

为什么 - 关于因果关系的新科学

Judea Pearl



分享人：白楚

□ 本书三位一体的使命

- 用**非数学**的语言阐述因果革命的**知识内涵**，说明它将怎样影响我们的生活和未来。
- 分享在解决重要的因果问题时，讲述前辈科学家走过的**英勇征程**，**无论成败**。
- 回溯因果革命在人工智能领域的发源地，介绍**如何开发出用我们的母语——因果语言进行交流的机器人**。

□ 本书三位一体的使命

- 用非数学的语言阐述因果革命的知识内涵，说明它将怎样影响我们的生活和未来。
- 分享在解决重要的因果问题时，讲述前辈科学家走过的英勇征程，无论成败。
- 回溯因果革命在人工智能领域的发源地，介绍如何开发出用我们的母语——因果语言进行交流的机器人。

□ 新一代机器人应该能够**表达和解释因果**

- 解释：事情**为何发生**
- 解释：**为何选择**以某种反应
- 解释：大自然**为何遵循**这样的运作方式
-
- 解释：人类思维**为什么**以这样的方式运行
- 解释**理性-感性的联结**：原因和结果、信任和遗憾、目的和责任究竟意味着什么

引言：思维胜于数据

引言

第一章：因果关系之梯

第十章：大数据、
人工智能和大问题

06

01

为什么

第二章：因果推断的起源
第三章：从证据到因
第四章：混杂与去混杂

04

02

第七章：超越统计调整-干预
第八章：探索假如的世界-反事实
第九章：寻找隐藏作用机制-中介

03

第五章：烟雾缭绕的争论
第六章：大量的悖论

引言：思维胜于数据

引言

第一章：因果关系之梯

第十章：大数据、
人工智能和大问题

06

01

为什么

第二章：因果推断的起源
第三章：从证据到因
第四章：混杂与去混杂

04

02

第七章：超越统计调整-干预
第八章：探索假如的世界-反事实
第九章：寻找隐藏作用机制-中介

03

第五章：烟雾缭绕的争论
第六章：大量的悖论

□ 新科学

■ 解决以下看似简单明了的问题

- 王婷老师的经常性提醒是否**有效**地促成了今天PPT的制作？
- **如果**王老师继续**执行**提醒操作，下次PPT能够更快完成吗？
- **假如**我昨天没有加班做完PPT，王老师今天会把我怎么样？

□ 老阻碍

- 提出因果问题的**词汇**和交流科学理论的**词汇**之间的鸿沟
- 传统数学公式可以表示关系，但是很难说清楚里面的因果
 - 加速度a与力F之间的关系： $F=ma$ ， $a=F/m$

□ 统计科学的不足

- 统计学盲目迷恋常识性的观察结论
- 没有告诉我们因果关系是什么

□ 数据不是万能的

- 数据可以告诉你服药的病人比不服药的病人康复得更快
- 数据却不能告诉你原因何在
- 影响效果的因素可能不是药物

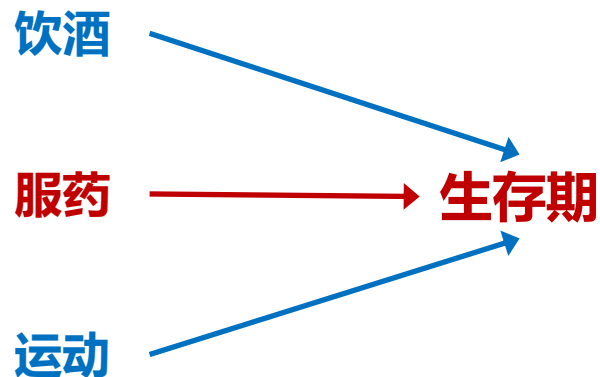
□ 因果关系演算法

□ 因果图

- 点：变量
- 箭头：变量之间已知或疑似存在的**因果关系**
- 示例：药物**D**对病人生存期(L)的影响

□ 符号式语言

药物**D**对病人生存期(L)的影响，记为 $P = (L | do(D))$
表示：如果病人服用了这种药,那么他在**L**年内存活**的概率P**
干预算子do(D)确保生存期**L**的变化完全归因于药物**D**

