# 逻辑与推理2

主讲：王亚星、刘夏雷、郭春乐
南开大学计算机学院

致谢：本课件主要内容来自浙江大学吴飞教授、
南开大学程明明教授

# 提纲

- 命题逻辑
- 谓词逻辑
- **知识图谱推理**
- 因果推理

# 教材勘误

- **P007页，A是B的充分条件意味着A蕴含B，即A是B的子集。**
- **P028页，定理2.3公式（1）和（4）表述错误，并不等价。**

**例 5.2** 证明：

(1) $\forall x(A(x) \lor B(x)) \not\Leftrightarrow \forall xA(x) \lor \forall xB(x)$

(2) $\exists x(A(x) \land B(x)) \not\Leftrightarrow \exists xA(x) \land \exists xB(x)$

其中 $A(x)$，$B(x)$ 为含 $x$ 自由出现的公式.

**证** (1) 取 $A(x)=F(x)$，$B(x)=G(x)$，并证明 $\forall x(F(x) \lor G(x)) \leftrightarrow \forall xF(x) \lor \forall xG(x)$ 不是永真式，其中 $F(x)$ 和 $G(x)$ 是谓词变项.

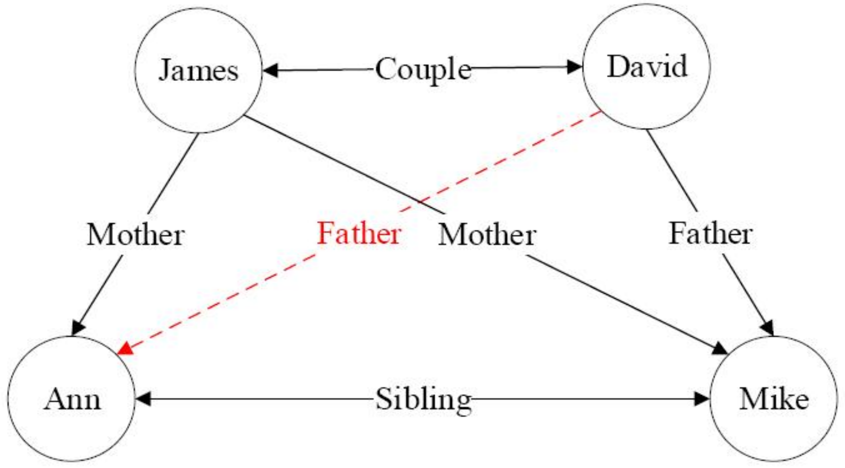
取解释 $I$ 为：个体域为自然数集合 $\mathbf{N}$，$\bar{F}(x)$：$x$ 是奇数，$\bar{G}(x)$：$x$ 是偶数，则 $\forall x(F(x) \lor G(x))$ 在解释 $I$ 下为真命题，而 $\forall xF(x) \lor \forall xG(x)$ 为假命题. 故 $\forall x(F(x) \lor G(x)) \leftrightarrow \forall xF(x) \lor \forall xG(x)$ 不是永真式.

可以类似地证明(2).

例 5.2 说明，全称量词“$\forall$”对“$\lor$”无分配律，存在量词“$\exists$”对“$\land$”无分配律，请初学者务必注意. 如果把式中的 $B(x)$ 改为没有 $x$ 自由出现的 $B$ 时，就得到量词辖域收缩与扩张等值式(5.3)和(5.4)中的第一个式子.

# 知识图谱：基本概念

- 知识图谱可视为包含多种关系的图。
  - 每个节点是一个**实体**（如人名、地名、事件和活动等），任意两个节点之间的边表示这两个节点之间存在的关系。
- 可将知识图谱中任意两个相连节点及其连接边表示成一个三元组（*triplet*）
  - 即 $(left_{node}, relation, right_{node})$
  - 例 $(David,\ Father,\ Mike)$



# 知识图谱：知识工程

- **Edward A. Feigenbaum于1977年提出了知识工程的概念**
    - 开发了基于知识的系统(knowledge-based systems)。该系统利用计算机程序包含大量知识、规则、和推理机制，为显示生活中的问题提供解决方案。
    - 基于知识的系统主要表现形式是用于模仿专家决策过程的专家系统。
    - Turing Award in 1994

Feigenbaum, E. A., The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering, *Proceedings of the International Joint Conference on Artificial Intelligence(IJCAI),1977*

# 知识图谱：以水利工程为例

- **概念之间层次化关系(ontology)：**
  - 如：工程→水利工程
  - 与Wordnet等早期本体知识构建不同，现有方法多在传统分类法(Taxonomy)中结合大众分类(Folksonomy)和机器学习来构建语义网络分类体系。
- **概念对应的例子或实体(instance/entity)**
  - 如三峡大坝和葛洲坝等属于水利工程这一概念。
  - 一般通过分类识别等手段实现。

# 知识图谱：以水利工程为例

- **概念或实体的属性：**
  - 属性是对概念或实体内涵的描述。
- **概念或实体之间的关系：**
  - 如三峡大坝和葛洲坝之间具有“梯级调节”关系。

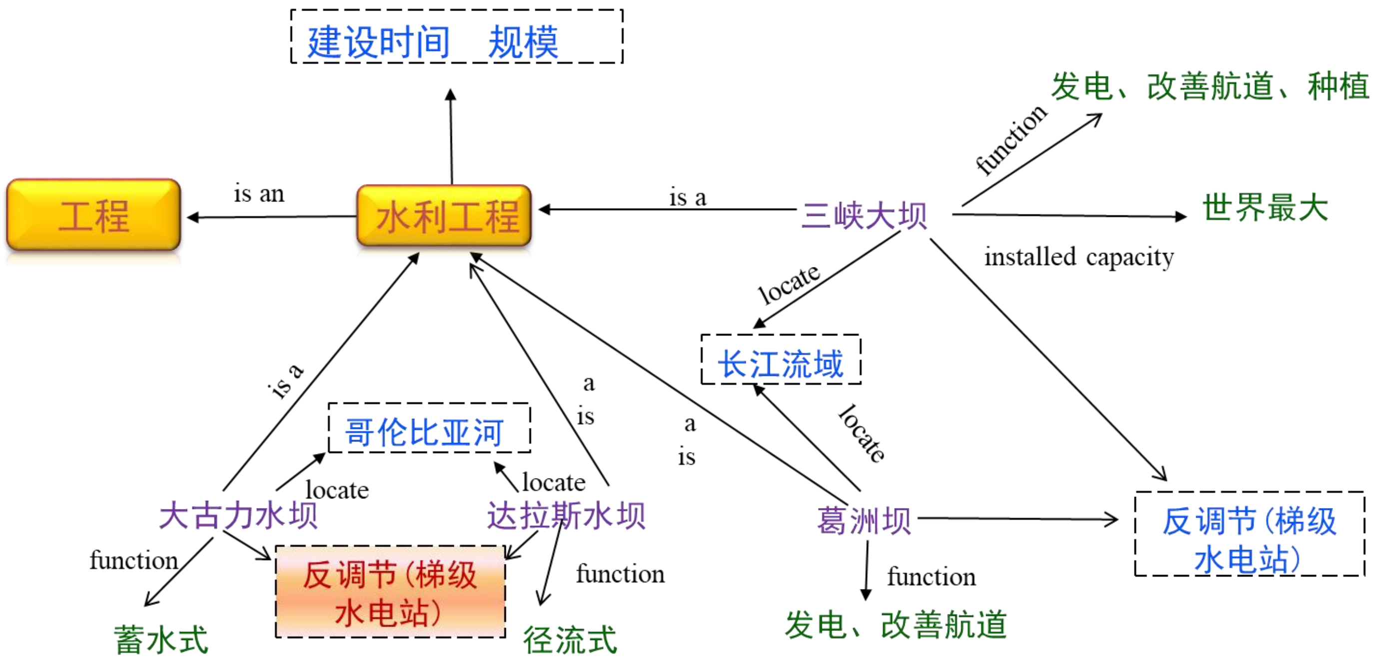
# 知识图谱：以水利工程为例

- **概念或实体的<u>属性描述和关系表达</u>一般通过三元组来表示：**
  - (entity, relation, entity)或(subject, predicate, object )

