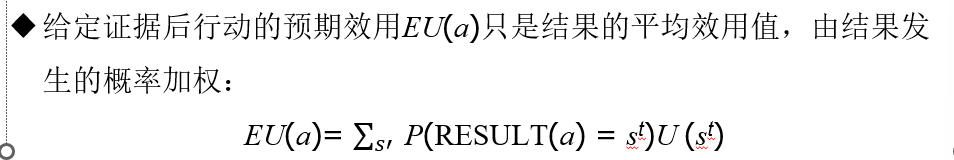
**16.1 在不确定性下结合信念与愿望**

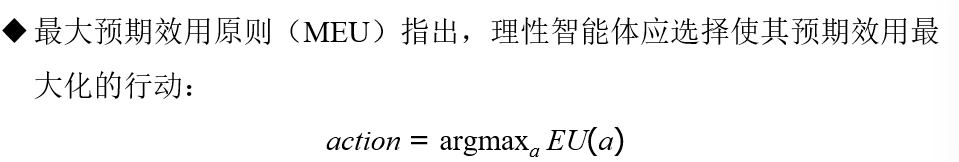
智能体为每个可能的当前状态s分配概率P(s) 。行动结果也可能存在不确定性；转换模型由 P(st | s, a)给出：在s状态下，智能体做出a动作，达到st状态的概率

做出动作a达到st状态的概率 P(RESULT(a) = st ) = ∑\_𝑠 P(s)P(st | s, a)

对某个状态的愿景U(s)