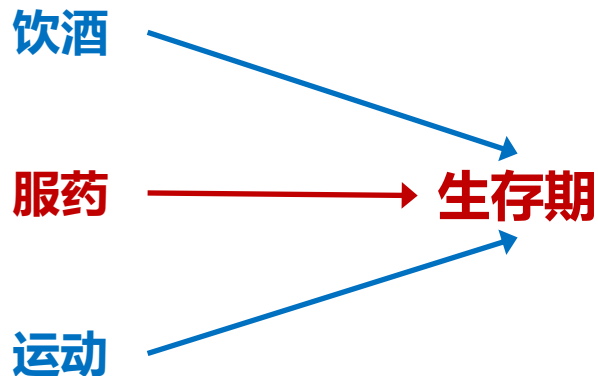


$P = (L | do(D))$
对比 $P = (L | D)$

乔在服用了药物一个月后死亡。

假如乔选择了放弃服药，他现在会活着吗？

What-ifs



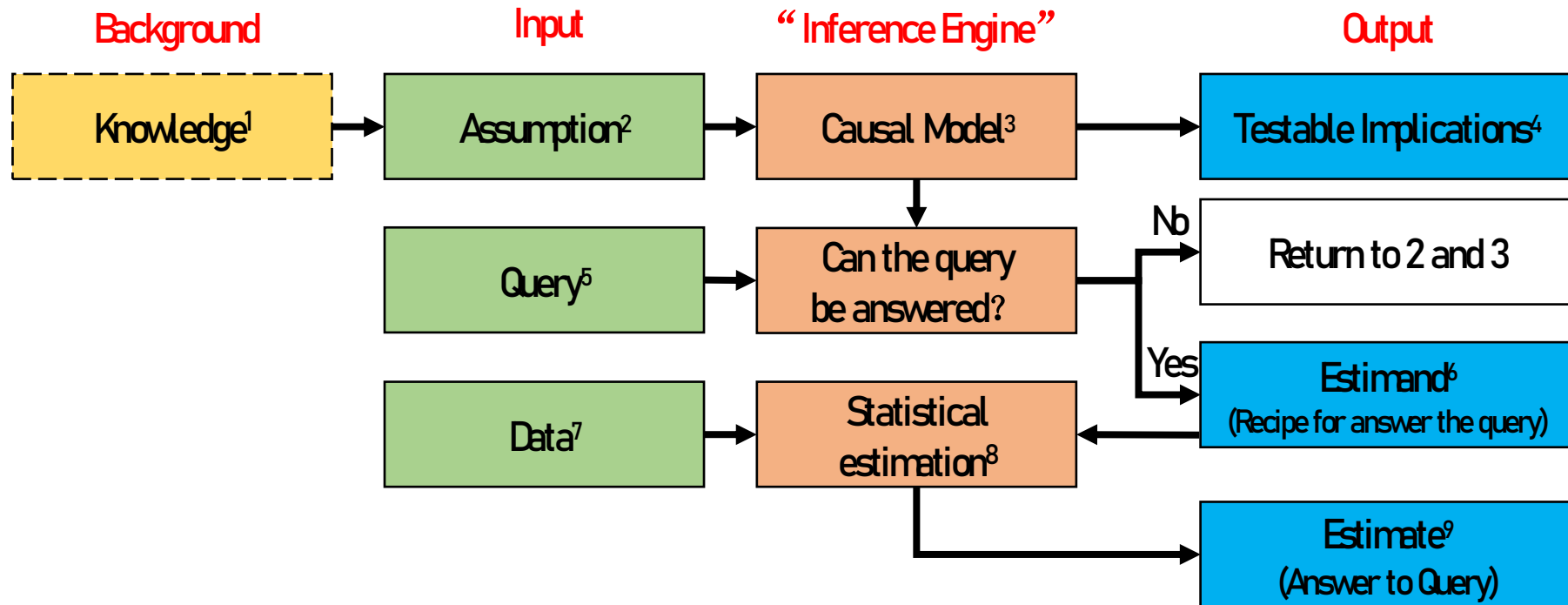
□ 经典统计学

- 关注**总结数据**，无法提供一种语言供我们提出上述问题

□ 反事实(Counterfactual)

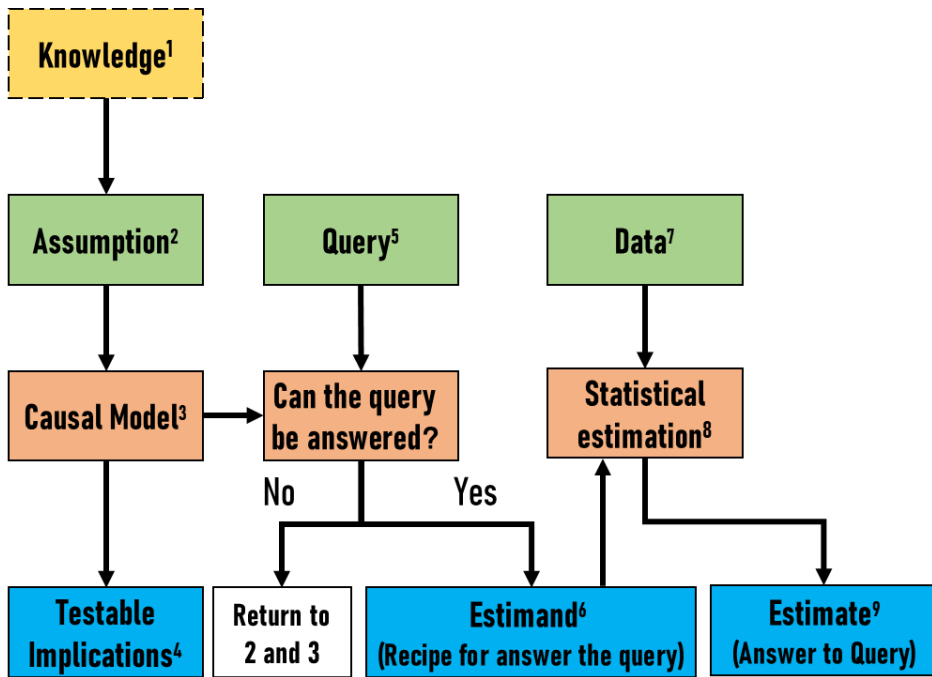
- 问题涉及反思性的思考——因果推断科学所独有的
- 因果推断则不仅提供了一种**表达符号系统**，还提供了一种**解决方案**

因果推断引擎

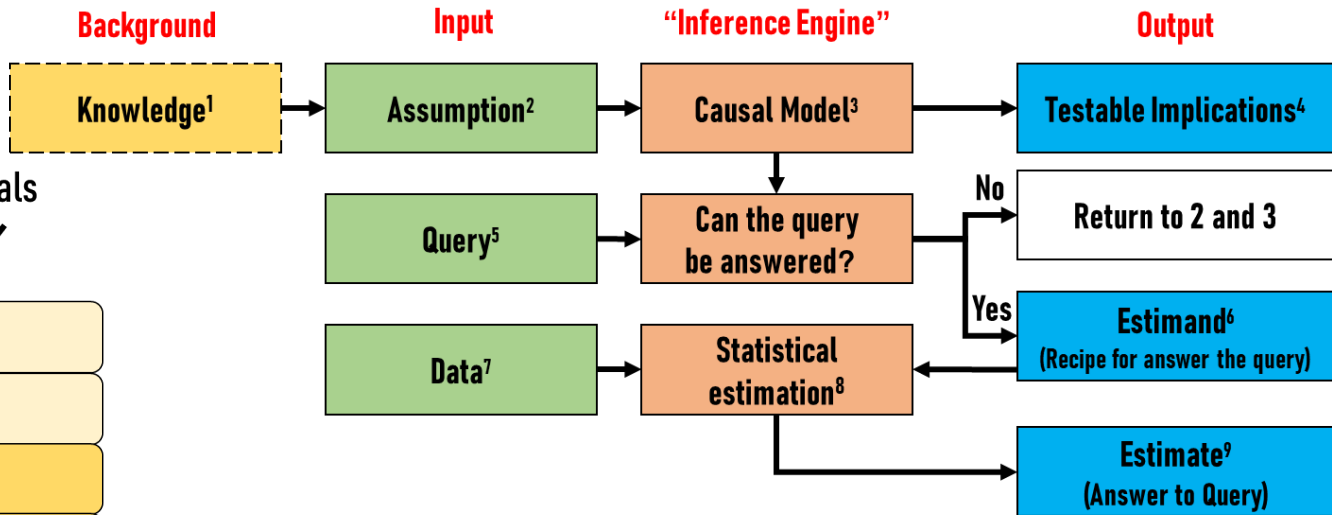
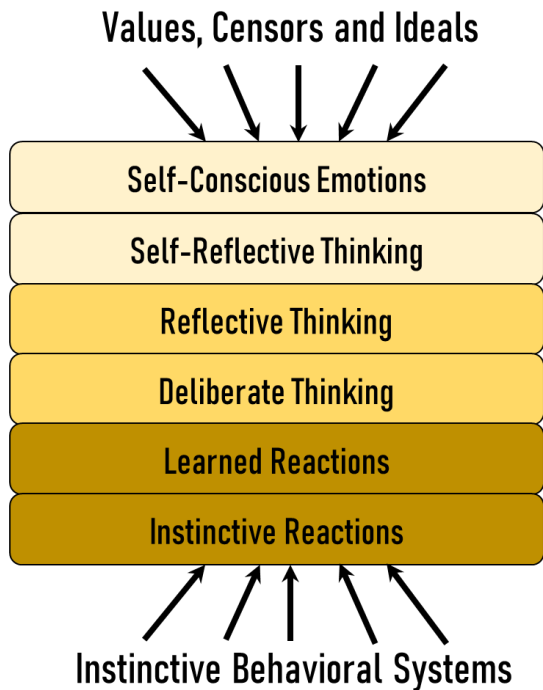