- **学习概念或实体属性描述及其关联关系是丰富知识图谱的关键！**

# 知识图谱：以水利工程为例

- **知识图谱推理 (inference)：**
  - 通过机器学习等方法对知识图谱所蕴含关系进行挖掘。



# 知识图谱的构成

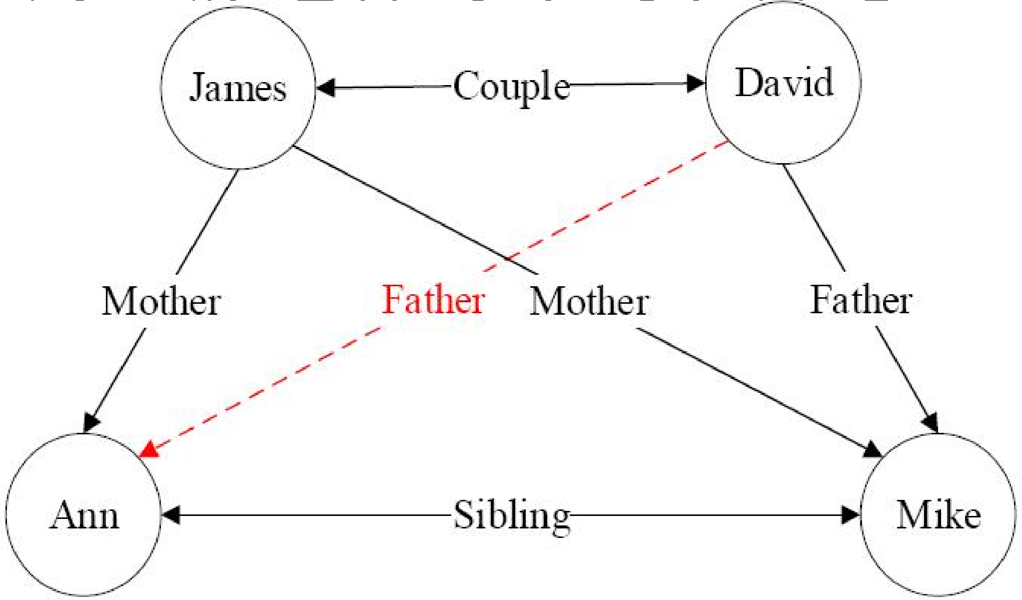
- **知识图谱一般可通过标注多关系图（labeled multi-relational graph）来表示。**
  - 概念：层次化组织
  - 实体：概念的示例化描述
  - 属性：对概念或实体的描述信息
  - 关系：概念或实体之间的关联
  - 推理规则：可产生语义网络中上述新的元素

# 知识图谱推理

- **知识图谱中存在连线的两个实体可表达为三元组形式**
  - < *left_node*, *relation*, *right_node* >
  - 这种三元组也可以表示为一阶逻辑(first order logic, FOL)的形式
  - 为基于知识图谱的推理创造了条件
  - 例如从<奥巴马，出生地，夏威夷>和<夏威夷，属于，美国>两个三元组，可推理得到<奥巴马，国籍，美国>。

# 知识图谱推理

- **可利用一阶谓词来表达刻画知识图谱中节点之间存在的关系**
  - $< James, Couple, David >$关系可用一阶逻辑的形式来描述
- $Couple(x, y)$**是一阶谓词**
  - $Couple$是实体之间具有的关系
  - $x$和$y$是谓词变量




- **图中可推知David和Ann具有父女关系**
  - 但这一关系在初始图(无红线)中并不存在，需要推理得到

# 知识图谱推理

- **问题：如何从知识图谱中推理得到**

$$father(David, Ann)$$

**归纳学习**

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \wedge Couple(x, z) \rightarrow Father(x, y))$$

# 知识图谱推理: 归纳学习

- **归纳逻辑程序设计 (inductive logic programming，ILP)算法**
  - ILP是机器学习和逻辑程序设计交叉领域的研究内容
  - ILP使用**一阶谓词逻辑**进行知识表示，通过修改和扩充逻辑表达式对现有知识归纳，完成推理任务。
  - 作为ILP的代表性方法，FOIL（First Order Inductive Learner）通过序贯覆盖实现规则推理。

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

- 推理手段: positive examples + negative examples + background knowledge examples $\Longrightarrow$ hypothesis

$$\underbrace{(\forall x)(\forall y)(\forall z)(Mother(z,y) \wedge Couple(x,z)}_{\text{前提约束谓词 (学习得到)}} \rightarrow \underbrace{Father(x,y)}_{\text{目标谓词 (已知)}})$$

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

- **目标谓词**：$Father(x, y)$
- **目标谓词只有一个正例**$Father(David, Mike)$
  - 反例在知识图谱中一般不会显式给出，但可从知识图谱中构造出来。如从知识图谱中已经知道$Couple(David, James)$成立
  - 则$Father(David, James)$可作为目标谓词$P$的一个反例， 记为 $\neg Father\ (David, James)$
- 只能在**已知两个实体的关系**且**确定其关系与目标谓词相悖**时，才能将这两个实体用于构建目标谓词的反例
  - 而不能在不知两个实体是否满足目标谓词前提下将它们来构造目标谓词的反例

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

- 目标谓词：$Father(x, y)$
- 背景知识：目标谓词以外的其它谓词实例化结果

| | |
|---|---|
| 背景知识样例集合 | Sibling(Ann, Mike)<br>Couple(David, James)<br>Mother(James, Ann)<br>Mother(James, Mike) |

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

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \wedge Couple(x, z) \rightarrow \boxed{Father(x, y)})$$

前提约束谓词(学习得到)

目标谓词(已知)

如何推理

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

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \wedge Couple(x, z) \rightarrow \boxed{Father(x, y)})$$

前提约束谓词(学习得到)

目标谓词(已知)

- 推理思路：从一般到特殊，**逐步给目标谓词**添加**前提约束谓词**，直到所构成的**推理规则不覆盖任何反例。**

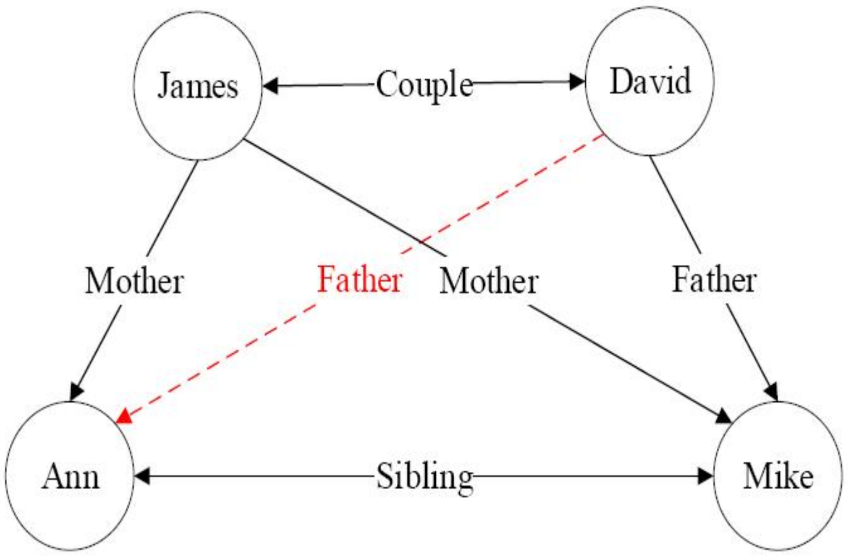
# 知识图谱推理: FOIL (First Order Inductive Learner)

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \land Couple(x, z) \rightarrow \boxed{Father(x, y)})$$

前提约束谓词(学习得到)

目标谓词(已知)

- **对目标谓词或前提约束谓词中的变量赋予值具体**
  - 如将上例这一推理规则所包含的目标谓词$Father(x, y)$中$x$和$y$分别赋值为$David$和$Ann$



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

- **哪些谓词好呢？可以作为目标谓词的前提约束谓词?**
  - FOIL中信息增益值(information gain)
- **FOIL信息增益值计算方法如下:**

$$FOIL\_Gain = \widehat{m}_+ \cdot (\log_2 \frac{\widehat{m}_+}{\widehat{m}_+ + \widehat{m}_-} - \log_2 \frac{m_+}{m_+ + m_-})$$

  - 其中，$\widehat{m}_+$和$\widehat{m}_-$是增加前提约束谓词后所得新推理规则覆盖的正例和反例的数量，$m_+$和$m_-$是原推理规则所覆盖的正例和反例数量

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

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \wedge Couple(x, z) \rightarrow \boxed{Father(x, y)})$$

前提约束谓词(学习得到)

目标谓词(已知)

- $Mother(\cdot, \cdot)$
- $Sibling(\cdot, \cdot)$
- $Couple(\cdot, \cdot)$

- **依次**将谓词加入到推理规则中作为前提约束谓词，并计算所得到新推理规则的FOIL增益值。
- 基于计算所得FOIL增益值来选择最佳前提约束谓词。

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

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \wedge Couple(x, z) \rightarrow \boxed{Father(x, y)})$$

前提约束谓词(学习得到)

目标谓词(已知)

| | | | |
|---|---|---|---|
| 背景知识样例集合 | Sibling(Ann, Mike)<br>Couple(David, James)<br>Mother(James, Ann)<br>Mother(James, Mike) | 目标谓词训练样例集合 | Father(David, Mike)<br>¬Father(David, James)<br>¬Father(James, Ann)<br>¬Father(James, Mike)<br>¬Father(Ann, Mike) |

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

- 给定目标谓词，此时推理规则只有目标谓词
- 因此推理规则所覆盖的正例和反例的样本数分别是训练样本中正例和反例的数量
  - $m_+ = 1$， $m_- = 4$

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 目标谓词 | 前提约束谓词 | 正例 | 反例 | 信息增益值 |
| $Father(x, y) \leftarrow$ | 空集 | $m_+ = 1$ | $m_- = 4$ | $FOIL\_Gain$ |
| $Father(x, y) \leftarrow$ | $Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 3$ | 0.32 |
| | $Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 2$ | 0.74 |
| | $Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\boldsymbol{Couple(x, z)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 1}$ | 1.32 |
| | $Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |

