

知识点总结（逻辑）

- 命题逻辑是应用一套形式化规则对以符号表示的描述性陈述进行推理的系统。
- 函数与谓词的区别
 - 函词中个体变元用个体常量（来自定义域）代入后结果仍是个体（值域）
 - 谓词中个体变元用个体常量带入后就变成了命题
- 知识图谱推理

- FOIL信息增益值计算方法如下：

$$FOIL_Gain = \widehat{m}_+ \cdot \left(\log_2 \frac{\widehat{m}_+}{\widehat{m}_+ + \widehat{m}_-} - \log_2 \frac{m_+}{m_+ + m_-} \right)$$

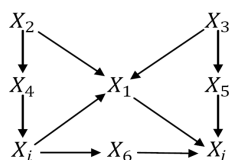
- FOIL的过程

知识图谱推理: FOIL (First Order Inductive Learner)

| FOIL算法 | |
|--------|--|
| 输入： | 目标谓词P，P的训练样例（正例集合E ⁺ 和反例集合E ⁻ ），其他背景知识 |
| 输出： | 推导得到目标谓词P的推理规则 |
| 1 | 将目标谓词作为所学习推理规则的结论 |
| 2 | 将其他谓词逐一作为前提约束谓词加入推理规则，计算所得到推理规则的FOIL信息增益值，选取最优前提约束谓词以生成新推理规则，并将训练样例集中与该推理规则不符的样例去掉 |
| 3 | 重复2过程，直到所得到的推理规则不覆盖任意反例 |

- 路径排序
 - 1. 对于目标关系Father，生成四组训练样例，一个为正例、三个为负例：正例：(David, Mike)负例：(David, James)，(James, Ann)，(James, Mike)
 - 2. 从知识图谱采样得到路径，每一路径链接上述每个训练样例中两个实体：
 - (David, Mike)对应路径：Couple→Mother
 - (David, James)对应路径：Father→Mother^{^(-1)}（Mother^{^(-1)}与 Mother为相反关系）
 - (James, Ann)对应路径：Mother→Sibling
 - (James, Mike)对应路径：Couple→Father
 - 3. 对于每一个正例/负例，判断上述四条路径可否链接其包含的两个实体，将可链接(记为1)和不可链接(记为0)作为特征，于是每一个正例/负例得到一个四维特征向量：
 - (David, Mike)：{[1, 0, 0, 0], 1}
 - (David, James)：{[0, 1, 0, 0], -1}
 - (James, Ann)：{[0, 0, 1, 0], -1}
 - (James, Mike)：{[0, 0, 1, 1], -1}
 - 4. 依据训练样本，训练分类器M

- 辛普森悖论
 - 在某些情况下，忽略潜在的“第三个变量”（如性别就是专业和身高之外的第三个变量），可能会改变已有的结论，而我们常常却一无所知。
- 因果推理的层级
 - 可观测性问题
 - 决策行动问题
 - 反事实问题(Counterfactual)
- 因果推理：有向无环图

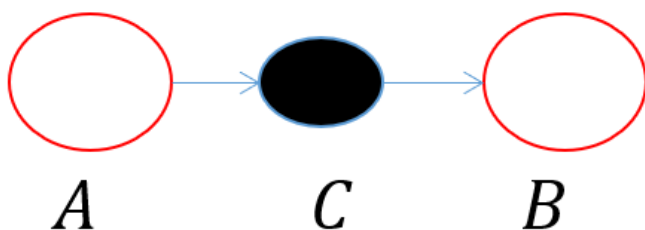


联合分布可表示为：

$$\begin{aligned}
 &P(X_1, X_2, X_3, X_4, X_5, X_6, X_i, X_j) \\
 &= P(X_2) \times P(X_3) \times P(X_1|X_2, X_3, X_i) \\
 &\quad \times P(X_4|X_2) \times P(X_5|X_3) \times P(X_6|X_i) \times P(X_j|X_5) \\
 &\quad \times P(X_i|X_4) \times P(X_j|X_5, X_6)
 \end{aligned}$$

- 因果推理： \emptyset -分离
 - \emptyset -分离用于判断集合 \emptyset 中变量是否与集合 \emptyset 中变量相互独立
 - 当C取值固定(可观测, observed), 有 $P(A)P(C|A) = P(C)P(A|C)$ serial connection
 - $P(A, B | C) = (P(A, B, C)) / (P(C)) = (P(A)P(C|A)P(B|C)) / (P(C)) = P(A | C)P(B|C)$

D-分离的例子 (serial connection)

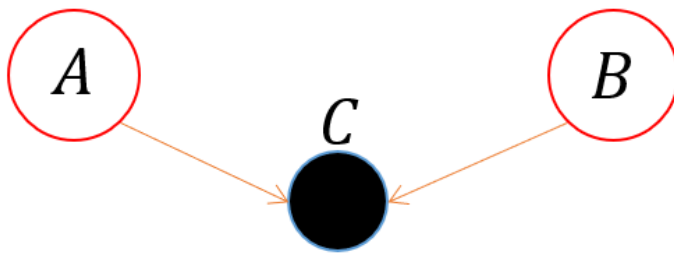


- 当C取值固定(observed), diverging connection
- $P(A, B | C) = (P(A, B, C)) / (P(C)) = (P(C)P(A|C)P(B|C)) / (P(C)) = P(A | C)P(B|C)$



- 可见A和B在C取值固定情况下，是条件独立的
 - 如果C不固定，则有 $P(A, B) = \sum_C P(A | B)P(B | C)P(C)$
 - 由于 $P(A, B) \neq P(A)P(B)$ ，因此A和B在条件C下不独立的

- V-structure connection

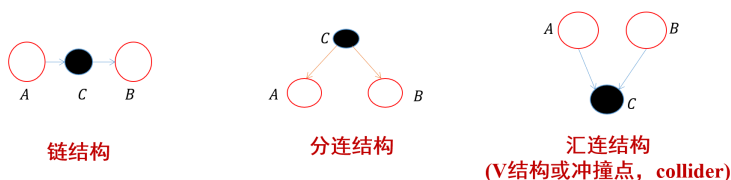


- 当C取值固定，A和B在条件C下是不独立的（是相关的）
- 当C不作为观测点，A和B在条件C下是独立的
- 结论

| 链结构(chain) | 分连结构(fork) | 汇连（或碰撞）结构 (collider) |
|----------------|----------------|----------------------|
| | | |
| Z和X是相关的 | X和Z是相关的 | Z和X是相关的 |
| Y和Z是相关的 | Y和Z是相关的 | Z和Y是相关的 |
| Y和X很有可能是相关的 | Y和X很有可能是相关的 | Y和X是相互独立的 |
| 给定Z时，Y和X是条件独立的 | 给定Z时，Y和X是条件独立的 | 给定Z时，Y和X是相关的 |

- 当两个节点是有向分离时，意味着这两个节点相互独立

• **D-分离**：如果A和B之间所有路径都是阻塞的，那么A和B就是关于C条件独立的；否则A和B不是关于C条件独立



- 在因果图上，若节点 α 和节点 β 之间的每一条路径都是阻塞的

- 称节点 α 和节点 β 是有向分离的
- 反之，称节点 α 和节点 β 是有向连接的

- 推理方法 推理方式 说明
- 归纳推理 如果 A_i (i 为若干取值)，那么B 从若干事实出发推理出一般性规律
- 演绎推理 如果A，那么B A是B的前提、但不是唯一前提，因此A是B的充分条件。当然，在特殊情况下A也可作为B的充分必要条件
- 因果推理 因为A，所以B A是B的唯一前提，因此“如果没有A，那么没有B”也成立。