现实的蓝图

- 1-**知识**：推理主体过去的经验
- 2-**假设**：要明确表述出来的陈述
- 3-**因果模型**：因果图、结构方程、逻辑语句
- 4-**可验证的蕴含**：数据中某种显而易见的模式或相关关系
- 5-**问题**：希望解答的问题，例如 $P(L|\text{do}(D))$
- 6-**被估量**：从数据中估算出来的统计量，用于合理地表示问题的答案。但实际上也是一种回答因果问题的方法， $P(L|D,Z) \times P(Z)$
- 7-**数据**：填充被估量的原料，提供数量信息
- 8-**统计估计**：最大似然估计、倾向评分、置信区间、显著性检验等
- 9-**估计值**：待解决的因果问题的答案



异曲同工?



The Book of Why: The New Science of Cause and Effect

The Emotion Machine: Commonsense thinking, artificial intelligence, and the future of the human mind

导言：思维胜于数据

导言

第一章：因果关系之梯

第十章：大数据、
人工智能和大问题

06

01

为什么

第二章：因果推断的起源
第三章：从证据到因
第四章：混杂与去混杂

04

02

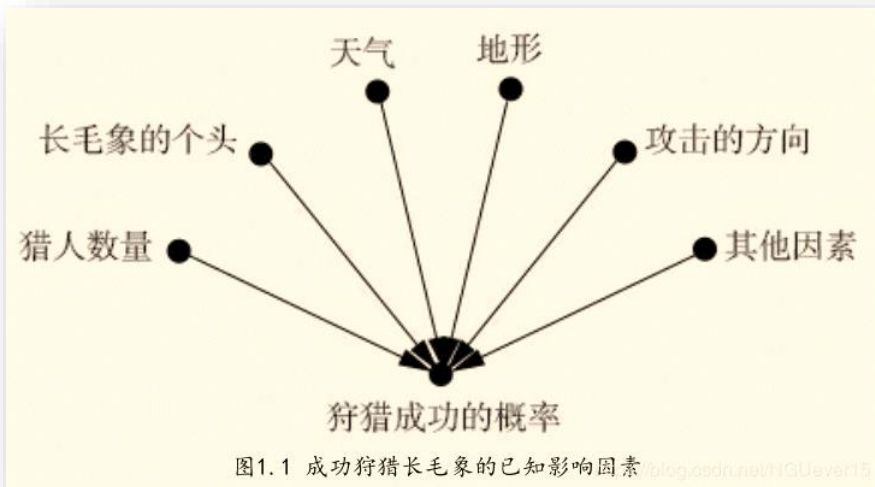
第七章：超越统计调整-干预
第八章：探索假如的世界-反事实
第九章：寻找隐藏作用机制-中介

03

第五章：烟雾缭绕的争论
第六章：大量的悖论

□ 因果想象力和因果图

- 是“**规划**”的基础
- **思维主体必须具备心理模型**
 - 可供参考
 - 可以**自主调整**
 - 是施展**想象**的“舞台”
 - 能够通过对模型局部的修改来**试验不同的情境**
- **因果图是心理模型的一种建构方式**
- **因果图是因果推理引擎的计算核心**



□ 因果关系的三个层级

■ 观察能力 (seeing)

- 发现环境中的规律

■ 行动能力 (doing)

- 预测对环境进行改变后的结果
- 并根据预测结果选择行为方案以催生出自己期待的结果

■ 想象能力 (imagining)

- 工具理论：为什么这种工具有效，无效该怎么做



□ 因果关系之梯——关联

- 许多动物和早期人类所共有，即通过**被动观察**发现环境中的规律
- 观察到事件A改变了观察到事件B的可能性，基于被动观察结果做出**预测**
- 当今人工智能在因果关系之梯的**最底层**
 - **深度学习**为数据拟合函数增加了复杂的层次
 - 深度学习具有了高超的**拟合能力**，而非智能
 - 当今的AI缺少**现实模型**

“购买牙膏的顾客同时购买牙线的可能性有多大？”
 $P(\text{牙线}|\text{牙膏})$



因果关系的梯——干预

- 少数物种表现这种能力，能**预测**对环境进行刻意改变后的结果
- 根据**预测结果**选择行动方案，以催生自己期待的结果
 - 预测干预结果的一种非常直接的方法是在**严格控制**的条件下进行实验
 - 即使没有实验，一个强大和准确的**因果模型**可以利用**第一层级(关联)**的数据来回答**第二层级(干预)**的问题

如果对牙膏另行定价，那么在某一价位销售牙线的概率是多少？

$$P(\text{牙线} | \text{do}(\text{牙膏}))$$



□ 因果关系之梯——反事实

- 目前只有人类表达这种能力，能通过**想象**回到过去，**改变历史的事实**
- 反事实与数据之间存在天然的矛盾
 - 数据是**发生的事实**，因此无法描述**虚构世界**里发生了什么
 - 但人类思维能够通过想象，可靠、重复地推断“**假如…会怎样**”这类假设性问题
 - 人类得以从**历史和他人的经历**中获取经验教训，超越其它物种，奠定了哲学、科学和技术创新的雏形——源自**想象**

