逻辑与推理2

主讲：郭春乐、刘夏雷

南开大学计算机学院

1. 和均是原子命题,"如果那么"是由和组合得到的复合命题。下面对"如果那么"这一复合命题描述不正确的是( )

A "如果那么"定义的是一种蕴涵关系(即充分条件)

B "如果那么"意味着命题包含着命题,即是的子集

C 无法用真值表来判断"如果那么"的真假

D 当不成立时,"如果那么"恒为真

提交

2. 下面哪个逻辑等价关系是
不成立的（ ）

A $\quad \forall x \neg P(x) \equiv \neg \exists x P(x)$

B $\quad \neg \forall x P(x) \equiv \exists x \neg P(x)$

C $\quad \forall x P(x) \equiv \neg \exists x \neg P(x)$

D $\quad \exists x P(x) \equiv \neg \forall x P(x)$

提交

3.下列句子中不是命题的是
（）。

A x>2。

B 今天天气真好啊！

C 天津大学是中国近代第一所大学。

D 所有实数的平方都大于或等于0。

提交

4.应用归结法证明以下命题集是不可满足的。

a)$\alpha \vee \beta$;

b)$\beta \rightarrow \gamma$;

c)$\neg\alpha \wedge \neg\gamma$;

解　应用归结法：

(1) $\alpha \vee \beta$ (已知);

(2) $\neg\beta \vee \gamma$ (b 进行蕴涵消除);

(3) $\alpha \vee \gamma$ (由 1 和 2);

(4) $\neg(\alpha \vee \gamma)$ (c 使用 De Morgan 定律)

(5) 3 和 4 矛盾，因此原命题集是不可满足的。

正常使用主观题需2.0以上版本雨课堂

作答

5.证明苏格拉底三段论"所有人都是要死的，苏格拉底是人，所以苏格拉底是要死的"。

解：设 $F(x)$：x 是人；$G(x)$：x 是要死的人；a 是苏格拉底。

前提：$(\forall x)(F(x) \rightarrow G(x))$；$F(a)$。

结论：$G(a)$。

证明：（1）$(\forall x)(F(x) \rightarrow G(x))$；

（2）$F(a) \rightarrow G(a)$（1的全称量词消去）；

（3）$F(a)$；

（4）$G(a)$（2和3的假言推理）。

正常使用主观题需2.0以上版本雨课堂

作答

勘误

- P007

A是B的充分条件意味着A蕴含B，即B是A的子集。

- P028

公式（1）和（4）是单项箭头。

（1） $((\forall x)A(x) \vee (\forall x)B(x) \Rightarrow (\forall x)(A(x) \vee B(x)))$
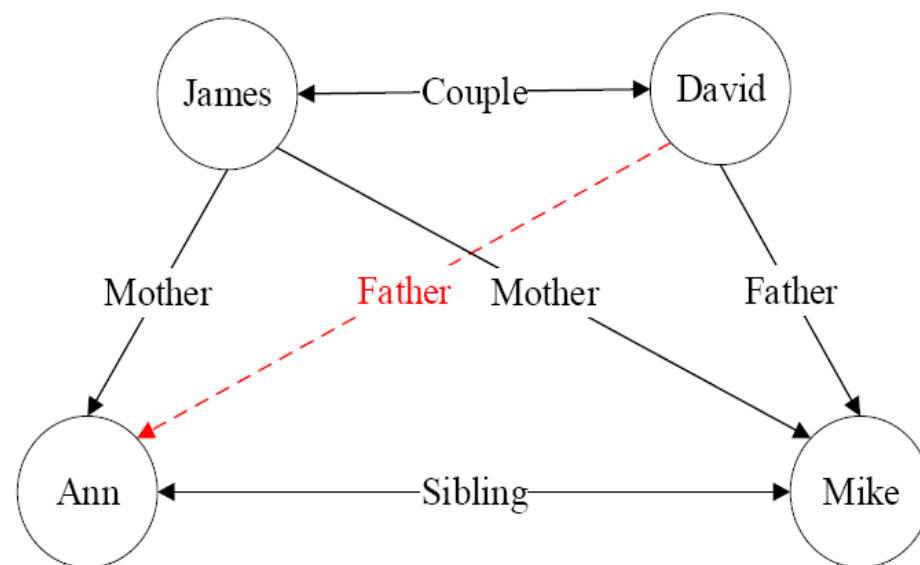
（4） $(\exists x)(A(x) \wedge B(x)) \Rightarrow (\exists x)A(x) \wedge (\exists x)B(x)$

提纲

- 命题逻辑

- 谓词逻辑

- **知识图谱推理**

- 因果推理

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

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，如水利工程具有建设时间和规模等属性、三峡大坝具有发电功能等属性。

- **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。

- **概念或实体的属性描述和关系表达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。

建设时间 规模

工程 ←is an— 水利工程 ←is a— 三峡大坝 —function→ 发电、改善航道、种植

三峡大坝 → 世界最大

installed capacity

三峡大坝 —locate→ 长江流域

is a

哥伦比亚河

大古力水坝

locate

达拉斯水坝

locate

function

蓄水式

反调节(梯级水电站)

function

径流式

长江流域 ←locate— 葛洲坝

葛洲坝 → 反调节(梯级水电站)