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

- 将$Mother(x, y)$作为前提约束谓词加入，可得到推理规则
  - $Mother(x, y) \rightarrow Father(x, y)$
- 背景知识中$Mother(x, y)$有两个实例
  - $Mother(James, Ann)$
  - $Mother(James, Mike)$
- 对于$Mother(James, Ann)$这一实例
  - $x = James$, $y = Ann$, 将$x$和$y$代入$Father(x, y)$，可知在训练样本中$Father(James, Ann)$是一

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 目标谓词 | 前提约束谓词 | 正例 | 反例 | 信息增益值 |
| $Father(x, y) \leftarrow$ | 空集 | $m_+ = 1$ | $m_- = 4$ | $FOIL\_Gain$ |
| $Father(x, y) \leftarrow$ | $Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 3$ | 0.32 |
| | $Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 2$ | 0.74 |
| | $Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\boldsymbol{Couple(x, z)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 1}$ | 1.32 |
| | $Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |

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

- **将$Mother(x, y)$作为前提约束谓词加入，可得到推理规则**
  - $Mother(x, y) \rightarrow Father(x, y)$
- **背景知识中$Mother(x, y)$有两个实例**
  - $Mother(James, Ann)$
  - $Mother(James, Mike)$
- **对于$Mother(James, \mathrm{Mike})$这一实例**
  - $x = James$，$y = Mike$，将$x$和$y$代入$Father(x, y)$，可知在训练样本中$Father(James, Mike)$是一个

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 目标谓词 | 前提约束谓词 | 正例 | 反例 | 信息增益值 |
| $Father(x, y) \leftarrow$ | 空集 | $m_+ = 1$ | $m_- = 4$ | $FOIL\_Gain$ |
| $Father(x, y) \leftarrow$ | $Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 3$ | 0.32 |
| | $Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 2$ | 0.74 |
| | $Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\boldsymbol{Couple(x, z)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 1}$ | 1.32 |
| | $Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |

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

$Mother(x,y) \rightarrow Father(x,y)$

- 覆盖正例和反例数量分别为 $\hat{m}_+ = 0, \ \hat{m}_- = 2$
- 由于$\hat{m}_+ = 0$，代入$FOIL_Gain$公式时会出现负无穷的情况，此时$FOIL_Gain$记为NA（Not Available）

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 目标谓词 | 前提约束谓词 | 正例 | 反例 | 信息增益值 |
| $Father(x,y) \leftarrow$ | 空集 | $m_+ = 1$ | $m_- = 4$ | $FOIL\_Gain$ |
| $Father(x,y) \leftarrow$ | $Mother(x,y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(x,z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(y,x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(y,z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z,x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z,y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 3$ | 0.32 |
| | $Sibling(x,y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(x,z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(y,x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(y,z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z,x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z,y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 2$ | 0.74 |
| | $Couple(x,y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\boldsymbol{Couple(x,z)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 1}$ | 1.32 |
| | $Couple(y,x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(y,z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(z,x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |

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

- **如果将$Couple(x, z)$作为前提约束谓词加入,**
  - 可得到如下推理规则

    $Couple(x, z) \rightarrow Father(x, y)$
- **在背景知识中，$Couple(x, z)$只有一个实例**
  - 即$x = David$，$z = James$
  - 将其代入$Father(x, y)$得到

    $Father(David, y)$

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 目标谓词 | 前提约束谓词 | 正例 | 反例 | 信息增益值 |
| $Father(x, y) \leftarrow$ | 空集 | $m_+ = 1$ | $m_- = 4$ | $FOIL\_Gain$ |
| | $Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 3$ | 0.32 |
| | $Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| $Father(x, y) \leftarrow$ | $Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 2$ | 0.74 |
| | $Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\boldsymbol{Couple(x, z)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 1}$ | 1.32 |
| | $Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |

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

- 训练样本中存在正例以及反例
  - $Father(David, Mike)$
  - $\neg Father(David, James)$,
  - 即$Couple(x, z) \rightarrow Father(x, y)$覆盖正反例数量分别为1和1。
  - 信息增益值为：

$$\hat{m}_+ \cdot (\log_2 \frac{\hat{m}_+}{\hat{m}_+ + \hat{m}_-} - \log_2 \frac{m_+}{m_+ + m_-})$$

$$= 1 \cdot (\log_2 \frac{1}{1+1} - \log_2 \frac{1}{1+4}) = 1.32$$

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 目标谓词 | 前提约束谓词 | 正例 | 反例 | 信息增益值 |
| $Father(x, y) \leftarrow$ | 空集 | $m_+ = 1$ | $m_- = 4$ | $FOIL\_Gain$ |
| $Father(x, y) \leftarrow$ | $Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |
| | $Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Mother(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 3$ | 0.32 |
| | $Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Sibling(z, y)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 2$ | 0.74 |
| | $Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\boldsymbol{Couple(x, z)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 1}$ | 1.32 |
| | $Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 2$ | NA |

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

| | |
|---|---|
| **Back-ground knowledge** | Sibling(Ann, Mike)<br>Couple(David, James)<br>Mother(James, Ann)<br>Mother(James, Mike) |
| **Positive and negative samples** | Father(David, Mike)<br>¬Father(David, James)<br>~~¬Father(James, Ann)~~<br>~~¬Father(James, Mike)~~<br>~~¬Father(Ann, Mike)~~ |

- $Couple(x, z)$ 加入后信息增益最大
- 将 $Couple(x, z)$ 加入推理规则，得到 $Couple(x, z) \rightarrow Father(x, y)$新推理规则
- 将训练样例中与该推理规则不符的样例去掉。这里不符指当$Couple(x, z)$中$x$取值为David时，与$Father$(David, )或¬$Father$(David, )无法匹配的实例。
- 训练样本集中只有正例$Father$(David, Mike)和负例¬$Father$(David, James)两个实例

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

$Mother(z, y) \wedge Couple(x, z)$

$\rightarrow Father(x, y)$

已知：

$Mother$(James, Ann)

$Couple$(David,James)

于是： x=David, z=James,因此y=Ann 或者Mike.

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 现有规则 | 拟加入前提约束谓词 | 正例 | 反例 | |
| $Father(x, y) \leftarrow Couple(x, z)$ | | $m_+ = 1$ | $m_- = 1$ | 1.32 |
| $Father(x, y) \leftarrow Couple(x, z)$ | $\wedge Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge \boldsymbol{Mother(z, y)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 0}$ | **1** |
| | $\wedge Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(z, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\wedge Couple(x, z)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 1$ | 0 |
| | $\wedge Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- =0$ | NA |
| | $\wedge Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |

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

- **推理**$Mother(z, y)$**加入信息增益最大**
- **将**$Mother(z, y)$**加入，得到新规则**

$$Mother(z, y) \wedge Couple(x, z)$$

$$\rightarrow Father(x, y)$$

- **当**$x = David$、$y = Mike$、$z = James$**时，该推理规则覆盖训练样本集合中正例**$Father(David, Mike)$**且不覆盖任意反例，因此算法学习结束。**

| 推理规则 | | 推理规则涵盖的正例和反例数 | | FOIL信息增益值 |
|---|---|---|---|---|
| 现有规则 | 拟加入前提约束谓词 | 正例 | 反例 | |
| $Father(x, y) \leftarrow Couple(x, z)$ | | $m_+ = 1$ | $m_- = 1$ | 1.32 |
| $Father(x, y) \leftarrow Couple(x, z)$ | $\wedge Mother(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Mother(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge \boldsymbol{Mother(z, y)}$ | $\boldsymbol{\hat{m}_+ = 1}$ | $\boldsymbol{\hat{m}_- = 0}$ | **1** |
| | $\wedge Sibling(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(x, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Sibling(z, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Couple(x, y)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 1$ | NA |
| | $\wedge Couple(x, z)$ | $\hat{m}_+ = 1$ | $\hat{m}_- = 1$ | 0 |
| | $\wedge Couple(y, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Couple(y, z)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |
| | $\wedge Couple(z, x)$ | $\hat{m}_+ = 0$ | $\hat{m}_- = 0$ | NA |

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

$$(\forall x)(\forall y)(\forall z)(Mother(z, y) \wedge Couple(x, z) \rightarrow \boxed{Father(x, y)})$$

前提约束谓词(学习得到)　　目标谓词(已知)

- **推理手段**: ***examples(positive*** + ***negative*** + ***background)*** ⟹ **hypothesis**

| 背景知识样例集合 | | 目标谓词训练样例集合 | |
|---|---|---|---|
| | Sibling(Ann, Mike) | | Father(David, Mike) |
| | Couple(David, James) | | ¬Father(David, James) |
| | Mother(James, Ann) | | ¬Father(James, Ann) |
| | Mother(James, Mike) | | ¬Father(James, Mike) |
| | | | ¬Father(Ann, Mike) |