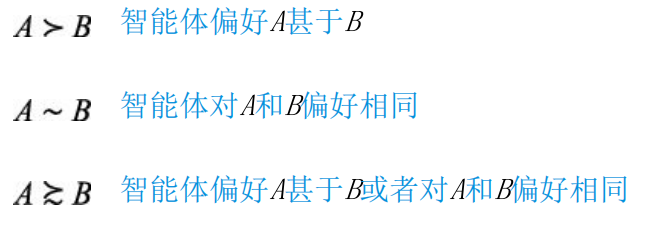
一个动作的期望效用：EU(s) a动作可能道道状态效用的加权平均值



****

**16.2 效用理论基础**

**16.2.1 理性偏好的约束**



Lottery彩票：每个动作的结果集合

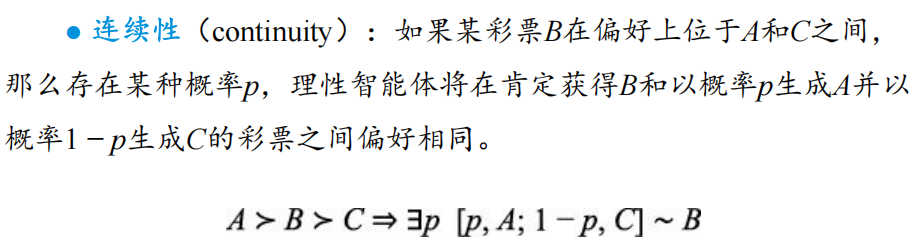
可能结果为S1, …, Sn，出现概率分别为p1, …, pn的彩票L可以写作

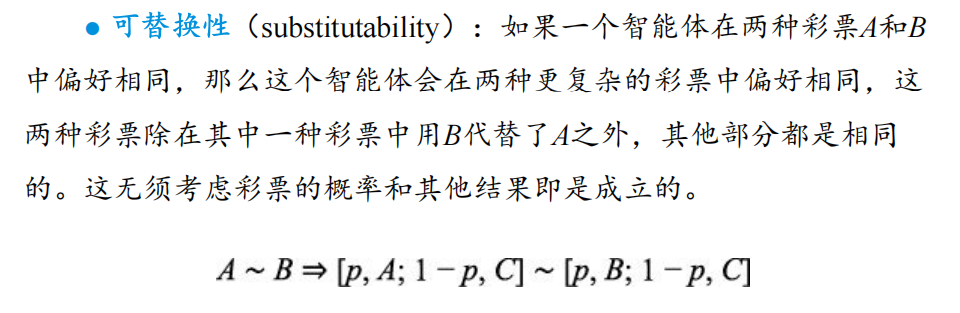
彩票：感觉可以看作一种混同策略

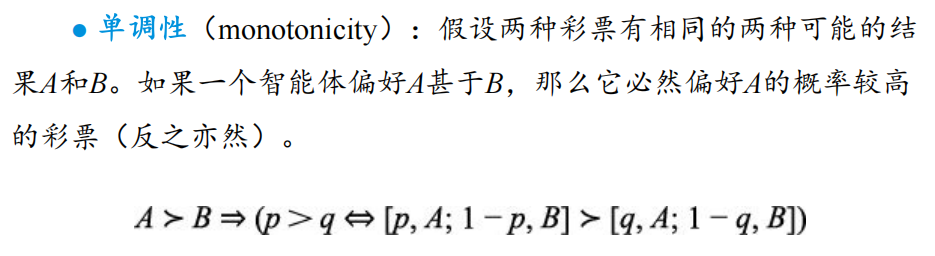
**效用理论的公理**

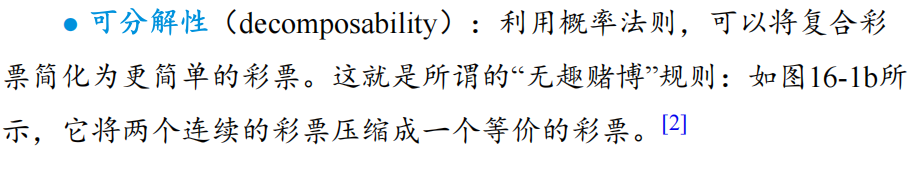
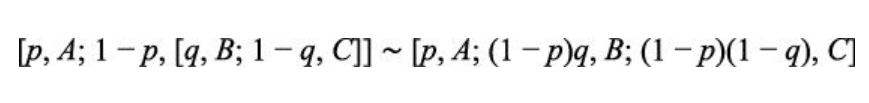
**6个任何合理的偏好关系都要满足的约束：**

有序性、传递性、连续性、可替换性、单调性、可分解性





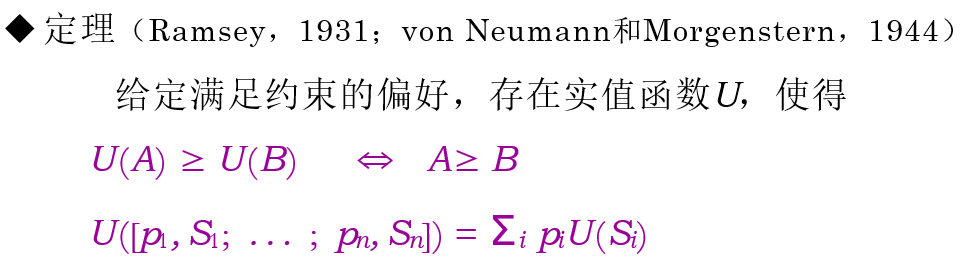




每一个公理都可以通过证明**违反它的智能体在某些情况下会表现出明显的非理性行为**来说明其提出的合理性。

**16.2.2 理性偏好导致效用**

**效用函数的存在性以及彩票的期望效用**



对于满足效用公理约束的派你好，一定存在一个效用函数刻画这种偏好

**16.3 效用函数**

**16.3.1 效用评估和效用尺度**

归一化效用（normalized utility）意味着u⊥ = 0且u⊤ = 1的尺度

规定结果的效用最好为1，最差为0

**16.3.2 金钱的效用**

金钱不是一种效用函数

一般来说，在所有其他因素相同的情况下，一个智能体偏好更多的钱。我们称智能体表现出对更多金钱的单调偏好（monotonic preference）。这并不意味着金钱表现为效用函数，因为它并没有提到涉及金钱的彩票之间的偏好。

在对实际效用函数的开创性研究中，格雷森（Grayson, 1960）发

现，**金钱的效用几乎与金额的对数成正比**。

拥有这种形状曲线的智能体是风险厌恶（riskaverse）的：它们更喜欢稳赚不赔的东西，而这种东西的收益要低于赌博的预期货币价值。但是，在图16-2b中，在大量负财富的“绝望”区域，这种行为呈现出风险寻求（risk-seeking）。智能体能接受的代替彩票的价值被称为彩票的确定性等价值（certainty equivalent）