“假如我们把牙膏的价格提高一倍，则之前买了牙膏的顾客会**怎样选择**？”

□ 迷你图灵测试

- 选择一个简单的**故事**
- 用某种方式将其**编码**并输入机器
- 测试机器能否正确回答人类能够回答的与故事相关的**因果问题**

■ 问题的表示

- “表示第一，获取第二”的范式
- 在寻求好的表示方法的过程中，关于如何获取知识的洞见就会自然产生
- 人类大脑肯定拥有某种**简洁的信息表示方式**

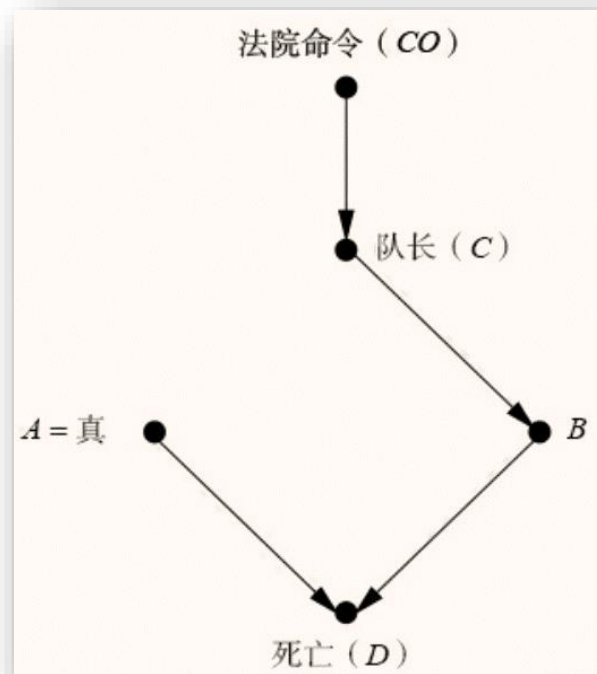
因果图

■ 关联

- 例如：如果A射击了，则B也射击

■ 干预

- 打破规则：**删除**指向某个事件的所有箭头
- 将该事件从所有其他影响因子中**解放**出来
- 例如：如果A自行决定射击，从C到A的箭头被删除，并且A被赋值为真，那么D发生，因此不能推断B是否射击



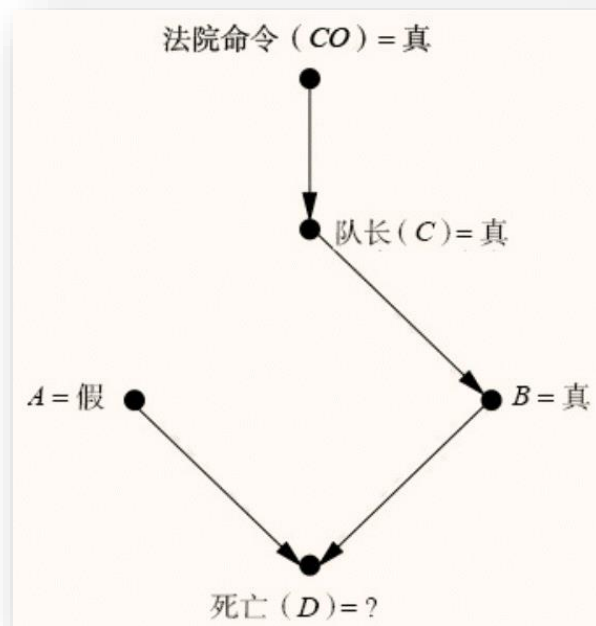
□ 因果图

■ 反事实

- 假设犯人已经倒地身亡
- 假如A没有射击，犯人会死吗？
- B会开枪，犯人依然会死

□ 超越简化模型

- 简化模型里排除了不确定因素和异常事件
- 世界是充满不确定——引入概率的重要性



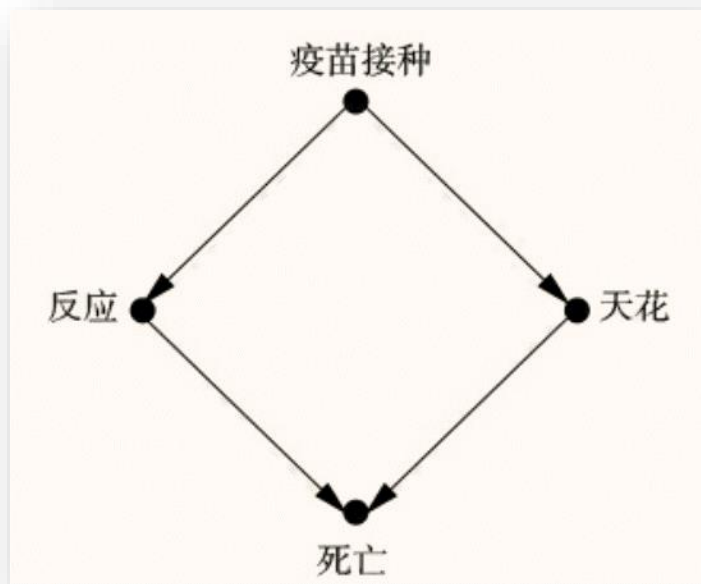
□ 概率的重要性——辩证的看

■ 举例：疫苗接种

- 100万个孩子，99万人接种疫苗
 - 死于疫苗不良反应的99人（不良反应致死1%）
 - 死于天花的40人（患天花概率2%，致死率20%）
- **反事实：假如不接种疫苗，会怎样？**
 - 接种率设为0
 - 2万人罹患天花，死亡4000人

■ 构建因果模型**不仅仅是画箭头**

- **箭头背后还隐藏着概率**



但是，**概率提高准则不能超越因果关系之梯的第一级，因果图的结构不能用概率替代因果图的结构（信息表示方式）**能让我们推测各种因果关系和反事实关系简单的或复杂的，确定的或概率的，线性的或非线性的（**信息的获取**）