给定目标谓词，FOIL算法从实例（正例、反例、背景样例）出发，不断测试所得到推理规则是否还包含反例，一旦不包含负例，则学习结束，展示了“**归纳学习**”能力。

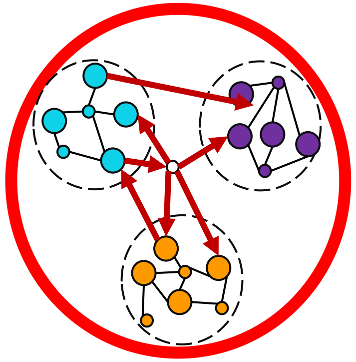
# 知识图谱推理: FOIL (First Order Inductive Learner)

| 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，并将训练样例集合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 |

# 知识图谱推理

推理机

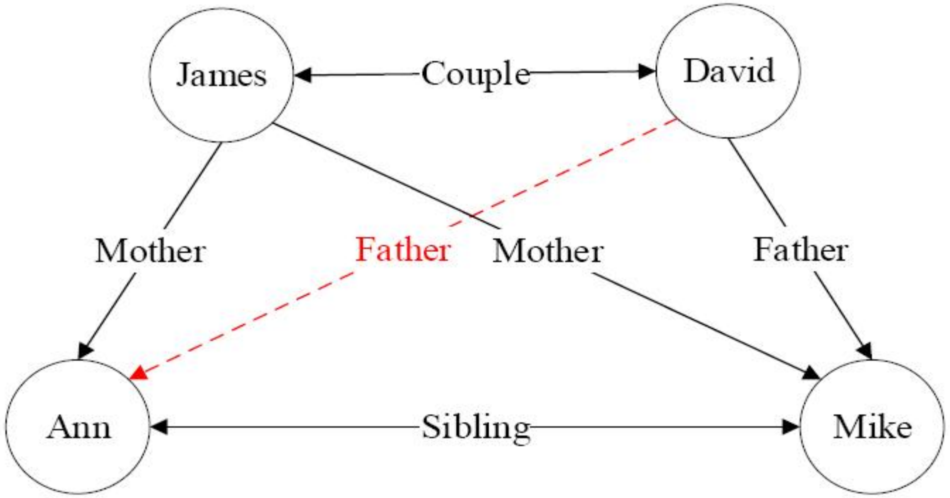
- 概念
- 实体
- 属性
- 关系



知识图谱的扩充

# 知识图谱推理：路径排序

**Score(*Father*(*David*, *Ann*))**




一个简单的家庭关系知识图谱

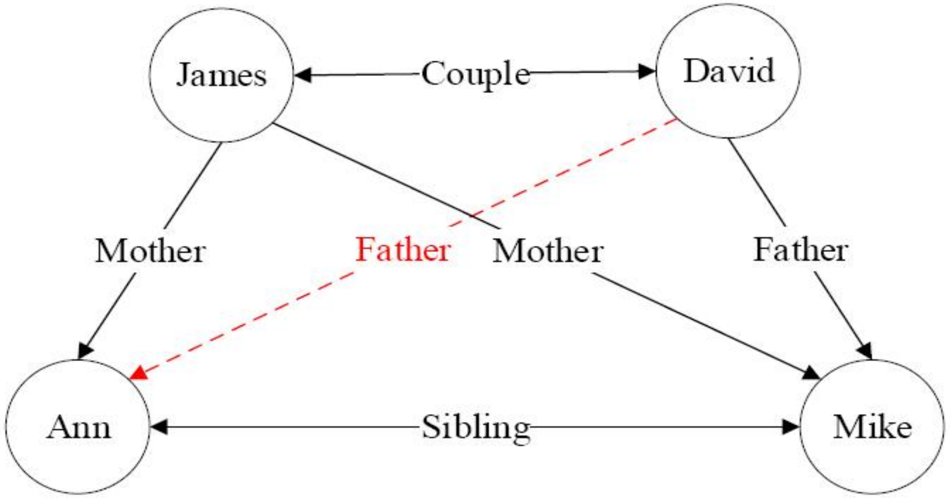
将实体之间的关联路径作为特征，来学习目标关系的分类器

即：判断David和Ann之间的路径关联是否足够支持表述*Father*这一关系。

Ni Lao, William W. Cohen, Relational retrieval using a combination of path-constrained random walks, *Machine learning*, 2010, 81(1): 53-67, 2010

# 知识图谱推理：路径排序

$\mathbf{Score}(Father(David, Ann))$




一个简单的家庭关系知识图谱

$$score(s,t) = \sum_{\pi_j \in p_l} \theta_j \, P(s \rightarrow t; \pi_j)$$

$p_l$是链接节点 $s$ 和节点 $t$ 的所有路径集合

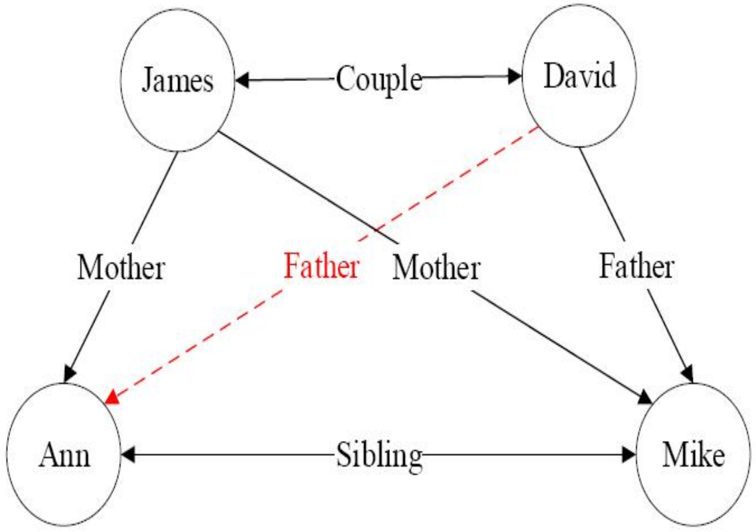
$\theta_j$ 是某一条路径 $\pi_j$的权重

$P$ 是路径$\pi_j$ 概率值大小

# 知识图谱推理：路径排序

$$score(s,t)=\sum_{\pi_j \in p_l} \theta_j\ P(s \to t;\pi_j)$$

Score($Father(David, Ann)$)

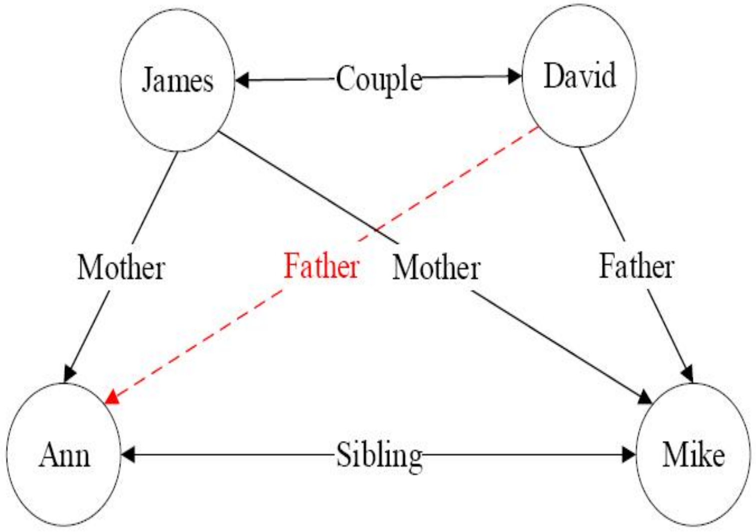



- **特征抽取**：生成并选择路径特征集合。生成路径的方式有随机游走（random walk）、广度优先搜索、深度优先搜索等。
- **特征计算**：计算每个训练样例的特征值$P(s \to t;\pi_j)$。该特征值可以表示从实体节点$s$出发，通过关系路径$\pi_j$到达实体节点$t$的概率；也可以表示为布尔值，表示实体$s$到实体$t$之间是否存在路径$\pi_j$；还可以是实体$s$和实体$t$之间路径出现频次、频率等。
- **分类器训练**：根据训练样例的特征值，为目标关系训练分类器。当训练好分类器后，即可将该分类器用于推理两个实体之间是否存在目标关系。

# 知识图谱推理：路径排序

$$score(s,t)=\sum_{\pi_j \in p_l} \theta_j \ father(s \rightarrow t;\pi_j)$$

Score($Father(David, Ann)$)




给定目标关系：$Father(s,t)$

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：$Couple \rightarrow Mother$