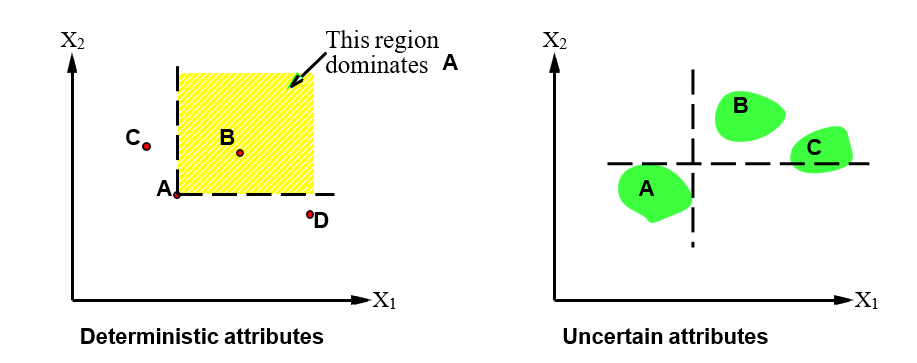
**16.3.3 期望效用与决策后失望**

**16.4 多属性效用函数**

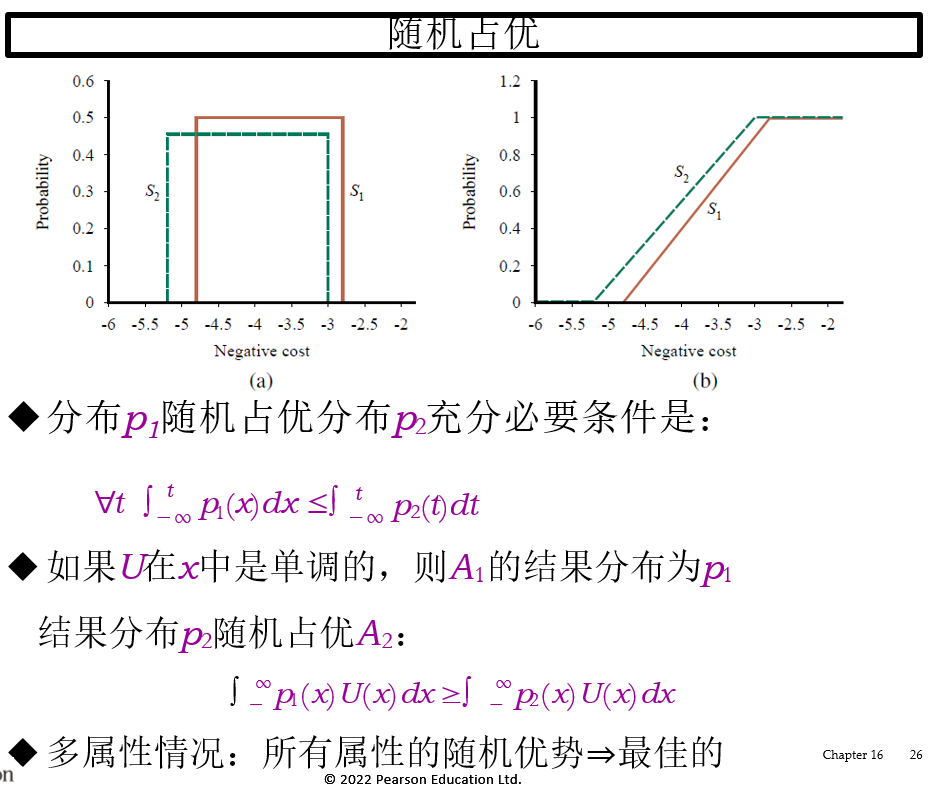
**16.4.1 占优**

**通常定义属性使U 在每个属性中都是单调的**

**严格占优：∀ i Xi(B) ≥ Xi(A) (因此 U (B) ≥ U (A))**



**随机占优**

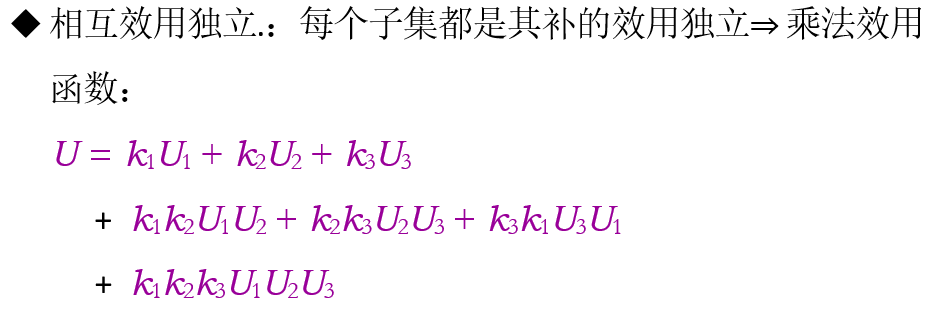


**X1 和 X2优先独立于X3**：如果(x1, x2, x3)和(xt1, xt2, x3)之间的优先不依赖于 x3

定理（Leontief，1947）：如果每对属性都是其补的偏好独立，那么每个属性子集都是其补的偏好独立：相互偏好独立

定理（Debreu，1960）：相互偏好独立⇒ ∃ 加值函数： V (S) = ΣiVi(Xi(S))

