 X_i 的概率依赖于它的父变量（父节点），这些父变量通过有向边连接到 X_i ，形成一个有向无环图（DAG）。

- 有向图模型可以用有向无环图（Directed Acyclic Graph, DAG）表示，其中命题变量作为节点，箭头从父节点指向子节点。

解释：DAG 是一种图形结构，用于可视化变量之间的因果或依赖关系。箭头表示信息流向，父节点影响子节点，方便表示复杂的概率关系。

原子独立性结构（Atomic Independence Structures）

（有向无环图（DAGs）暗示条件独立性，但不暗示依赖性！）

概述

- 对于一元和二元变量图，条件独立性是显而易见的。
- 对于三元子图（涉及三个变量），有三种可能的结构：

(i) 链式结构

- **图：** $A \rightarrow B \rightarrow C$
- **因子分解：**

$$P(A, B, C) = P(C|B) \cdot P(B|A) \cdot P(A)$$
- **含义：** $A \perp C|B$ （在已知 B 的条件下， A 和 C 独立），但不是，例如， $A \perp C$ （ A 和 C 总体上不独立）。

解释：在这个结构中， B 是中介变量， A 通过 B 影响 C 。在知道 B 的值后， A 和 C 之间没有直接相关性，因此条件独立。

(ii) 树形结构

- 图： $A \leftarrow B \rightarrow C$
- 因子分解：
$$P(A, B, C) = P(A|B) \cdot P(C|B) \cdot P(B)$$
- 含义： $A \perp C|B$ （在已知 B 的条件下， A 和 C 独立），但不是，例如， $A \perp C$ （ A 和 C 总体上不独立）。

解释： B 是共同原因，影响 A 和 C 。在知道 B 的值后， A 和 C 之间的关系被“屏蔽”，因此条件独立。

(iii) 发散结构

- 图： $A \rightarrow C \leftarrow B$
- 因子分解：
$$P(A, B, C) = P(B|A, C) \cdot P(C) \cdot P(A)$$
- 含义： $A \perp C$ （ A 和 C 总体上独立），但不是，例如， $A \perp C|B$ （在已知 B 的条件下， A 和 C 不独立）。

解释： C 是共同结果，由 A 和 B 分别影响。总体上 A 和 C 独立，但如果知道 B 的值， A 和 C 可能通过 B 产生依赖性。

总结

- 有向无环图（DAG）通过其结构隐含条件独立性，但不保证依赖性。
- 这些结构展示了变量之间的因果关系和独立性如何通过图结构和概率分解表示。

解释：DAGs 是一种强大的工具，用于表示变量之间的依赖和独立关系，三种结构展示了不同因果关系下的条件独立性或依赖性。