引言：思维胜于数据

引言

第一章：因果关系之梯

第十章：大数据、
人工智能和大问题

06

01

为什么

第二章：因果推断的起源
第三章：从证据到因
第四章：混杂与去混杂

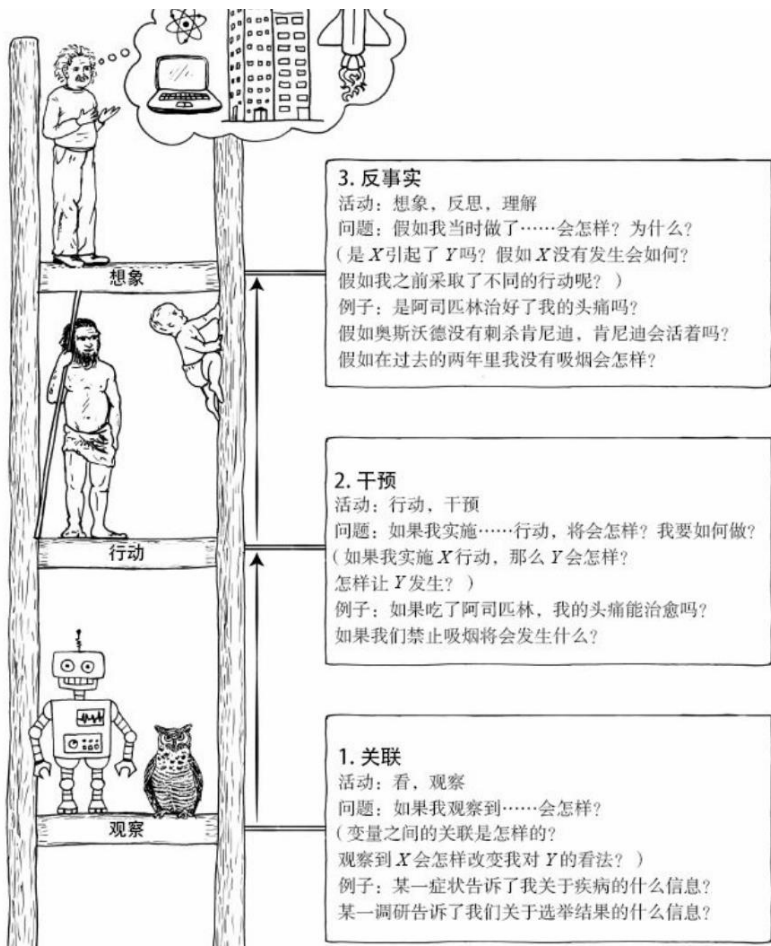
04

02

第七章：超越统计调整-干预
第八章：探索假如的世界-反事实
第九章：寻找隐藏作用机制-中介

03

第五章：烟雾缭绕的争论
第六章：大量的悖论



□ 从关联-到干预巅峰需要迈过的台阶

■ 第二章：因果推断的起源

- 遗传学家休厄尔·赖特于上世纪20年代绘制了世界上第一张因果图

■ 第三章：从证据到因

- 作者如何通过对人工智能的研究，特别是对贝叶斯网络的研究，皈依了因果论

■ 第四章：混杂与去混杂

- 统计学对因果推断的主要贡献：随机对照试验
- 随机对照试验是一个进行**干预的工具**，用以解答 $P(L|\text{do}(D))$ 问题
- 提出一种**去混杂**的极简方法（10分钟内学会）