**X与 Y无关** 的充分必要条件是：**X中的彩票偏好不依赖于Y**



**决策网络**

使用的节点有三种类型：

机会节点（椭圆形）：随机变量，就像它们在贝叶斯网络中一样

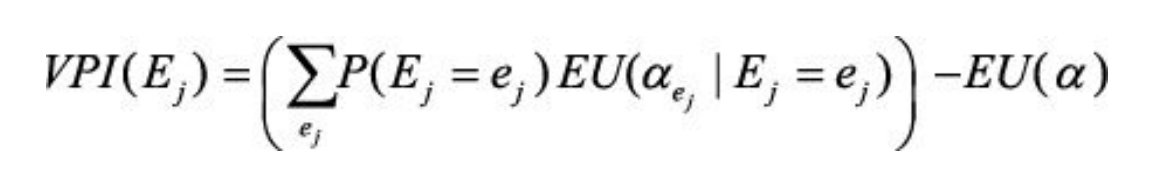
决策节点（矩形）：决策者可以选择行动的点

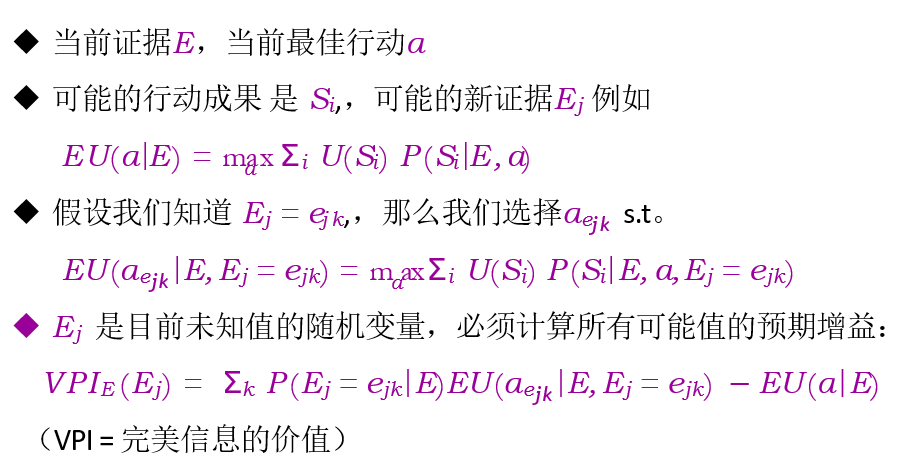
效用节点（菱形）：智能体的效用函数