葛洲坝 —function→ 发电、改善航道
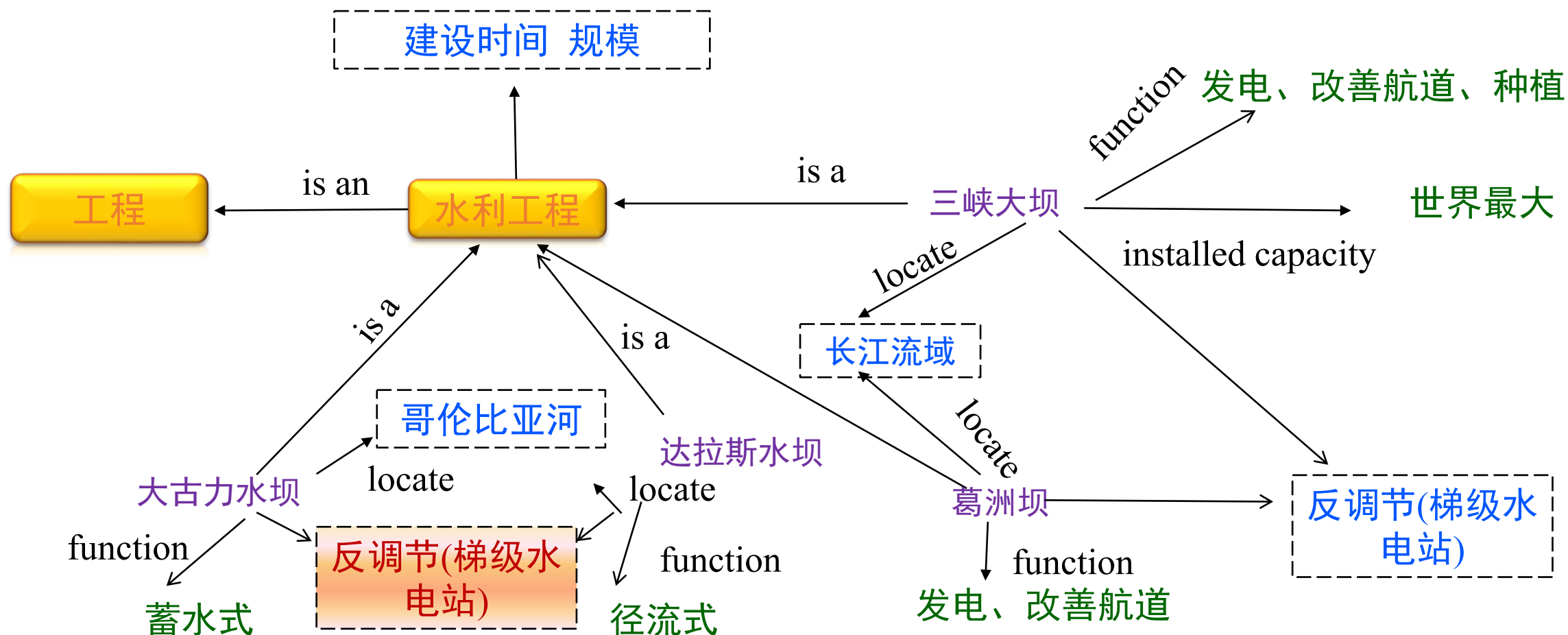
知识图谱的构成

- 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
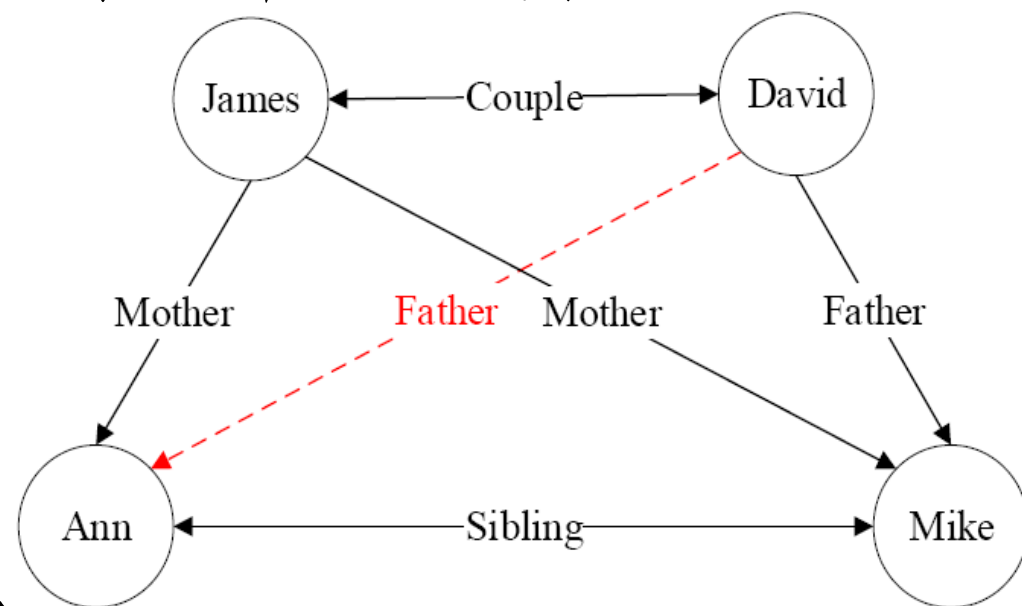
知识图谱推理

- 问题：如何从知识图谱中推理得到

$$father(David, Ann)$$

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

如果能够学习得到这条规则，该有多好？（从具体例子中学习是归纳推理的范畴）
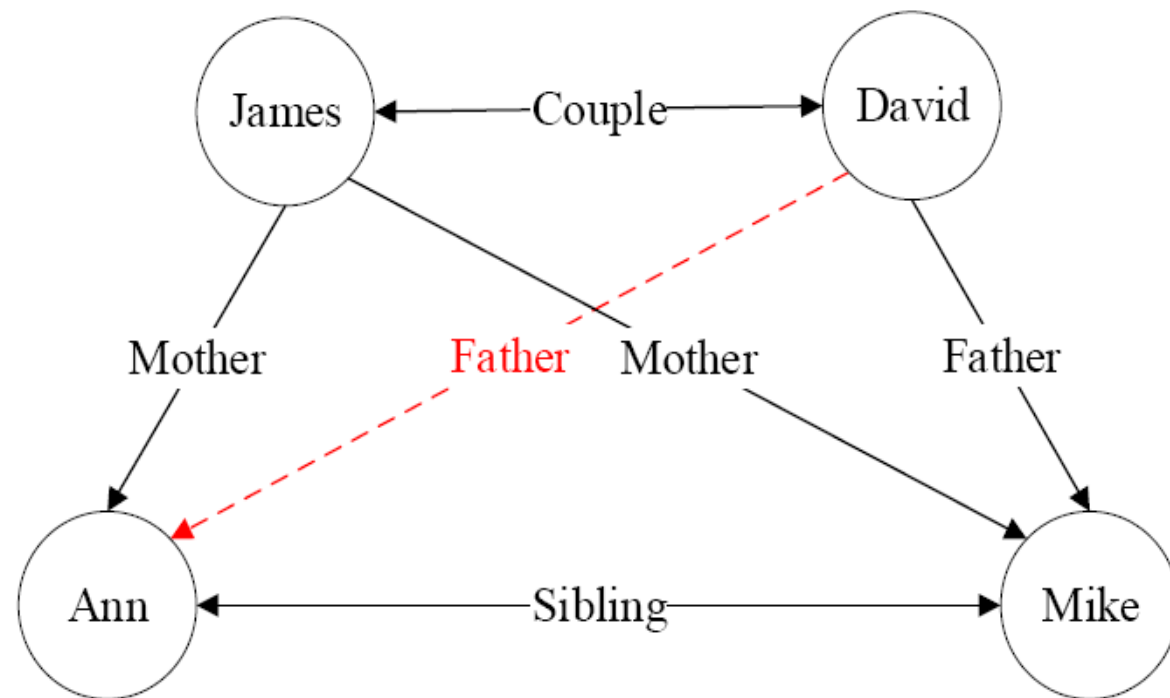
一个简单的家庭关系知识图谱

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 ⟹ hypothesis

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

前提约束谓词
（学习得到）

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：知识图谱中目标谓词以外的其他谓词实例化结果

  - 如$Sibling(Ann, Mike)$
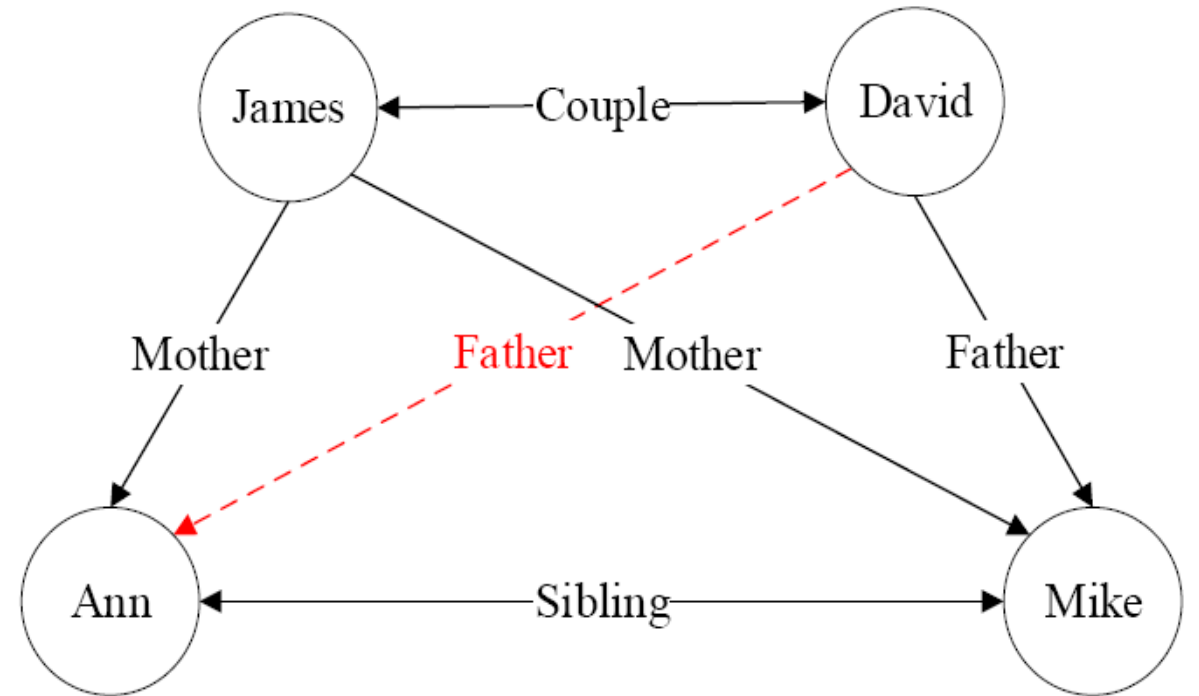
一个简单的家庭关系知识图谱

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

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。**

- **对目标谓词或前提约束谓词中的变量赋予具体值**

  - 如将 $(\forall x)(\forall y)(\forall z)(Mother(z, y) \land Couple(x, z) \rightarrow Father(x, y))$ 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

- 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：

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

  - 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)