(David, James)对应路径：$Father \rightarrow Mother^{-1}$

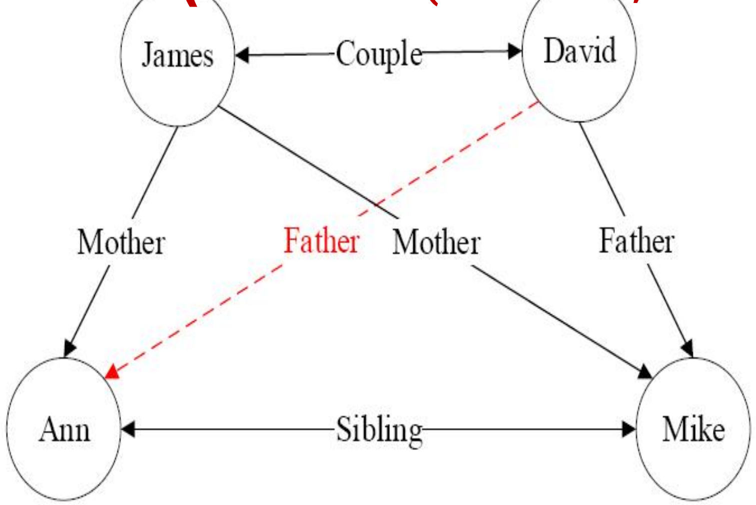
($Mother^{-1}$与 $Mother$为相反关系)

(James, Ann)对应路径：$Mother \rightarrow Sibling$

(James, Mike)对应路径：$Couple \rightarrow Father$

# 知识图谱推理：路径排序

Score($Father(David, Ann)$) $score(s,t)=\sum_{\pi_j\in p_l}\theta_j\ father(s\to t;\pi_j)$




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)：$\{[0,0,1,1],-1\}$

(David, Mike)对应路径：$Couple\to Mother$

(David, James)对应路径：$Father\to Mother^{-1}$

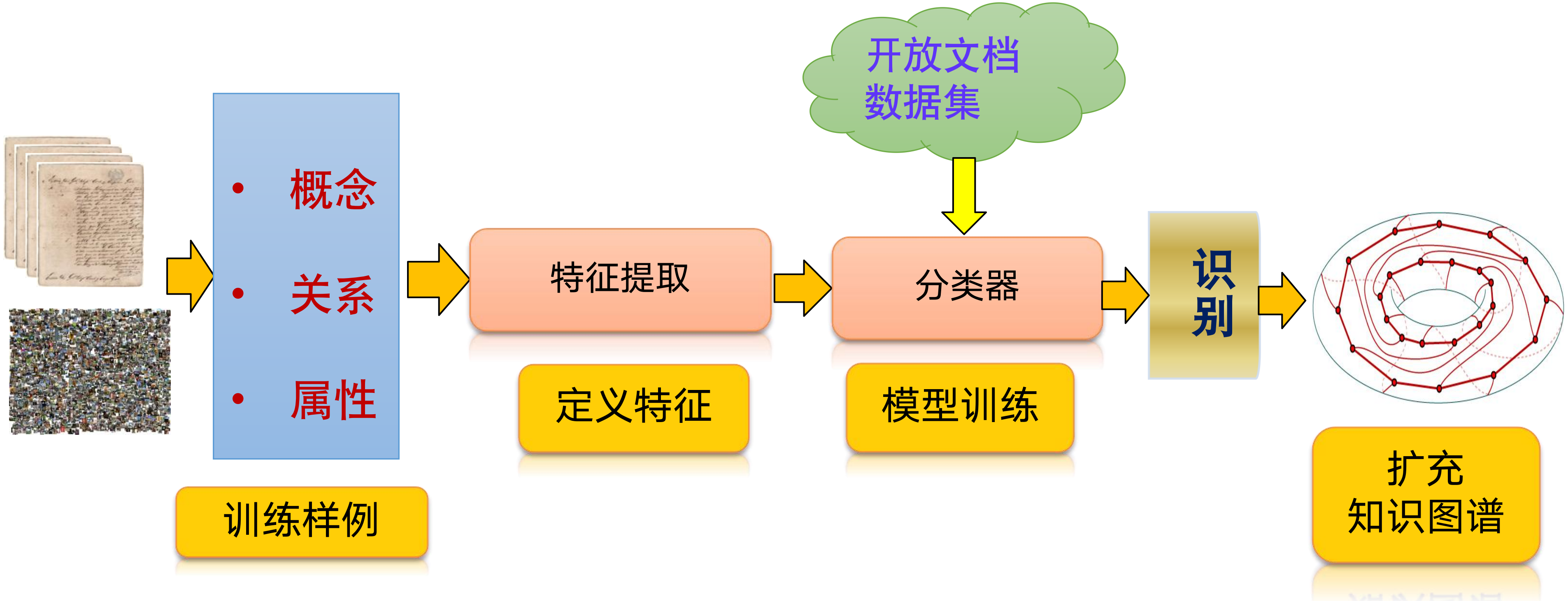
($Mother^{-1}$与 $Mother$为相反关系)

(James, Ann)对应路径：$Mother\to Sibling$

(James, Mike)对应路径：$Couple\to Father$

4. 依据训练样本，训练分类器$M$

# 知识图谱推理：机器学习

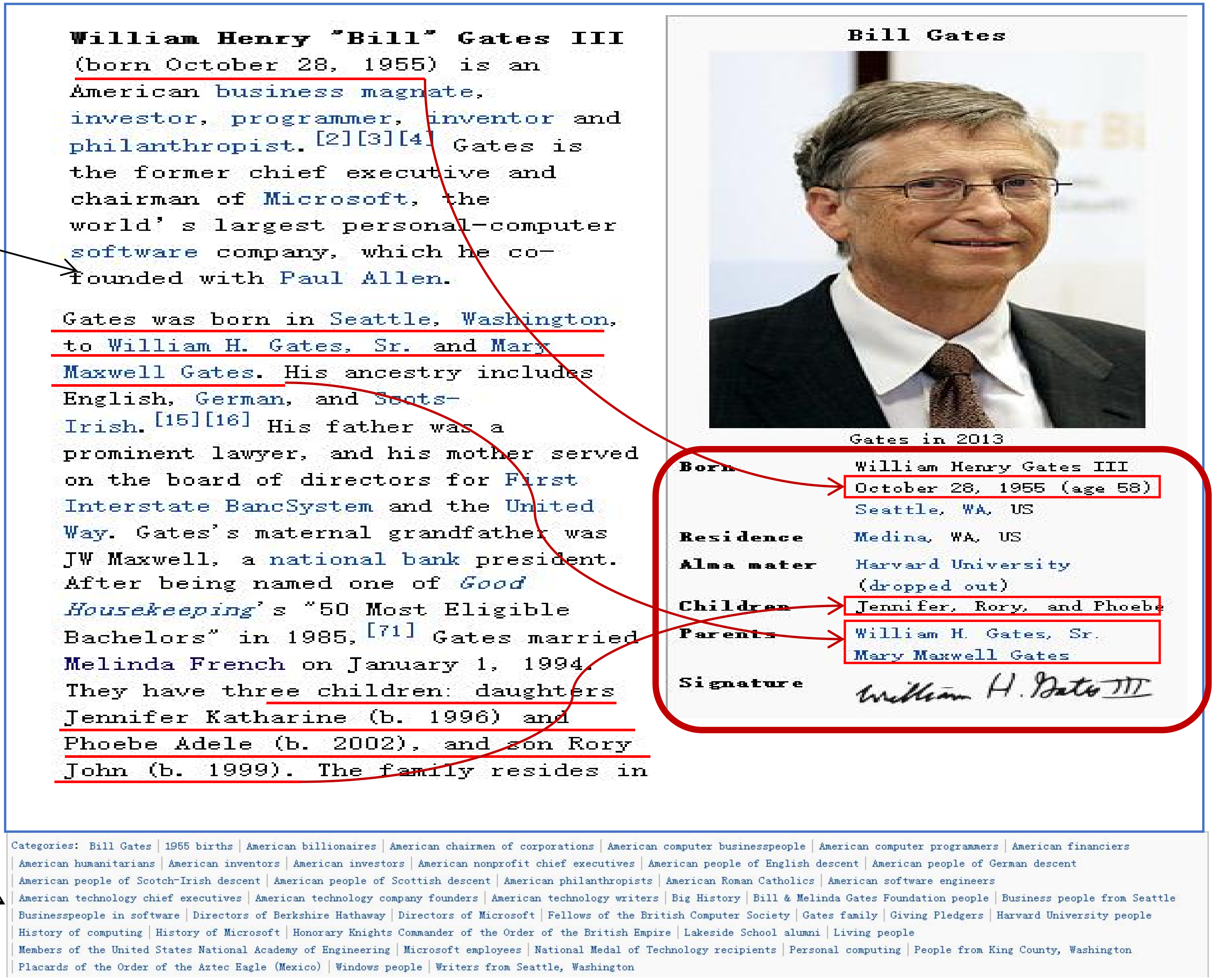


# 知识图谱构造流程：以Wiki为例子

正文描述

属性定义
及描述

44个
类别
标签



William Henry "Bill" Gates III (born October 28, 1955) is an American business magnate, investor, programmer, inventor and philanthropist.[2][3][4] Gates is the former chief executive and chairman of Microsoft, the world's largest personal-computer software company, which he co-founded with Paul Allen.