□ 从关联-到干预巅峰需要迈过的台阶

■ 第四章：混杂与去混杂

- **随机对照试验**：因果关系引擎的特例
- **后门调整**：采集阻断所有**后门**路径变量的数据

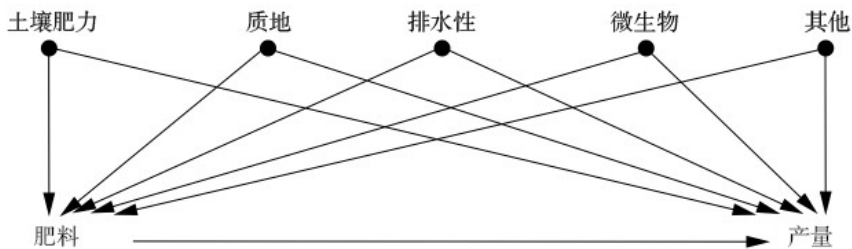


图4.4 模型1：一个错误的对照试验

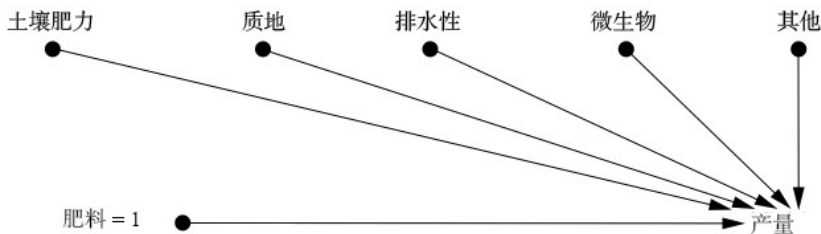


图4.5 模型2：我们想知道的世界

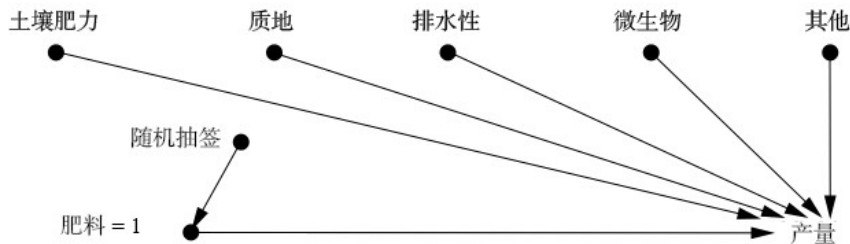
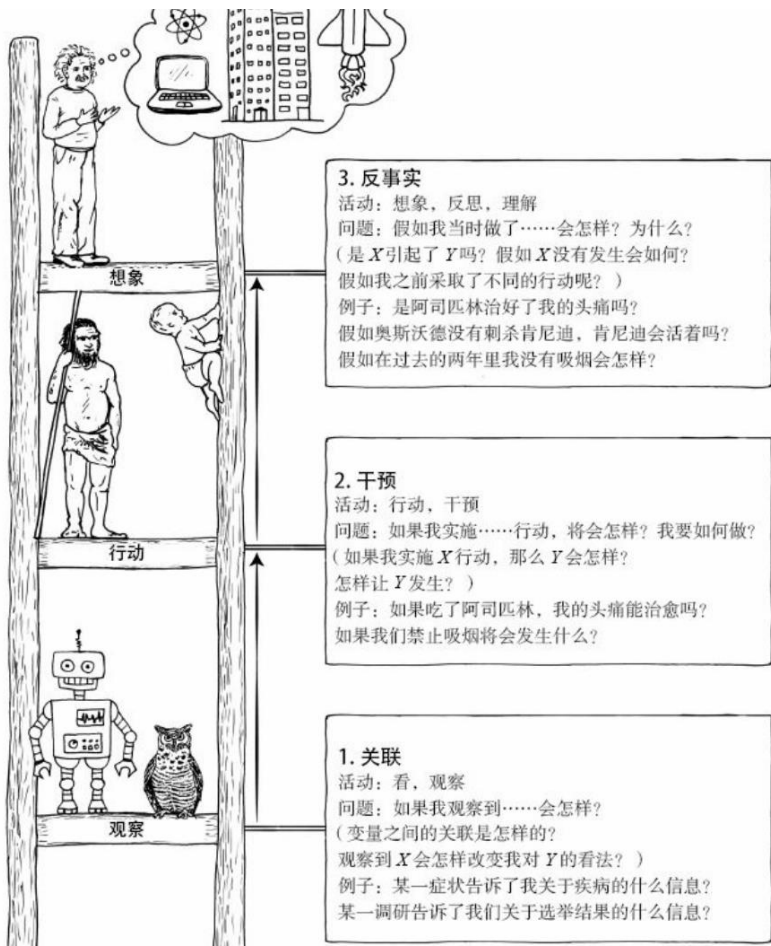


图4.6 模型3：由随机对照试验模拟的世界。



□ 从关联-到干预巅峰需要迈过的台阶

■ 第四章：混杂与去混杂

- 讨论**随机对照试验**作为一种**干预工具**，如何解答 $P(L|\text{do}(D))$ 问题
- 并提出一种**去混杂**的极简方法

■ 第五章：烟雾缭绕的争论

- 上世纪50-60年代，因果论发展史的重要时刻：“吸烟是否会导致肺癌”问题对统计学发起冲击。由于无法使用随机对照试验，人们在是否接受某方结论上始终难以达成一致。

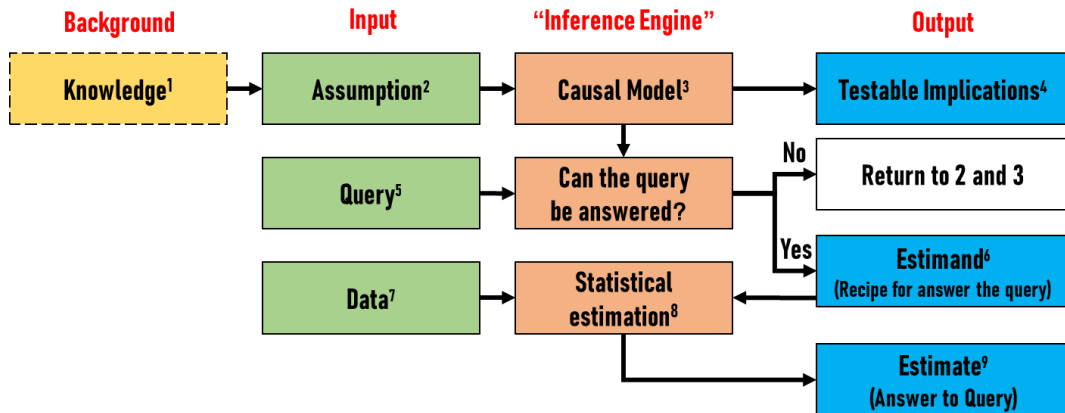
■ 第六章：大量的悖论

- **反因果直觉**的悖论及其内在构造
- **蒙提·霍尔悖论**、**辛普森悖论**、**伯克森悖论**等

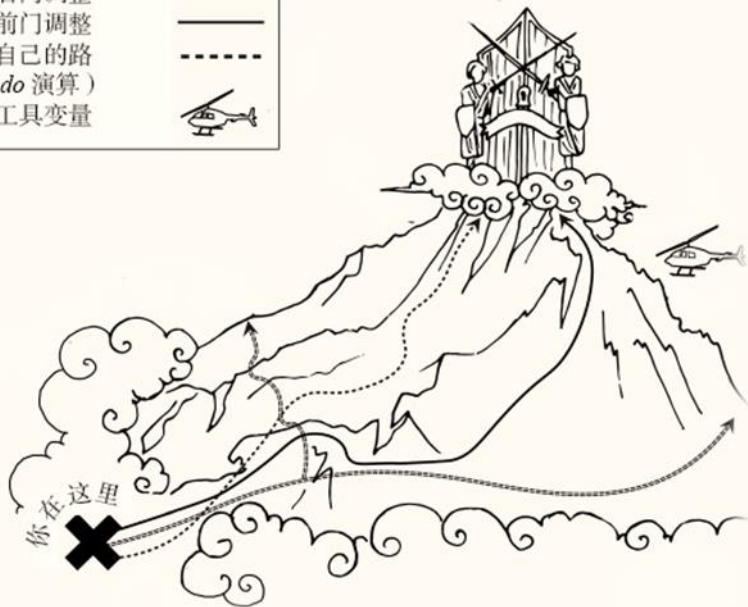
□ 从关联-到干预巅峰需要迈过的台阶

■ 第七章：超越统计调整，征服干预之峰

- 实现do类型问题解答的**自动化**
- 梳理在科学实践中进行**因果推断的主要工具**
- 深入分析“**因果推断引擎**”的本质，如何能够产生“是/否”问题的答案及被估量



后门调整
前门调整
走自己的路
(do 演算)
工具变量



人类具有设想那个不存在的世界的能力，正是这一能力将我们与类人猿祖先以及地球上的其他生物区分开来

□ 反事实推理的古老故事

在地震频繁发生的同时，在埃维亚岛的奥罗比亚地区，爆发了巨大海啸，古希腊史学家修昔底德在《伯罗奔尼撒战争史》中记载公元前426年的一场海啸，有这样的描述：假如没有地震，那我无从理解这种灾难是如何发生的。