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

前提约束谓词（学习得到）　　　　目标谓词（已知）

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

$\Rightarrow$
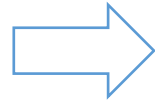
- 依次将谓词加入到推理规则中作为前提约束谓词，并计算所得到新推理规则的FOIL增益值。

- 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)

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

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) Couple(David, James) Mother(James, Ann) Mother(James, Mike)	目标谓词训练样例集合	Father(David, Mike) ¬Father(David, James) ¬Father(James, Ann) ¬Father(James, Mike) ¬Father(Ann, Mike)

- 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	$m_+ = 1$	$m_- = 4$	$FOIL\_Gain$
$Father(x, y) \leftarrow$	$Mother(x, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Mother(x, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Mother(y, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(y, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(z, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(z, y)$	$\widehat{m_+} = 1$	$\widehat{m_-} = 3$	0.32
	$Sibling(x, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Sibling(x, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Sibling(y, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(y, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(z, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(z, y)$	$\widehat{m_+} = 1$	$\widehat{m_-} = 2$	0.74
	$Couple(x, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$\boldsymbol{Couple(x, z)}$	$\boldsymbol{\widehat{m_+} = 1}$	$\boldsymbol{\widehat{m_-} = 1}$	1.32
	$Couple(y, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Couple(y, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Couple(z, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Couple(z, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA

- 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$Father(James, Ann)$是一个反例

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	$m_+ = 1$	$m_- = 4$	$FOIL\_Gain$
$Father(x, y) \leftarrow$	$Mother(x, y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Mother(x, z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Mother(y, x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(y, z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(z, x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(z, y)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 3$	0.32
	$Sibling(x, y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Sibling(x, z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Sibling(y, x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(y, z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(z, x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(z, y)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 2$	0.74
	$Couple(x, y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$\boldsymbol{Couple(x, z)}$	$\boldsymbol{\widehat{m}_+ = 1}$	$\boldsymbol{\widehat{m}_- = 1}$	1.32
	$Couple(y, x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Couple(y, z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Couple(z, x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Couple(z, y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA

- **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$Mother(James, \text{Mike})$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$Father(James, Mike)$是一个反例

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	$m_+ = 1$	$m_- = 4$	$FOIL\_Gain$
$Father(x,y) \leftarrow$	$Mother(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Mother(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Mother(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(z,y)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 3$	0.32
	$Sibling(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Sibling(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Sibling(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(z,y)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 2$	0.74
	$Couple(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$\boldsymbol{Couple(x,z)}$	$\boldsymbol{\widehat{m}_+ = 1}$	$\boldsymbol{\widehat{m}_- = 1}$	1.32
	$Couple(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Couple(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Couple(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Couple(z,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	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 $\widehat{m_+} = 0, \ \widehat{m_-} = 2$

- 由于 $\widehat{m_+} = 0$，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	$m_+ = 1$	$m_- = 4$	$FOIL\_Gain$
$Father(x,y) \leftarrow$	$Mother(x,y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Mother(x,z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Mother(y,x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(y,z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(z,x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(z,y)$	$\widehat{m_+} = 1$	$\widehat{m_-} = 3$	0.32
	$Sibling(x,y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Sibling(x,z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Sibling(y,x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(y,z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(z,x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(z,y)$	$\widehat{m_+} = 1$	$\widehat{m_-} = 2$	0.74
	$Couple(x,y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$\boldsymbol{Couple(x,z)}$	$\boldsymbol{\widehat{m_+} = 1}$	$\boldsymbol{\widehat{m_-} = 1}$	1.32
	$Couple(y,x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Couple(y,z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Couple(z,x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Couple(z,y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA

- **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	$m_+ = 1$	$m_- = 4$	$FOIL\_Gain$
$Father(x,y) \leftarrow$	$Mother(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Mother(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Mother(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Mother(z,y)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 3$	0.32
	$Sibling(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Sibling(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$Sibling(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Sibling(z,y)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 2$	0.74
	$Couple(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$\boldsymbol{Couple(x,z)}$	$\boldsymbol{\widehat{m}_+ = 1}$	$\boldsymbol{\widehat{m}_- = 1}$	1.32
	$Couple(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Couple(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$Couple(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 2$	NA
	$Couple(z,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA

- **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：

$$\widehat{m_+} \cdot \left(\log_2 \frac{\widehat{m_+}}{\widehat{m_+} + \widehat{m_-}} - \log_2 \frac{m_+}{m_+ + m_-} \right)$$

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

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	$m_+ = 1$	$m_- = 4$	$FOIL\_Gain$
$Father(x, y) \leftarrow$	$Mother(x, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Mother(x, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Mother(y, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(y, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(z, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Mother(z, y)$	$\widehat{m_+} = 1$	$\widehat{m_-} = 3$	0.32
	$Sibling(x, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Sibling(x, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$Sibling(y, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(y, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(z, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Sibling(z, y)$	$\widehat{m_+} = 1$	$\widehat{m_-} = 2$	0.74
	$Couple(x, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA
	$\boldsymbol{Couple(x, z)}$	$\boldsymbol{\widehat{m_+} = 1}$	$\boldsymbol{\widehat{m_-} = 1}$	1.32
	$Couple(y, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Couple(y, z)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 0$	NA
	$Couple(z, x)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 2$	NA
	$Couple(z, y)$	$\widehat{m_+} = 0$	$\widehat{m_-} = 1$	NA

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

- 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
$Father(x,y)$ $\leftarrow Couple(x,z)$	$\wedge Mother(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge \boldsymbol{Mother(z,y)}$	$\widehat{\boldsymbol{m}}_+ = \boldsymbol{1}$	$\widehat{\boldsymbol{m}}_- = \boldsymbol{0}$	$\boldsymbol{1}$
	$\wedge Sibling(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(z,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$\wedge Couple(x,z)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 1$	0
	$\wedge Couple(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(z,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA

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

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

已知：
$Mother(James, Ann)$
$Couple(David, James)$

于是：$Father(David, Ann)$

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
$Father(x,y)$ $\leftarrow Couple(x,z)$	$\wedge Mother(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Mother(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge \boldsymbol{Mother(z,y)}$	$\boldsymbol{\widehat{m}_+ = 1}$	$\boldsymbol{\widehat{m}_- = 0}$	**1**
	$\wedge Sibling(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(x,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Sibling(z,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(x,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 1$	NA
	$\wedge Couple(x,z)$	$\widehat{m}_+ = 1$	$\widehat{m}_- = 1$	0
	$\wedge Couple(y,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(y,z)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(z,x)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA
	$\wedge Couple(z,y)$	$\widehat{m}_+ = 0$	$\widehat{m}_- = 0$	NA

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

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	Sibling(Ann, Mike) Couple(David, James) Mother(James, Ann) Mother(James, Mike)	目标 谓词 训练 样例 集合	Father(David, Mike) ¬Father(David, James) ¬Father(James, Ann) ¬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

Score($Father(David, Ann)$)
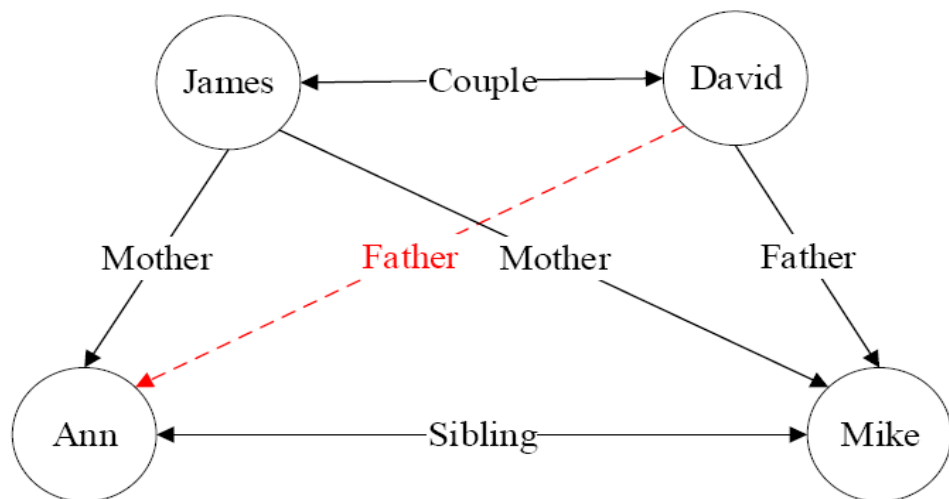
一个简单的家庭关系知识图谱