Gates was born in Seattle, Washington, to William H. Gates, Sr. and Mary Maxwell Gates. His ancestry includes English, German, and Scots-Irish.[15][16] His father was a prominent lawyer, and his mother served on the board of directors for First Interstate BancSystem and the United Way. Gates's maternal grandfather was JW Maxwell, a national bank president. After being named one of *Good Housekeeping*'s "50 Most Eligible Bachelors" in 1985,[71] Gates married Melinda French on January 1, 1994. They have three children: daughters Jennifer Katharine (b. 1996) and Phoebe Adele (b. 2002), and son Rory John (b. 1999). The family resides in

**Bill Gates**

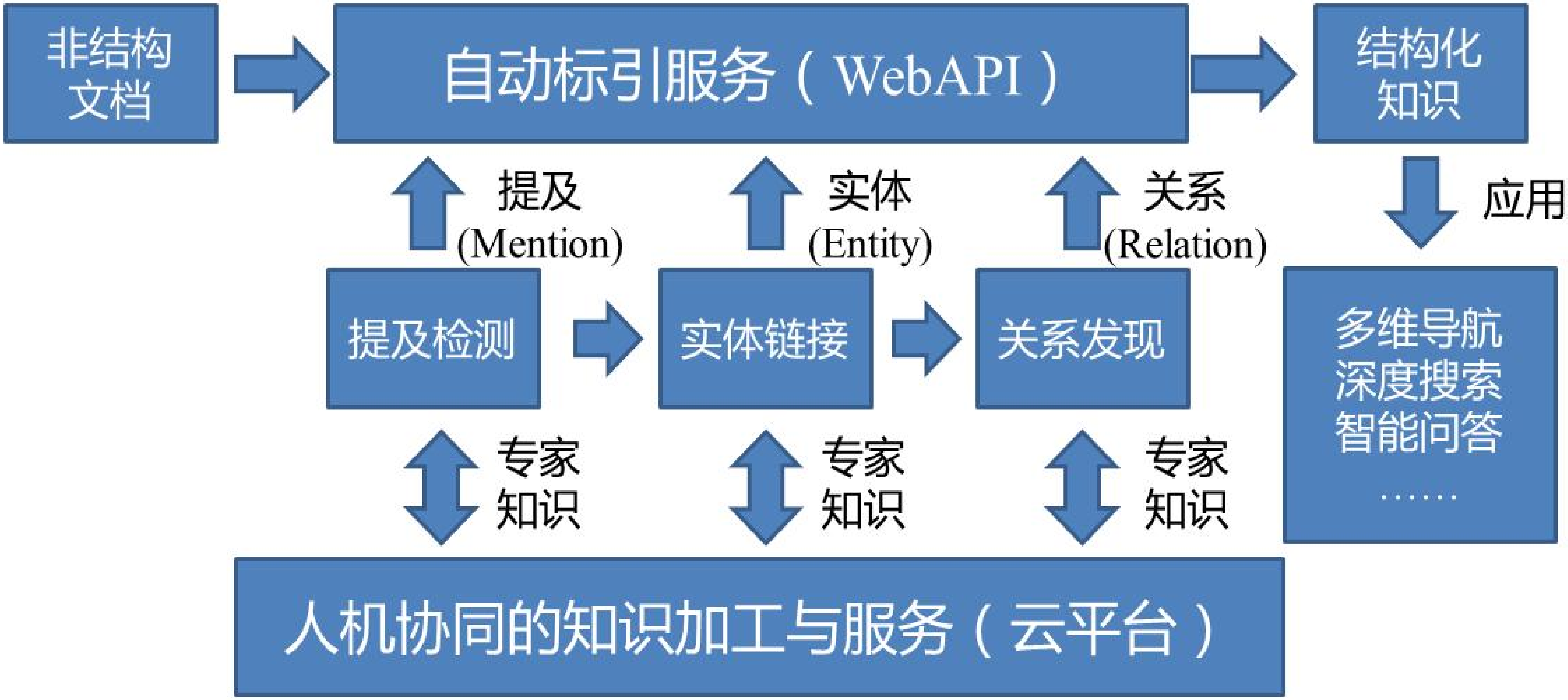
Gates in 2013

| | |
|---|---|
| **Born** | William Henry Gates III<br>October 28, 1955 (age 58)<br>Seattle, WA, US |
| **Residence** | Medina, WA, US |
| **Alma mater** | Harvard University (dropped out) |
| **Children** | Jennifer, Rory, and Phoebe |
| **Parents** | William H. Gates, Sr.<br>Mary Maxwell Gates |
| **Signature** | William H. Gates III |

Categories: Bill Gates | 1955 births | American billionaires | American chairmen of corporations | American computer businesspeople | American computer programmers | American financiers | American humanitarians | American inventors | American investors | American nonprofit chief executives | American people of English descent | American people of German descent | American people of Scotch-Irish descent | American people of Scottish descent | American philanthropists | American Roman Catholics | American software engineers | American technology chief executives | American technology company founders | American technology writers | Big History | Bill & Melinda Gates Foundation people | Business people from Seattle | Businesspeople in software | Directors of Berkshire Hathaway | Directors of Microsoft | Fellows of the British Computer Society | Gates family | Giving Pledgers | Harvard University people | History of computing | History of Microsoft | Honorary Knights Commander of the Order of the British Empire | Lakeside School alumni | Living people | Members of the United States National Academy of Engineering | Microsoft employees | National Medal of Technology recipients | Personal computing | People from King County, Washington | Placards of the Order of the Aztec Eagle (Mexico) | Windows people | Writers from Seattle, Washington

Wiki中用户对〝Bill Gates〞这个实例的标注

# 知识图谱：从数据到知识、从知识到决策



# 提纲

- 命题逻辑
- 谓词逻辑
- 知识图谱推理
- 因果推理

# 因果推理：Simpson’s Paradox (辛普森悖论)

1973年伯克利本科生录取率

| | 男生 | | 女生 | |
|---|---|---|---|---|
| | 申请数 | 录取率 | 申请数 | 录取率 |
| 整体 | 8442 | 44% | 4321 | 35% |

男生录取率(44%)**远高于**女生(35%)

| 学院 | 男生 | | 女生 | |
|---|---|---|---|---|
| | 申请数 | 录取率 | 申请数 | 录取率 |
| A | 825 | 62% | 108 | 82% |
| B | 560 | 63% | 25 | 68% |
| C | 325 | **37%** | 593 | 34% |
| D | 417 | 33% | 375 | 35% |
| E | 191 | **28%** | 393 | 24% |
| F | 373 | 6% | 341 | 7% |

6个最大的院系中，4个院系女生录取率大于男生。如果按照这样的分类，**女生实际上比男生的录取率还高一点点。**

女生更愿意申请那些竞争压力很大的院系（比如英语系），但是男生却更愿意申请那些相对容易进的院系（比如工程学系）。

P.J. Bickel, E.A. Hammel, J.W. O’Connell, Sex bias in graduate admissions: Data from Berkeley, *Science*, 187(4175):398-404,1975

# 因果推理：Simpson’s Paradox (辛普森悖论)

- **计算机学院学生的高个率高于文学院（左表）。**
  - 分别比较两所学院男生和女生身高时，却发现计算机学院男生和女生的高个率均低于文学院
- **在总体样本上成立的某种关系却在分组样本里恰好相反。**

| 身高(cm) | 计算机 | 文学院 |
|---|---|---|
| 矮个人数 | 60 | 80 |
| 高个人数 | 290 | 270 |
| 高个率(%) | 82.9 | 77.1 |

左：计算机学院和文学院学生的身高情况

| | 计算机学院 | | 文学院 | |
|---|---|---|---|---|
| 身高(cm) | 男生 | 女生 | 男生 | 女生 |
| 矮个人数 | 35 | 25 | 10 | 70 |
| 高个人数 | 235 | 55 | 80 | 190 |
| 高个率(%) | 87 | 68.9 | 88.9 | 73.1 |

右：以性别分组后的计算机学院和文学院的学生身高情况

$$\frac{b}{a} < \frac{d}{c}, \frac{b'}{a'} < \frac{d'}{c'}$$

$$\Downarrow$$

$$\frac{b+b'}{a+a'} > \frac{d+d'}{c+c'}$$

# 因果推理：Simpson’s Paradox (辛普森悖论)

- **右表体现了男生比女生个子高这一现象**
  - 如计算机学院和文学院男生高个率都比女生高个率要大
  - 性别会影响专业选择，计算机男生多，而文学院女生多。因此，当计算机学院的样本中包含更多的男生，就会看到左表所呈现的情况：计算机学院的高个率高于文学院。

$$\frac{b}{a} < \frac{d}{c}, \frac{b'}{a'} < \frac{d'}{c'}$$

$$\Downarrow$$

$$\frac{b+b'}{a+a'} > \frac{d+d'}{c+c'}$$

| 身高(cm) | 计算机 | 文学院 |
|---|---|---|
| 矮个人数 | 60 | 80 |
| 高个人数 | 290 | 270 |
| 高个率(%) | 82.9 | 77.1 |

左：计算机学院和文学院学生的身高情况

| | 计算机学院 | | 文学院 | |
|---|---|---|---|---|
| 身高(cm) | 男生 | 女生 | 男生 | 女生 |
| 矮个人数 | 35 | 25 | 10 | 70 |
| 高个人数 | 235 | 55 | 80 | 190 |
| 高个率(%) | 87 | 68.9 | 88.9 | 73.1 |