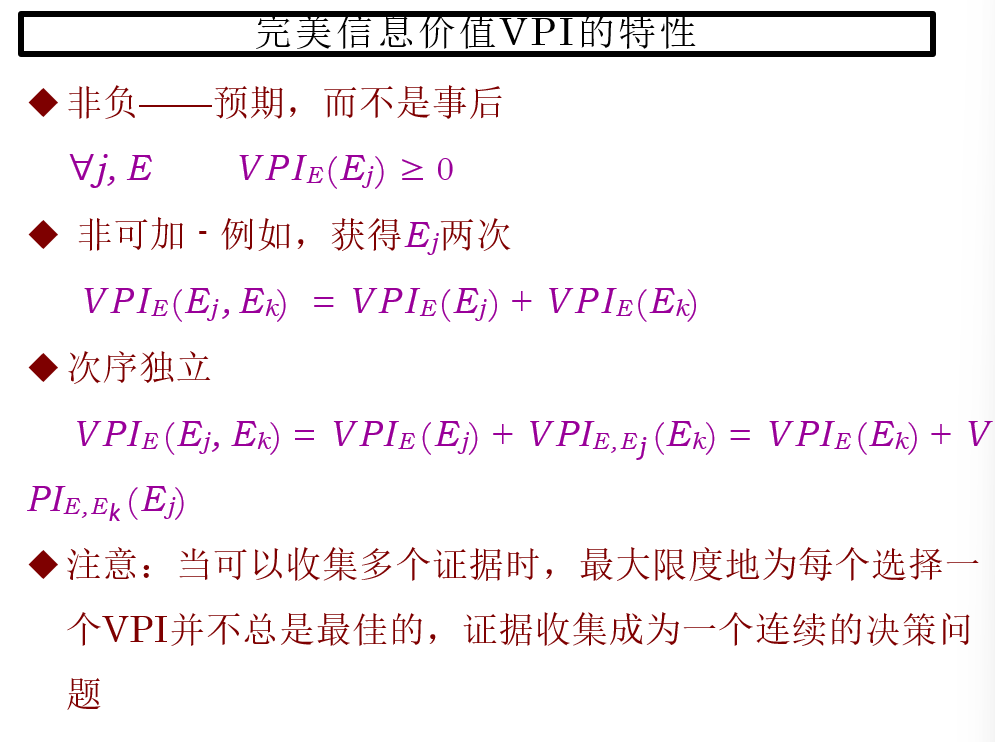
**信息价值**

给定信息的最佳行动的预期值减去没有信息的最佳行动的预期值

完美信息的一般公式？？？？？

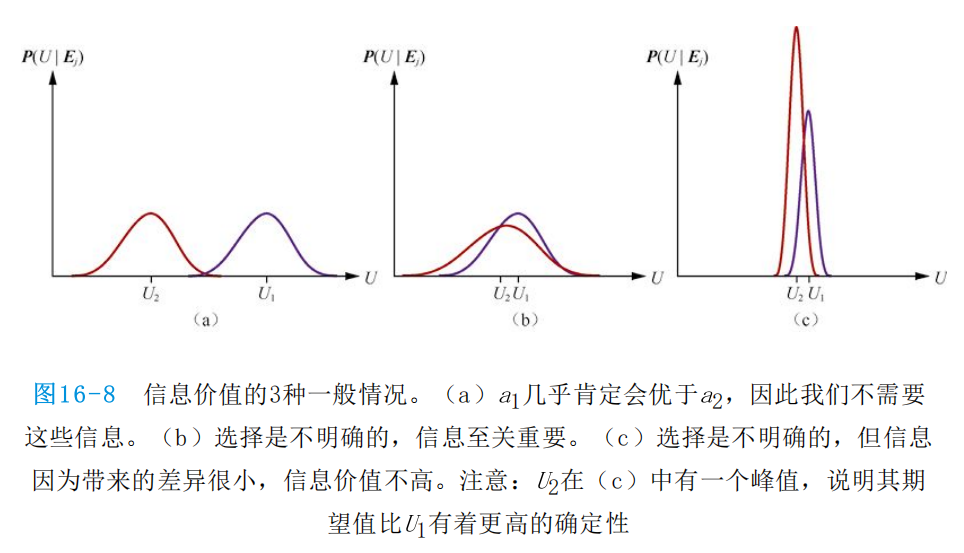






信息的作用：缩小可能结果的范围/缩小效用函数的可能取值范围，即使获得更具确定性的效用估计（实际的效用与估计效用的偏差减小）

**信息价值的定性比较**



（图为效用U1、U2各自的概率分布）