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

Score($Father(David, Ann)$)
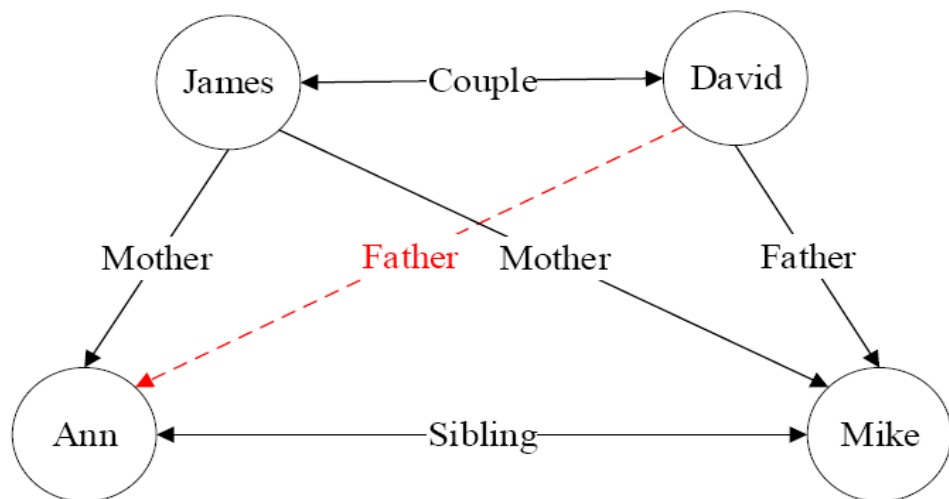
一个简单的家庭关系知识图谱

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

p_l是链接节点 s 和节点t的 所有路径集合

θ_j 是某 一条路径 π_j的权重

P 是路径π_j 概率 值大小

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

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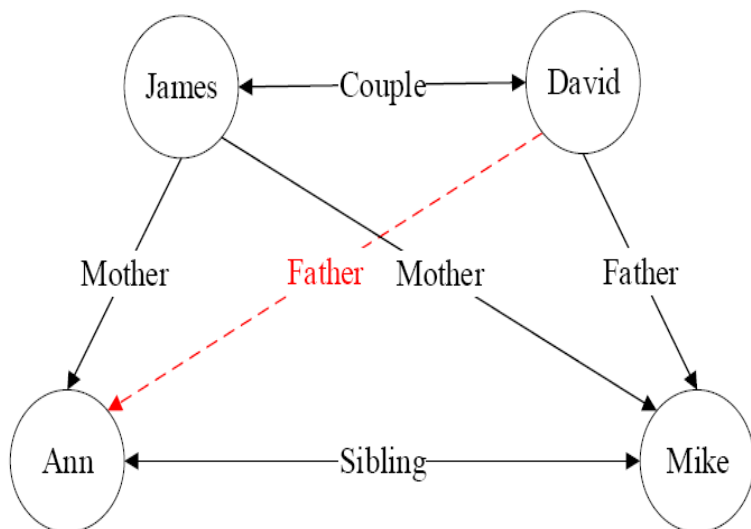
- **特征抽取：** 生成并选择路径特征集合。生成路径的方式有随机游走（random walk）、广度优先搜索、深度优先搜索等。
- **特征计算**：计算每个训练样例的特征值$P(s \rightarrow t; \pi_j)$。该特征值可以表示从实体节点$s$出发，通过关系路径$\pi_j$到达实体节点$t$的概率；也可以表示为布尔值，表示实体$s$到实体$t$之间是否存在路径$\pi_j$；还可以是实体$s$和实体$t$之间路径出现频次、频率等。
- **分类器训练**：根据训练样例的特征值，为目标关系训练分类器。当训练好分类器后，即可将该分类器用于推理两个实体之间是否存在目标关系。

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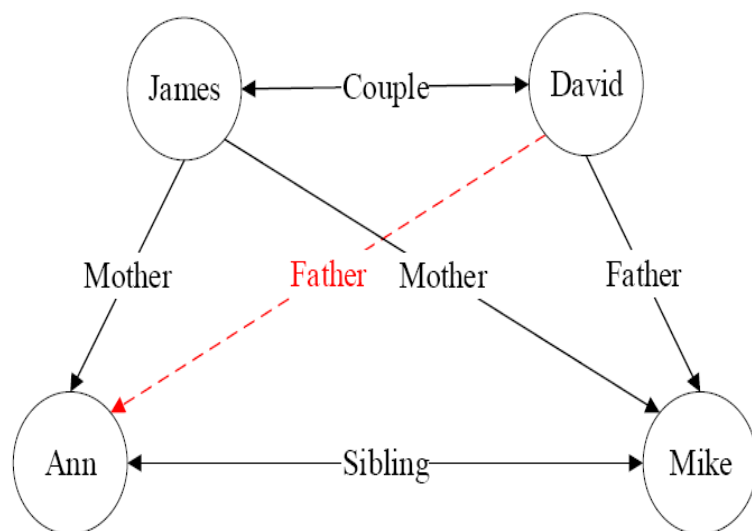
Score(*Father*(*David*, *Ann*)**)**

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* → *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*

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

Score($Father(David, Ann)$)
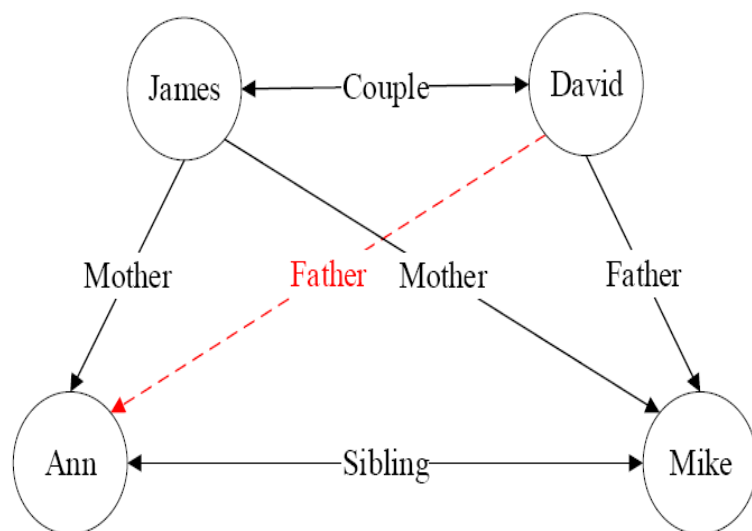
3. 对于每一个正例/负例，判断上述四条路径可否链接其包含的两个实体，将可链接(记为1)和不可链接(记为0)作为特征，于是每一个正例/负例得到一个四维特征向量：

(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

知识图谱推理：机器学习

概念
关系
属性

训练样例

特征提取

定义特征

开放文档
数据集

分类器

模型训练

识别

扩充
知识图谱

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

Wiki中用户对"Bill Gates"这个实例的标注

知识图谱构造流程：以Wiki为例子

- 从 概 念 的 专 业 分 类 (taxonomy) 到 大 众 分 类 (folksonomy)，即用户趋向于用自我定义的标签对内容进行组织和分类。

- 44 个 类 别 标 签：Bill Gates,1955 births, American billionaires, American chairmen of corporations, American computer, business people, 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, Business people in software, Directors of Berkshire Hathaway…..