右：以性别分组后的计算机学院和文学院的学生身高情况

# 因果推理：Simpson’s Paradox (辛普森悖论)

- **辛普森悖论表明**，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。从观测结果中寻找引发结果的原因，由果溯因，就是本节要介绍的因果推理

**不能只满足于数字或图表，必须考虑数据生成过程----因果模型**

# 因果推理的层级

- **Actions: B will be true if we do A.**
- **Counterfactuals: B would be different if A were true**

| | | |
|---|---|---|
| **可观测性问题(关联)** | What if we see A | $P(y\|A)$ |
| **决策行动问题(干预)** | What if we do A | $P(y\|do(A))$<br>(如果采取A行为，则B真) |
| **反事实问题(Counterfactual)** | What if we did things differently | $P(y^{'}\|A)$<br>(如果A为真，则B将不同) |
| Options: with what probability | | |

# 因果推理模型: 结构因果模型和因果图

- **结构因果模型(structural causal model, SCM)**
  - 也被称为因果模型（causal model）或Neyman–Rubin因果模型
  - Jerzy Neyman在1923年博士论文中(波兰语)提出的“潜在结果”（potential outcome）的概念
  - 之后，Donald Rubin发展了“潜在结果”这一概念，并将其和缺失数据的理论联系在一起。
- **因果图(causal diagram)由Judea Pearl于 1995年提出。**
- **每个结构因果模型$M$都与一个因果图$G$相对应**

# 因果推理：有向无环图

- **有向无环图(directed acyclic graphs, DAG)**
  - 一个无回路的有向图，即从图中任意一个节点出发经过任意条边，均无法回到该节点。刻画了图中所有节点之间的依赖关系。

- **DAG 可用于描述数据的生成机制**
  - 这样描述变量联合分布或者数据生成机制的模型，被称为“贝叶斯网络”（Bayesian network）

# 因果推理：有向无环图（DAG）

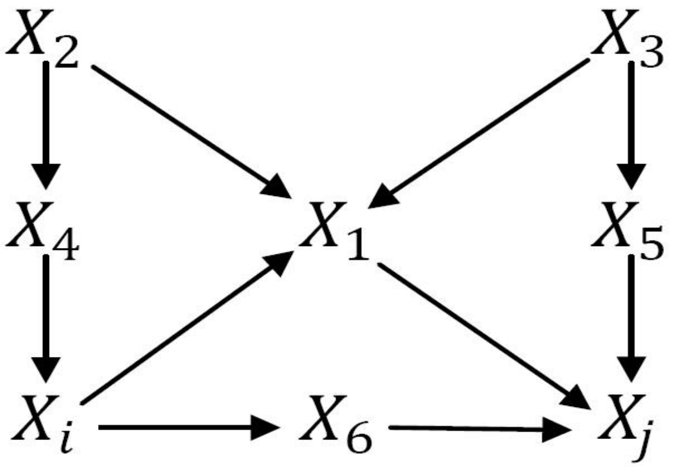
- 对于任意的DAG，模型中$d$个变量的联合概率分布由每个节点与其父节点之间条件概率$P(child|parents)$的乘积给出：

$$P(x_1, x_2, \cdots, x_d) = \prod_{j=1}^{d} P(x_j | x_{pa(j)})$$

其中，$x_{pa(j)}$表示节点$x_j$的父节点集合（所有指向$x_j$的节点）

# 因果推理：有向无环图（DAG）

- **一个有向无环图唯一地决定了一个联合分布**
- **一个联合分布不能唯一地决定有向无环图**
  - 反过来的结论不成立
  - 如联合分布$P(X_1, X_2) = P(x_1)P(x_2|x_1) = P(x_2)P(x_1|x_2)$



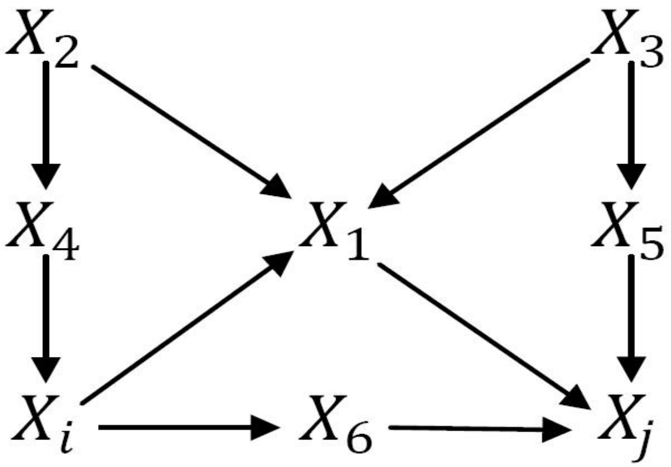

有向无环图DAG

联合分布可表示为：

$$P(X_1, X_2, X_3, X_4, X_5, X_6, X_i, X_j)$$
$$= P(X_2) \times P(X_3) \times P(X_1|X_2, X_3, X_i)$$
$$\times P(X_4|X_2) \times P(X_5|X_3) \times P(X_6|X_i) \times P(X_i|X_4)$$
$$\times P(X_j|X_1,\ X_5,\ X_6)$$

# 因果推理：有向无环图（DAG）

- **一个有向无环图唯一地决定了一个联合分布**
- **一个联合分布不能唯一地决定有向无环图**
  - 反过来的结论不成立
  - 如联合分布$P(X_1, X_2) = P(x_1)P(x_2|x_1) = P(x_2)P(x_1|x_2)$




有向无环图DAG

联合分布可表示为：

$$P(X_1, X_2, X_3, X_4, X_5, X_6, X_i, X_j)$$
$$= P(X_2) \times P(X_3) \times P(X_1|X_2, X_3, X_i)$$
$$\times P(X_4|X_2) \times P(X_5|X_3) \times P(X_6|X_i) \times P(X_i|X_4)$$
$$\times P(X_j|X_1, X_5, X_6)$$

# 因果推理：有向无环图（DAG）

- **例：假设某个有向无环图中存在一条依赖路径 $X \to Y \to Z$，其中$X$表示气候好，$Y$表示水果产量高，$Z$表示水果价格低，给出$P$(气候好，水果产量高，水果价格低)的联合概率。**
  - 使用乘积分解规则，将$P$(气候好，产量高，价格低)转换为：

$$P(\text{气候好}) \times P(\text{产量高}|\text{气候好}) \times P(\text{价格低}|\text{产量高})$$

假设：

$P(\text{气候好}) = 0.5$

$P(\text{产量高}|\text{气候好}) = 0.8$

$P(\text{价格低}|\text{产量高}) = 0.9$

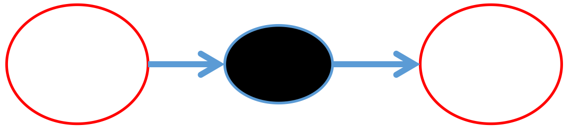
$P(\text{气候好，产量高，价格低})$
$= 0.5 \times 0.8 \times 0.9 = 0.36$

# 因果推理：D分离

- D分离用于判断集合A中变量是否与集合B中变量相互独立（给定集合C），记为$A \perp B \mid C$

D-分离的例子

(serial connection)




- 当C取值固定(可观测, observed)，有

$$P(A,B|C)=\frac{P(A,B,C)}{P(C)}=\frac{P(A)P(C|A)\,P(B|C)}{P(C)}=P(A|C)P(B|C)$$

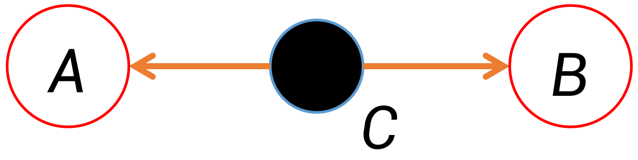
- 可见A和B在C取值固定情况下，是条件独立的
  - 注：上式利用了$P(A)P(C|A)=P(C)P(A|C)$

# 因果推理：D分离

- *D*分离用于判断集合*A*中变量是否与集合*B*中变量相互独立（给定集合*C*），记为$A \perp B \mid C$

D-分离的例子

(diverging connection)




- 当C取值固定(observed)，有

$$P(A,B|C) = \frac{P(A,B,C)}{P(C)} = \frac{P(C)P(A|C)\,P(B|C)}{P(C)} = P(A|C)P(B|C)$$

- 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

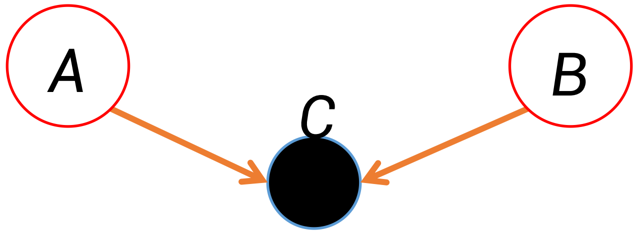
# 因果推理：D分离

- *D*分离用于判断集合*A*中变量是否与集合*B*中变量相互独立（给定集合*C*），记为$A \perp B \mid C$

$$P(A,B,C) = P(A)P(B)P(C|A,B)$$

D-分离的例子

(V-structure connection)