亚伯拉罕走上前来，说：无论善恶，您都要剿灭吗？假若那城里有50个义人，您还要剿灭那地方，而不为城里那50个义人饶恕那地方吗？上帝说：我若在所多玛城里见有50义人，我就为他们的缘故饶恕那地方的众人。《创世纪》

□ 亚里士多德构建了一个因果关系分类

□ 包括质料因、形式因、动力因、目的因，未提及反事实的因果关系

因此我们记得曾见过我们称之为‘火焰’的事物，记得曾感受过我们称之为‘热’的事物。我们还会回想起在过去所有的经历中它们的恒常联结。就这样，我们称一个为因，另一个为果，并从一个的存在推断出另一个的存在。”

□ 大卫·休谟对因果关系的第一种定义——“规律性”

- 倾向于将因果关系仅仅视为一种**相关关系**
- 认为第一层级的关联就是我们需要的全部
- 切断了因果关系之梯第一层级与第二、三层级的联系

□ 大卫·休谟对因果关系的第二种定义——“反事实”

- 使用了反事实语言，因此它位于第三层级
- **两种定义出自同一个人，最终却演变为了两种互不相容的思想**

我们可以给一个因下定义说，它先行于、接近于另一个对象的一个对象，而且在这里，凡与前一个对象类似的一切对象都和与后一个对象类似的那些对象处在类似的先行关系和接近关系中。或者，换言之，假如没有前一个对象，那么后一个对象就不可能存在。

- 1973年，大卫·刘易斯的《反事实》呼吁学界**放弃规律性定义**
 - 应该将“A导致B”解释为“假如没有A，则B就不会发生”
 - 我们是通过想象一个或多个**“假如世界”**来实现这一点的，在这些假如世界中，反事实陈述为真
 - **认为可能的世界就相当于真实存在的世界**，而认为我们是通过比较我们的世界和在其他方面与现实世界**最相似**的那个假如世界来评估反事实陈述的
- 1974：唐纳德·鲁宾将**潜在结果**作为一种回答因果问题的语言
 - **打破了75年来统计学界对因果关系问题避而不谈的局面**
 - 变量Y的一个潜在结果就是“假如X的值为 x ，那么Y在个体 u 上的取值”，记为 $Y_{X=x}(u)$ 或 $Y_x(u)$
 - 假如X的值为 x ，那么变量Y一定会取某个与之相对应的值，其客观存在性与Y在现实中实际取的值相当
 - 潜在结果，或反事实，是在**个体层面**而非总体层面上定义的

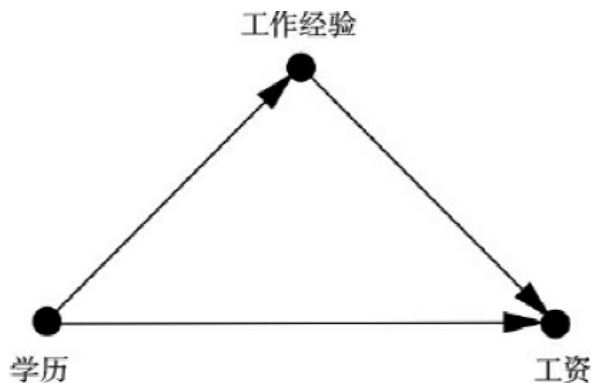
- 统计学家保罗·霍兰德曾称该问题为“因果推断的基本问题”
- Pearl认为将因果推断问题看作一个数据缺失“基本问题”会造成非常严重的误导

雇员 (u)	$EX(u)$ (单位: 年)	$ED(u)$	$S_0(u)$ (单位: 美元)	$S_1(u)$ (单位: 美元)	$S_2(u)$ (单位: 美元)
爱丽丝	6	0	81 000	?	?
伯特	9	1	?	92 500	?
卡罗琳	9	2	?	?	97 000
戴维	8	1	?	91 000	?
厄内斯特	12	1	?	100 000	?
弗朗西斯	13	0	97 000	?	?
等等

- 第一种推断方法: **匹配**
- 第二种推断方法: **线性回归** $S = 65000 + 2500 \times EX + 5000 \times ED = 85000$

■ 数据驱动的方法

- 所有数据都是通过检查表格中的其他值来填充的
- 根据因果关系之梯，没有那种纯粹基于数据的方法（第一层级）可以回答反事实的问题（第三层级）
- 数据是事实的呈现，反事实的推断不用数据，用什么呢？



$$S = 65000 + 2500 \times EX + 5000 \times ED + U_S$$

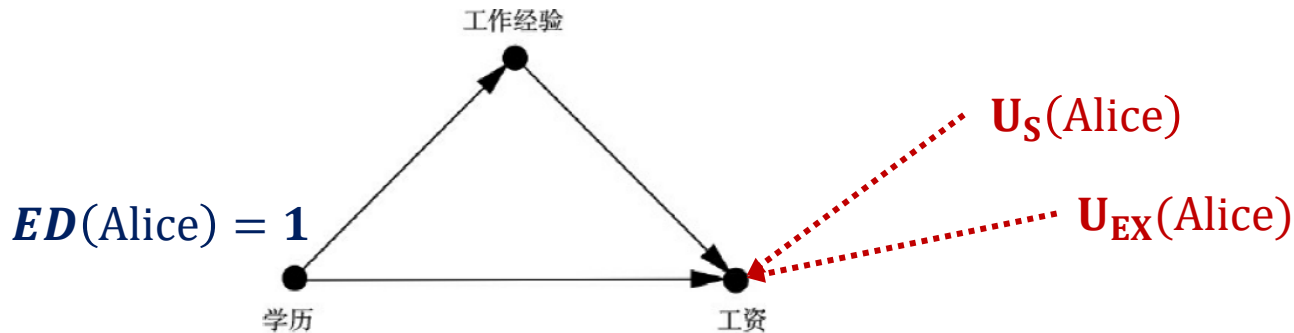
更高学历的员工花了更长时间去接受教育，因此假如爱丽丝和波特一样都是本科学位，爱丽丝则会比伯特有更少的工作经验。

$$EX = 10 - 4 \times ED + U_{EX}$$

工作经验是一个中介物，禁止统计调整

■ 从结构模型中推导反事实，三个步骤

- **外展**：利用关于各个员工的数据来估计Alice的特质因子： $U_S(\text{Alice})$ 和 $U_{EX}(\text{Alice})$
- **干预**：利用do算子改变模型，以反映我们提出的反事实假设，在这个案例中即，假如爱丽丝有大学学位： $ED(\text{Alice}) = 1$
- **预测**：利用修改后的模型及有关**外生变量**的更新信息 $U_S(\text{Alice})$ 和 $U_{EX}(\text{Alice})$ 和 $ED(\text{Alice})$ 来估算爱丽丝的工资水平。新的工资水平就等于 $S_{ED=1}(\text{Alice})$
 - $S = 65000 + 2500 \times EX + 5000 \times ED + U_S = 76000$
 - $S = 65000 + 2500 \times EX + 5000 \times ED = 85000$