知识图谱构造流程：以Wiki为例子

- □ 手工构造的种子知识(infobox)
- □ 比尔盖茨这个实体用了13个属性描述、奥巴马这个实体用了20多个属性描述(由于两者均属于persons这个类别，有些属性是共享的)
- □ Freebase中定义的cities这个概念时使用了将近200个属性，但是仍然远远不够。

□ **如何发现和学习新的属性？**

知识图谱构造流程：以Wiki为例子

- 基于机器学习算法进行概念、属性和关系学习，需要大量良好标注数据

- Wikipedia当中，英文文章占总文章数的38.95%，中文文章占4.75%，其他语言文章占56.30%。其中，38.60%的英文文章以Infobox方式被标注 (如实体属性或实体之间的关系)，21.43%的中文文章被标注，平均28.57%的其他语言文章被标注。

用户对wiki中Bill Gates文章所标注的知识

知识图谱构造流程：以Wiki为例子

- 给定无结构化文档数据，通过机器学习方法对实体描述内容进行分类，同时提取描述实体的属性和对应属性值。

嫦娥三号类别： 2013 in China; Space probes launched in 2013; Chinese Lunar Exploration Program; Chinese space probes; Lunar rovers; Missions to the Moon; Spacecraft that orbited the Moon; Soft landings on the Moon; Space probes

机器学习

种子知识

来自InfoBox
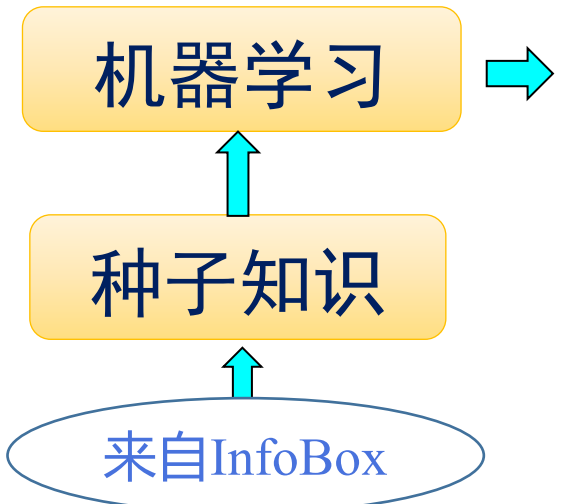
名称	嫦娥三号月球探测器
所属国家	中华人民共和国
构成	着陆器、"玉兔号"月球车
重量	3,750千克
关键技术	7500牛变推力发动机、热控两相流体回路、可变热导热管等

类别分类、实体识别、属性填充

知识图谱构造流程：属性识别与填充

纳尔逊·罗利赫拉赫拉·曼德拉（Nelson Rolihlahla Mandela）1918年7月18日出生于南非特兰斯凯一个大酋长家庭，先后获南非大学文学士和威特沃特斯兰德大学律师资格，当过律师。曼德拉自幼性格刚强，崇敬民族英雄。他是家中长子而被指定为酋长继承人。但他表示："决不愿以酋长身份统治一个受压迫的部族"，而要"以一个战士的名义投身于民族解放事业"。他毅然走上了追求民族解放的道路。

南非政府2013年12月6日（北京时间）宣布南非前总统曼德拉去世，享年95岁。

曼德拉，南非黑人领袖，因其在废除南非种族歧视政策方面作出了巨大贡献而于1993年荣获诺贝尔和平奖。

1918年7月18日，曼德拉出生于南非特兰斯凯，曼德拉自幼性格刚强，崇敬民族英雄。他是家中长子而被指定为酋长继承人。但他表示："决不愿以酋长身份统治一个受压迫的部族"，而要"以一个战士的名义投身于民族解放事业"，他毅然……
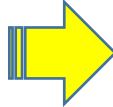
无结构化文档

人物	
中文名	纳尔逊·罗利赫拉赫拉·曼德拉
外文名	Nelson Rolihlahla Mandela
国籍	南非
出生地	南非特兰斯凯
出生日期	1918年7月18日
职业	政治 南非总统
主要成就	诺贝尔和平奖得主

所提取的实体及其属性值形式的知识

知识图谱推理: 机器学习

George Washington

George Washington by Gilbert Stuart, 1797

从数据中学习
人名这个类别
所定义的各个
属性分类器

Thomas Jefferson

3rd President of the United States

infoboxes

Born: February 22, 1732, Westmoreland County, Virginia, United States

Died: December 14, 1799, Mount Vernon, Virginia, United States

Spouse: Martha Washington (m. 1759–1799)

Presidential term: April 30, 1789 – March 4, 1797

Siblings: Lawrence Washington (1718–1752), more

Born:

Died:

Party:

Presidential term:

Children:

知识图谱推理: 机器学习

George Washington

George Washington by Gilbert Stuart, 1797

infoboxes

Born: February 22, 1732, Westmoreland County, Virginia, United States

Died: December 14, 1799, Mount Vernon, Virginia, United States

Spouse: Martha Washington (m. 1759–1799)