- 当*C*取值固定(observed)

$$P(A,B|C) = \frac{P(A,B,C)}{P(C)} = \frac{P(A)P(B)P(C|A,B)}{P(C)} \neq P(A)P(B)$$

- A和B在条件C下是不独立的（是相关的）

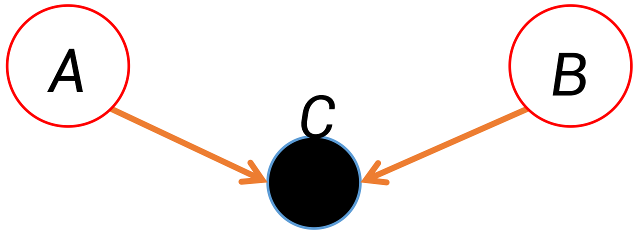
# 因果推理：*D*分离

- *D*分离用于判断集合*A*中变量是否与集合*B*中变量相互独立（给定集合*C*），记为$A \perp B \mid C$

$$P(A,B,C) = P(A)P(B)P(C|A,B)$$

D-分离的例子

(V-structure connection)

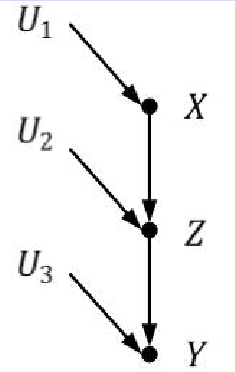
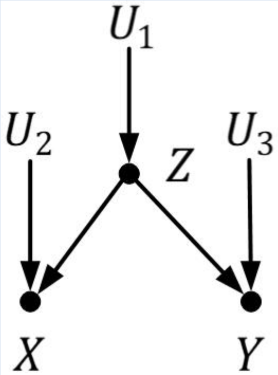
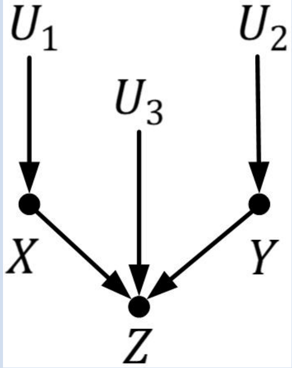



- 当*C*不作为观测点，A和B是独立的

$$P(A,B) = \sum_{C}(P(A)P(B)P(C|A,B)) = P(A)P(B)\boxed{\sum_{C}P(C|A,B)}$$

$= 1$

# 因果推理：D分离

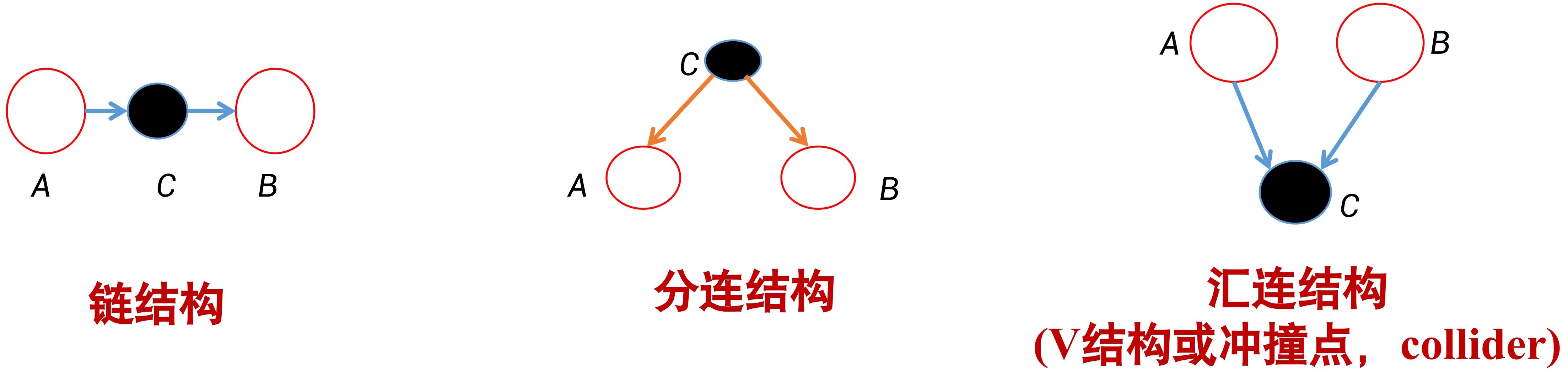
| 链结构(chain) | 分连结构(fork) | 汇连（或碰撞）结构(collider) |
|---|---|---|
|  |  |  |
| Z和X是相关的 | X和Z是相关的 | Z和X是相关的 |
| Y和Z是相关的 | Y和Z是相关的 | Z和Y是相关的 |
| Y和X很有可能是相关的 | Y和X很有可能是相关的 | Y和X是相互独立的 |
| **给定Z时，Y和X是条件独立的** | **给定Z时，Y和X是条件独立的** | **给定Z时，Y和X是相关的** |

# 因果推理：D分离

- **D-分离**：对于一个DAG图，如果$A$、$B$、$C$是三个集合（可以是单独的节点或者是节点的集合），为了判断$A$和$B$ 是否是 $C$ 条件独立的， 在DAG图中考虑所有$A$和$B$之间的路径(不管方向)。对于其中的一条路径，如果满足以下两个条件中的任意一条，则称这条路径是阻塞（block）的：路径中存在节点$X$
  - $X$是链结构或分连结构节点，且$X \in C$
  - $X$是汇连结构节点，并且$X$或$X$ 后代不包含在 C中

# 因果推理：D分离

- **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



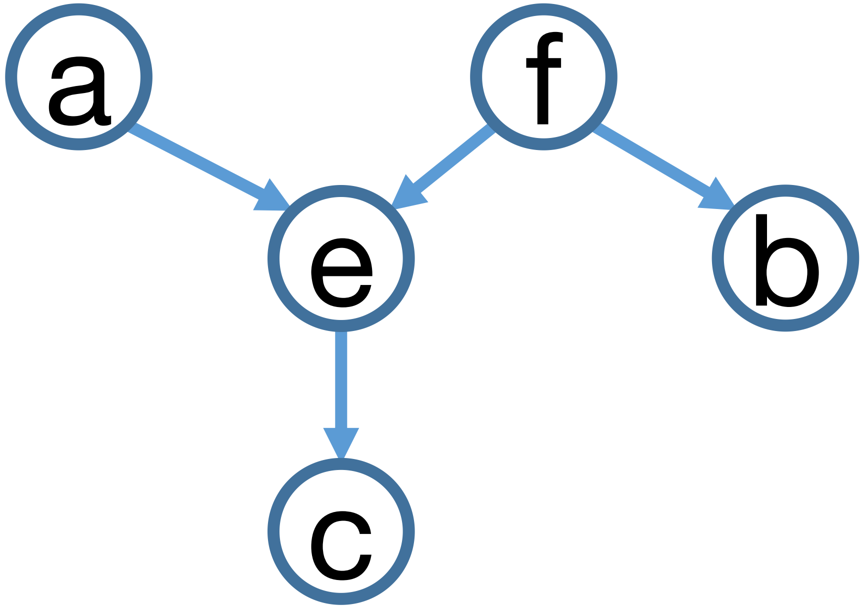
设置

下面描述中正确的有

A a和b是条件c下独立的

B a和b是条件e下独立的

C a和b是条件f下独立的




提交

# 因果推理：*D*分离 (directional separation)

- *D*分离方法可用于判断因果图上任意变量间相关性和独立性
- 在因果图上，若节点*X*和节点*Y*之间的每一条路径都是阻塞的
  - 称节点*X*和节点*Y*是有向分离的
  - 反之，称节点*X*和节点*Y*是有向连接的
- 当两个节点是**有向分离**时，意味着这两个节点**相互独立**

# 因果推理：干预(intervention)和$do$算子(do-calculus)

- DAG中具有链接箭头的节点之间存在某种“因果关系”。
- 要在 DAG 上引入“因果”的概念，需要引进*do* 算子
  - do-calculus的意思可理解为 “干预” （intervention）
- 在 DAG 中， $do(X_i) = x_i'$，表示将DAG中指向节点$X_i$的有向边全部切断，并且将$X_i$的值固定为常数$x_i'$
- 在这样操作后，所得到新的DAG中变量联合分布为：

$$P(x_1, x_2, \cdots, x_d | do(X_i) = \boldsymbol{x_i'})$$

# 因果推理：反事实推理 (counterfactual model)

- “反事实”框架是科学哲学家大卫·刘易斯（David Lewis）等人提出的推断因果关系的标准。
- 事实是指在某个特定变量$(A)$的影响下可观测到的某种状态或结果$(B)$。“反事实”是指在该特定变量$(A)$ 取负向值时可观测到的状态或结果$(B')$
- 条件变量对于结果变量的因果性就是$A$成立时$B$的状态与$A$取负向值时“反事实”状态$(B')$之间的差异。
- 如果这种差异存在且在统计上是显著的，说明条件变量与结果变量存在因果关系。

# 谢谢!