表象与本质

■ 从结构模型中推导反事实，三个步骤

■ 外展-干预-预测：因果推断第一定律

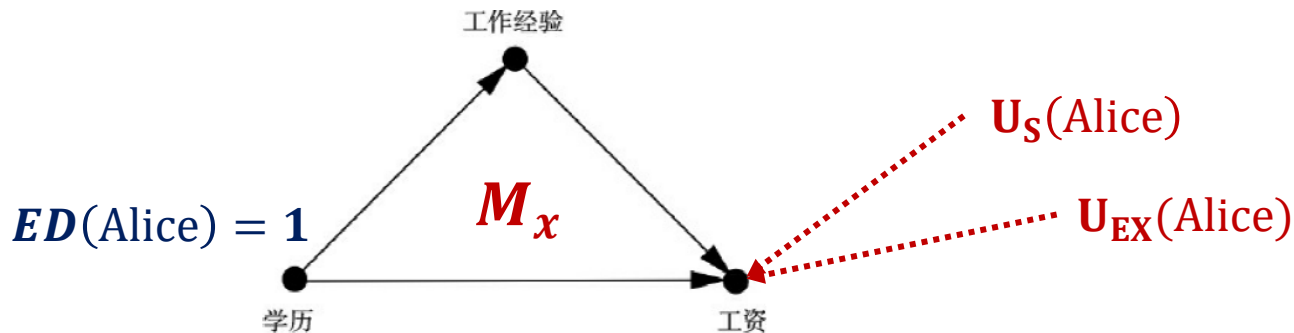
$$Y_x(u) = Y_{M_x}(u)$$

■ 然而，世上没有免费的午餐，我们所得强有力结果的前提是强有力的假设

- $S = 65000 + 2500 \times EX + 5000 \times ED + U_S = 76000$

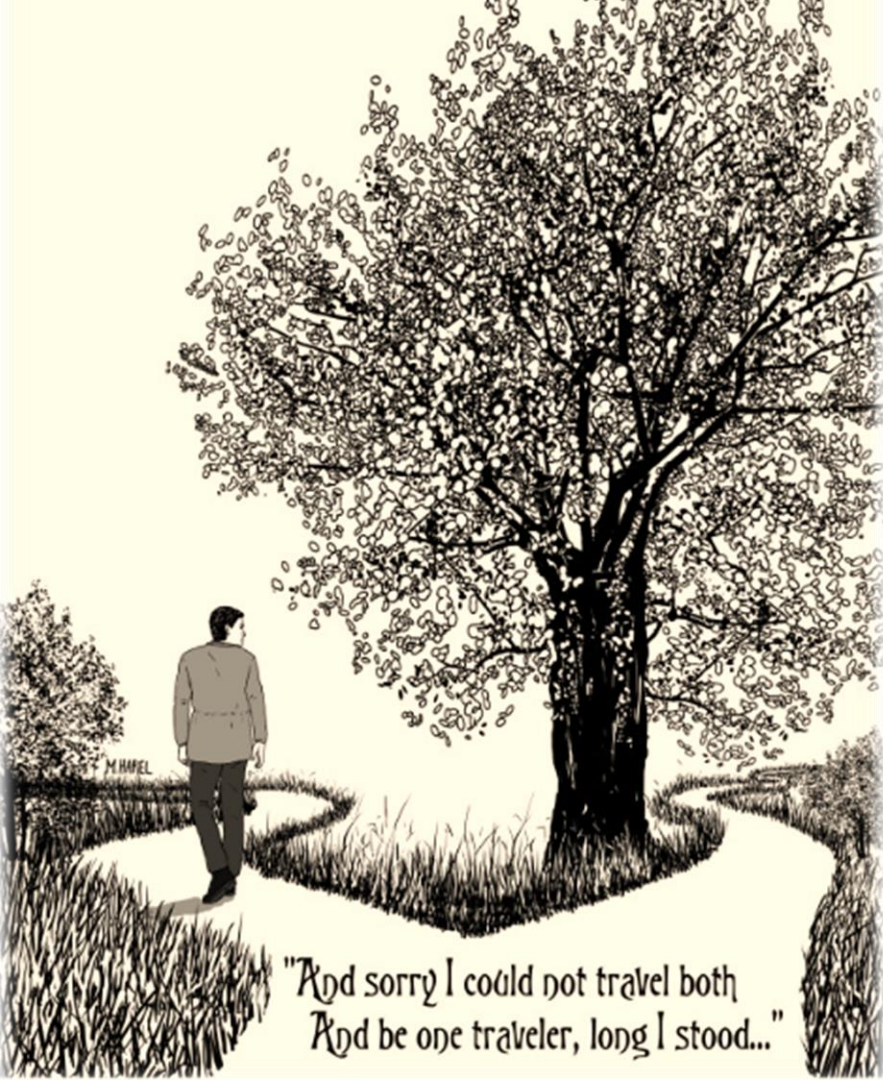
- 变量间的因果关系已知、线性函数关系成立...

- 事实上不总能获得这些函数信息，但对许多实际应用案例，部分指定模型所能给出的概率答案已经足够好了，比如贝叶斯网络



表象与本质

□ **反事实是我们人类认识世界，认识人类行为如何影响世界的基本途径之一**



"And sorry I could not travel both
And be one traveler, long I stood..."

Two roads diverged in a yellow wood
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth

Then took the other, as just as fair
And having perhaps the better claim
Because it was grassy and wanted wear
Though as for that the passing there
Had worn them really about the same

And both that morning equally lay
In leaves no step had trodden black
Oh, I kept the first for another day
Yet knowing how way leads on to way
I doubted if I should ever come back

I shall be telling this with a sigh
Somewhere ages and ages hence
Two roads diverged in a wood, and I
I took the one less traveled by
And that has made all the difference



Shining a light in the darkest night