Presidential term: April 30, 1789 – March 4, 1797

Siblings: Lawrence Washington (1718–1752), more

训练得到的
属性分类器

Thomas Jefferson

3rd President of the United States

Born: April 13, 1743, Shadwell, Virginia, United States

Died: July 4, 1826, Monticello, Virginia, United States

Party: Democratic-Republican Party

Presidential term: March 4, 1801 – March 4, 1809

Children: Martha Jefferson Randolph, Eston Hemings,

知识图谱：从数据到知识、从知识到决策

提纲
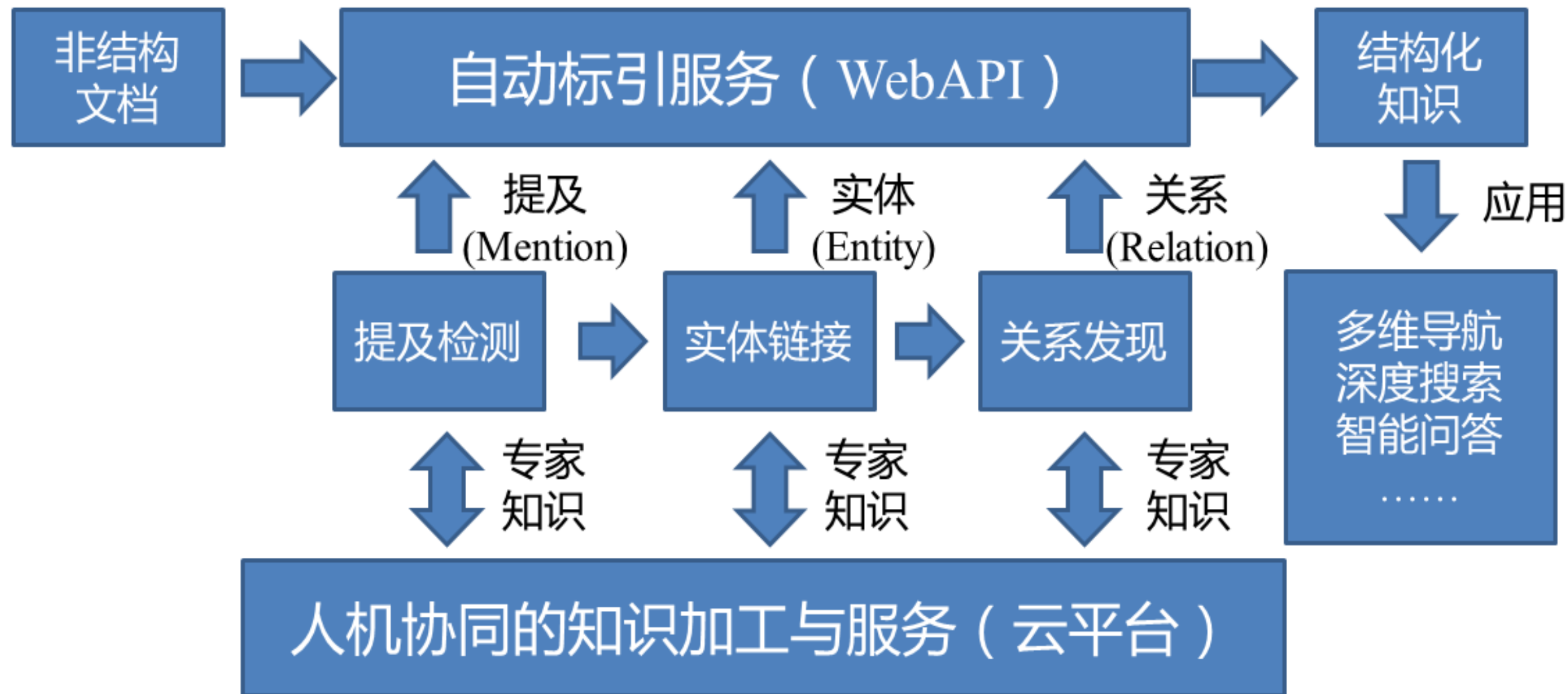
- 命题逻辑

- 谓词逻辑

- 知识图谱推理

- 因果推理

相关不意味着因果

致酒行 （唐-李贺）

零落栖迟一杯酒，主人奉觞客长寿。

主父西游困不归，家人折断门前柳。

吾闻马周昔作新丰客，天荒地老无人识。

空将笺上两行书，直犯龙颜请恩泽。

我有迷魂招不得，**雄鸡一声天下白。**

少年心事当拿云，谁念幽寒坐呜呃。

公鸡打鸣与太阳升起

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

身高(cm)	计算机学院		文学院	
	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

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

$$\Downarrow$$

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

- **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。

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

	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

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

$$\Downarrow$$

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

左：计算机学院和文学院学生的身高情况

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：Causal Inference

- **传统以统计建模为核心的推理手段**
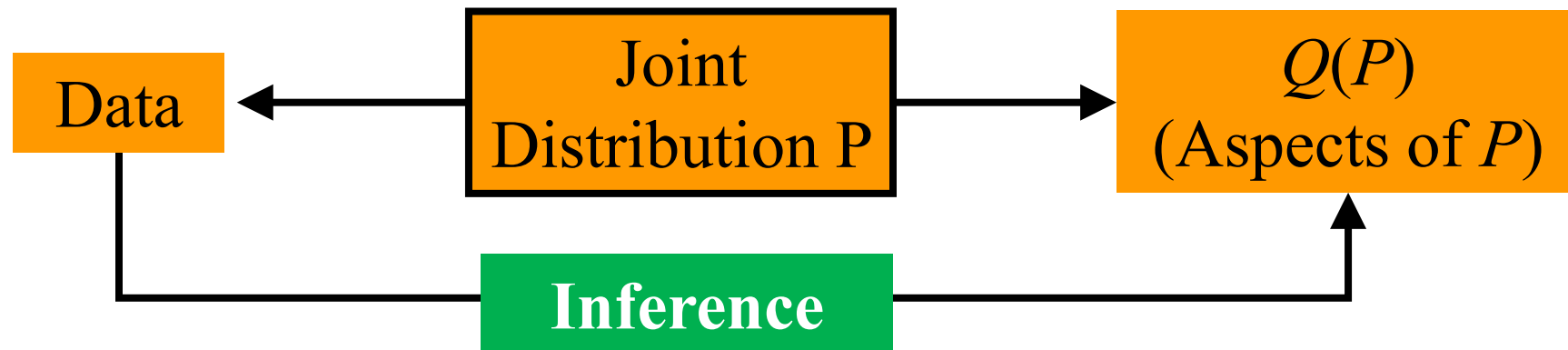
"The object of statistical methods is the reduction of data" (Fisher 1922).

购买了A商品的顾客是否会购买B商品(对A和B的联合分布建模)

$$Q = P(B \mid A)$$

因果推理：Causal Inference

● 如果商品价格涨价一倍，预测销售量P′(sales)的变化

● 如果放弃吸烟，预测癌症P′(cancer) 的概率
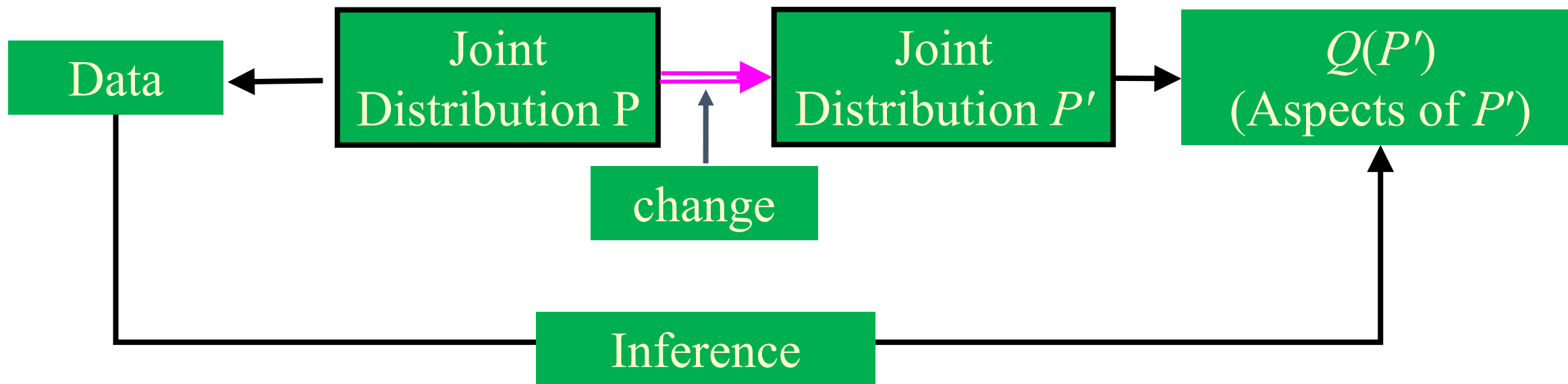
数据分布从 P 变换到 P′

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	What if we see A	$P(y|A)$
决策行动问题	What if we do A	$P(y|do(A))$ （如果采取A行为，则B真）
反事实问题 **(Counterfactual)**	What if we did things differently	$P(y'|A)$ （如果A为真，则B将不同）
Options: with what probability		

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）

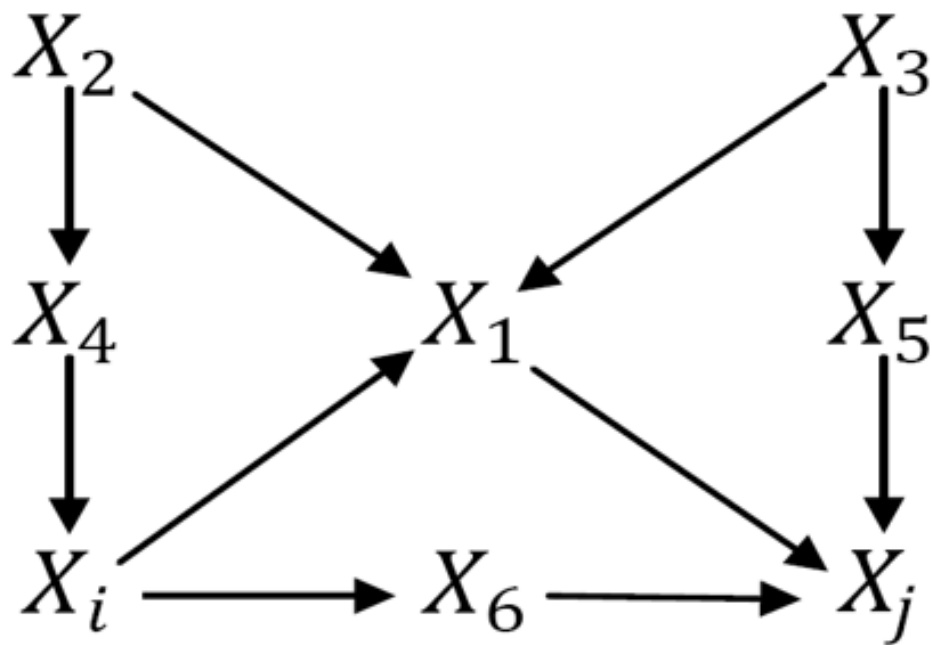- 对于任意的DAG，模型中d个变量的联合概率分布由每个节点与其父节点之间条件概率$P(child|parents)$的乘积给出：

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

- 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

  - 如联分布 $P(X_1, X_2) = P(x_1)P(x_2|x_1) = P(x_2)P(x_1|x_2)$

请写出下列有向无环图的联合概率形式，并区分内生变量和外生变量。

有向无环图DAG

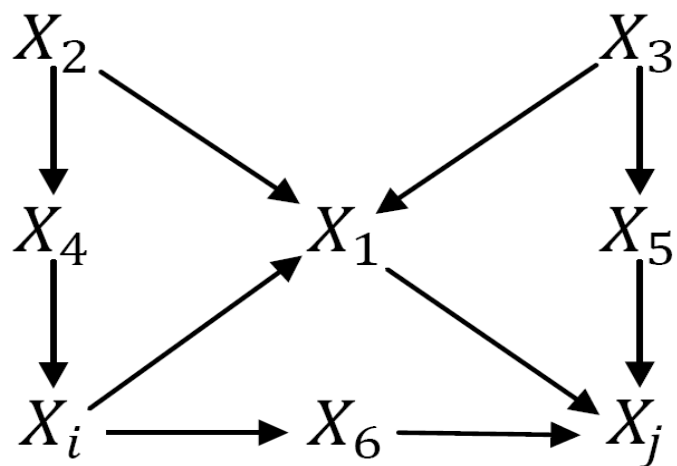正常使用主观题需2.0以上版本雨课堂

作答

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：
  $P($气候好$) \times P($产量高$|$气候好$) \times P($价格低$|$产量高$)$

假设：
$P($气候好$) = 0.5$
$P($产量高$|$气候好$) = 0.8$
$P($价格低$|$产量高$) = 0.9$

$P($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)

A C B
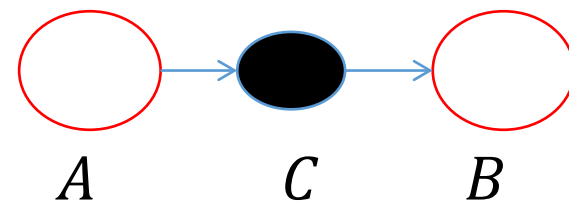
- 当C取值固定(可观测, **observed)**，有

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

- 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

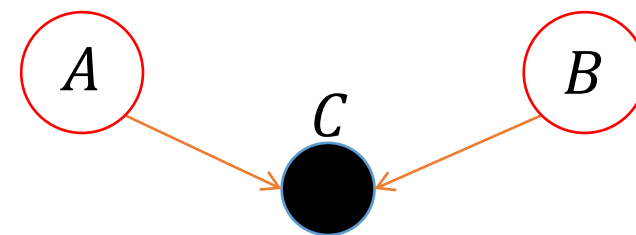- D-分离用于判断集合A中变量是否与集合B中变量相互独立（给定集合C），记为$A \perp B \mid C$

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

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**

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

$= 1$
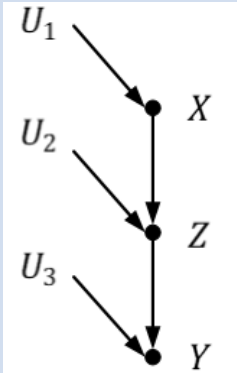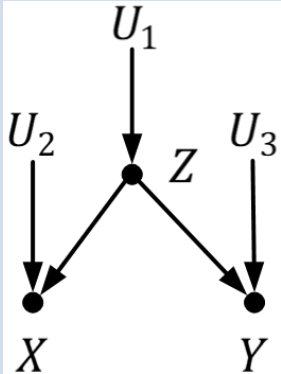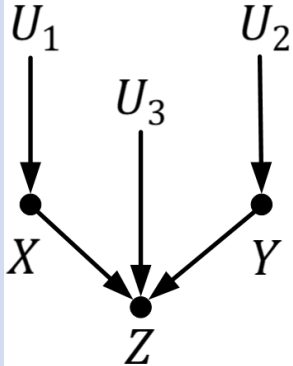
- **A和B在条件C下是独立的**
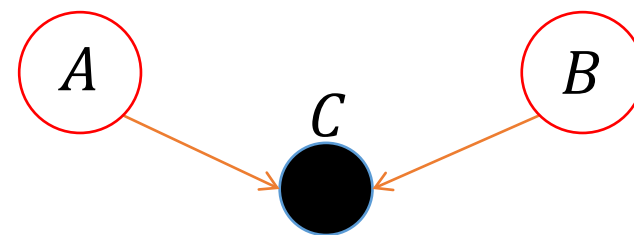
因果推理：*D*-分离

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
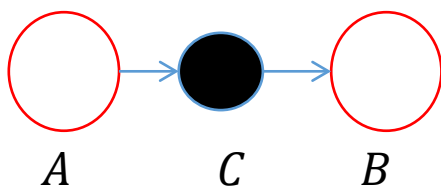
链结构
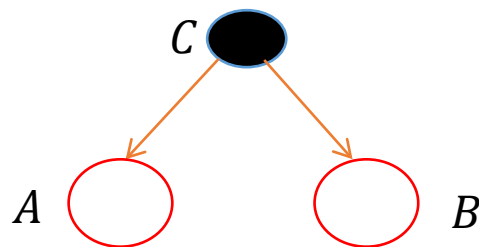
分连结构
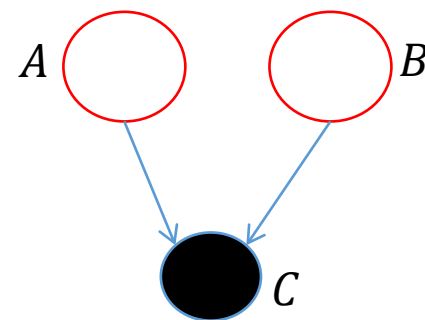
汇连结构
(V结构或冲撞点，collider)

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
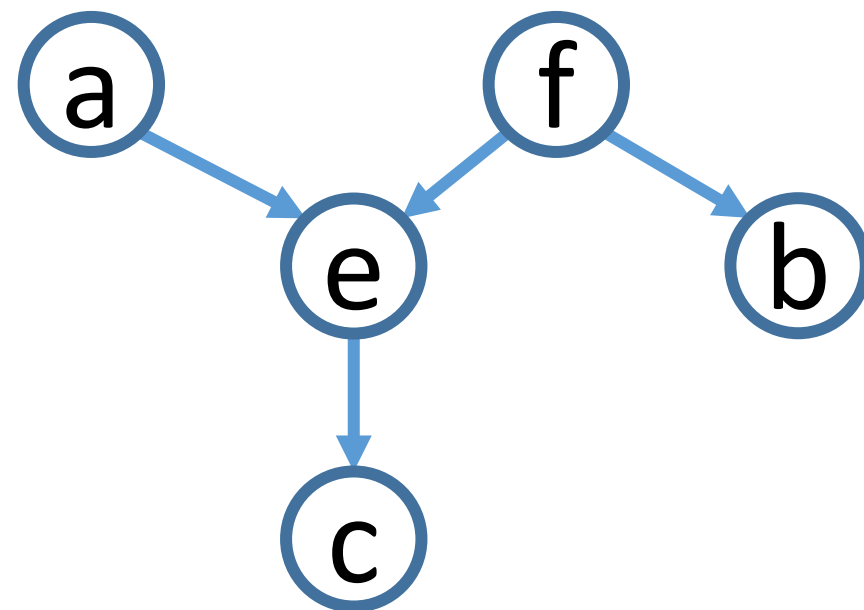
提交

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：

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

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

- **干预**指的是固定系统中某个变量，观察其他变量的变化

- 为了与自然取值进行区分，在对X进行干预时，引入do**算子**
  - 因此，$P(Y=y|X=x)$表示当X取值为x时，$Y=y$的概率；
  - 而$P(Y=y|do(X=x))$表示对X取值进行了干预，固定其取值为x时，$Y=y$的概率。

- $P(Y=y|X=x)$**反映了在取值为**x**的个体**X**上，**Y**的总体分布；**

- $P(Y=y|do(X=x))$**反映的是如果将**X**每一个取值都固定为**x**时，**Y**的总体分布。**

因果推理：操纵图模型和操纵概率

- 对学院变量X进行干预并固定其取值为x时，可将所有指向X的边均移除。

- 因果效应$P(Y = y \mid do(X = x))$等价于引入干预的**操纵图模型** (manipulated model)中条件概率$P_m(Y = y \mid X = x)$
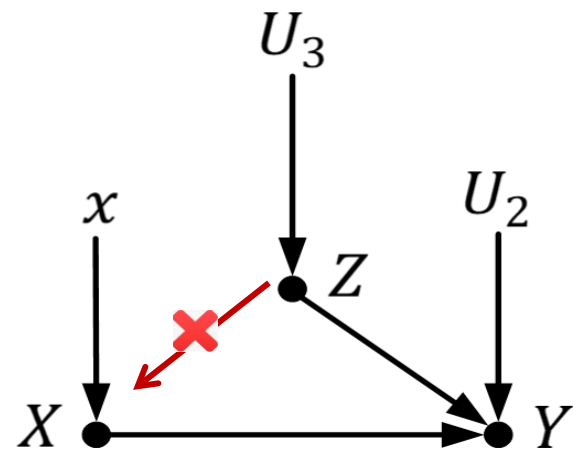
对X干预后操纵图模型

因果推理：操纵图模型和操纵概率

- 计算因果效应关键在于计算操纵概率(manipulated probability) P_m。P_m与无干预条件下概率P在如下两个方面的取值不变：

  - 边缘概率$P(Z = z)$不随干预而变化，因为Z的取值不会因为去掉从Z到X的箭头而变化，即：

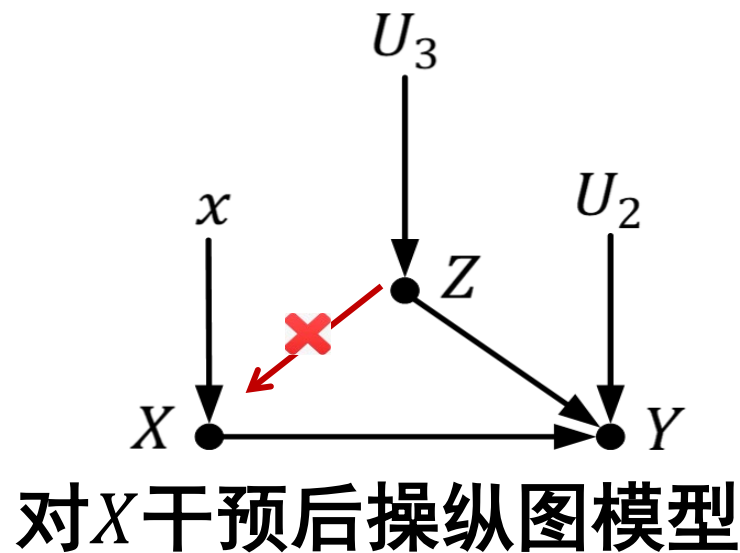 $$P_m (Z = z) = P(Z = z)$$

  - 条件概率$P(Y = y \mid X = x, Z = z)$不变因为$Y$关于$X$和$Z$的函数$f_Y = (X, Z)$并未改变，即：

  $$P_m (Y = y \mid X = x, Z = z) = P(Y = y \mid X = x, Z = z)$$

对X干预后操纵图模型

因果推理：调整公式

- 在干预图中，X和Z是D-分离的，因此是彼此独立的，即：

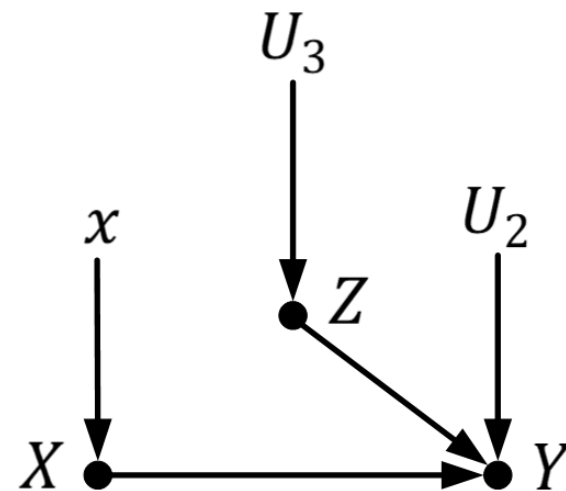$$P_m(Z = z | X = x) = P_m(Z = z) = P(Z = z)$$

- 因果效应$P(Y = y | do(X = x))$有：

$$P(Y = y | do(X = x)) = P_m(Y = y | X = x)$$

$$= \sum_z P_m(Y = y | X = x, Z = z)P_m(Z = z | X = x)$$

$$= \sum_z P_m(Y = y | X = x, Z = z)P_m(Z = z)$$

$$= \sum_z P(Y = y | X = x, Z = z)P(Z = z)$$

对X干预后操纵图模型

X和Z相对于Y构成了V结构

因果推理：调整公式

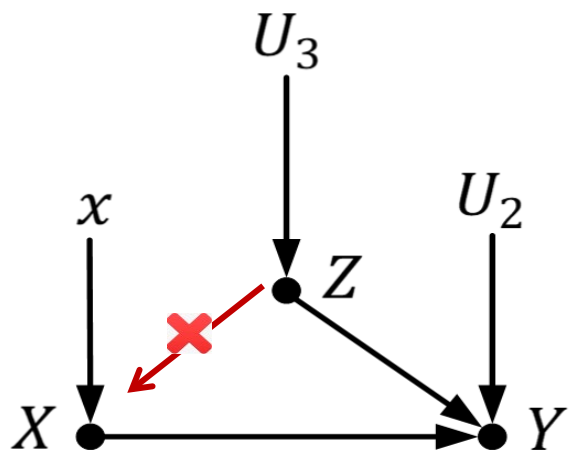- 因果效应$P(Y = y|do(X = x))$有：

$$P(Y = y|do(X = x)) = \sum_{z} P(Y = y|X = x, Z = z)P(Z = z)$$

- **该式被称为调整公式（adjustment formula）**

  - 对于Z的每一个取值z，该式计算X和Y的条件概率并取均值，这个过程称为"Z调整"(adjusting for Z) 或"Z控制"(controlling for Z)

- **该式的右端只包含正常(无干预)条件下的概率P**

  - 即可用正常(无干预)条件下的条件概率来计算干预后的条件概率。

因果推理：调整公式

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

以性别分组后的计算机学院和文学院的学生身高情况

对X干预后操纵图模型
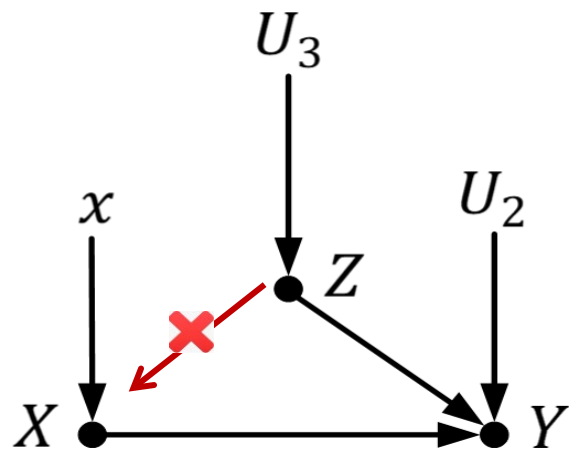
$$P(Y = y|do(X = x))$$

$$= \sum_z P(Y = y|X = x, Z = z)P(Z = z)$$

下面将调整公式用于计算对X取值进行干预后计算机学院/文学院高个子率，其中$X = 1$表示计算机学院，$Y = 1$表示高个，$Z = 1$表示表示男生，则有：

$$P(Y = 1|do(X = 1))$$
$$= P(Y = 1|X = 1, Z = 1)P(Z = 1)$$
$$+ P(Y = 1|X = 1, Z = 0)P(Z = 0)$$

因果推理：调整公式

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

以性别分组后的计算机学院和文学院的学生身高情况

对X干预后操纵图模型

X干预取值为计算机学院的因果效应为：

$$P(Y = 1|do(X = 1))$$
$$= 0.87 \times \frac{(35 + 235 + 10 + 80)}{(350 + 350)} + 0.689 \times \frac{(25 + 55 + 70 + 190)}{(350 + 350)}$$
$$= 0.782$$

X干预取值为文学院的因果效应为：

$$P(Y = 1|do(X = 0))$$
$$= 0.889 \times \frac{(35 + 235 + 10 + 80)}{(350 + 350)} + 0.731 \times \frac{(25 + 55 + 70 + 190)}{(350 + 350)}$$
$$= 0.812$$

其因果效应差：

$$ACE = P(Y = 1|do(X = 1)) - P(Y = 1|do(X = 0))$$
$$= 0.782 - 0.812 = -0.03$$

可见计算机学院的高个率不如文学院的高个率。即学院这一变量影响了高个率。

因果推理：因果图的不足

- **实际中难以得到一个完整的DAG**
  - 用于阐述变量之间的因果关系或者数据生成机制，使得 DAG 的应用受到的巨大的阻碍
  - 从观测数据学习 DAG 的结构，是充满挑战的问题。

- **Pearl 引入 *do* 算子是因果推断领域最主要贡献。**
  - 从系统之外人为控制某些变量。但是，这依赖于一个假定：干预某些变量并不会引起 DAG 中其他结构的变化。

- **DAG 作为一种简化的模型，在复杂系统中可能不完全适用**
  - 需要将其拓展到动态系统（如时间序列），还有待研究。

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。

推理总结

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。

谢谢!