不可访问信息理论[1]

Jacopo Surace[1,2]

[1]周界理论物理研究所，31 Caroline Street North，滑铁卢，ON N2L 2Y5，加拿大[2]ICFO - 光子科学研究所，巴塞罗那科学技术研究所，卡斯特尔德费尔斯（巴塞罗那），08860，西班牙

arXiv:2305.05734v2 [quant-ph] 2 Sep 2024

如果我们探索世界的能力存在基本极限，会产生什么后果？在这项工作中，我们认真探讨这个问题。我们首先假设存在一些陈述，其真值无法通过实验来检验。也就是说，即便在理论上，也没有办法直接检验这些陈述究竟为真还是为假。我们进一步构建了一种理论，其中实验上可及的陈述是固定最小数量的不可访问陈述的并集。例如，陈述 a 和 b 的真值不可访问，但陈述“a 或 b”的真值则可以访问。我们不直接假设概率论，而是仅定义实验上可及和不可访问的陈述，并以经典逻辑的规则为基础进行推导。我们发现，由此涌现出一种有趣的结构。通过发展这一理论，我们放宽逻辑结构，自然而然地推导出一种结构丰富、受约束的准概率理论，并将其命名为不可访问信息理论。令人惊讶的是，不可访问信息理论最简单的模型就是量子力学中的量子比特。在构建这一理论的过程中，我们刻画并研究了一类乘法信息度量，并将其称为不可访问性度量。

1 引言

从发现不确定性关系开始，并发展出内在量子随机性的概念 [1, 2, 3, 4], 量子力学告诉我们，通过实验所能获取的信息存在一个根本限制。我们对量子力学是否会继续作为我们能够构建的最根本理论知之甚少。然而，它给予我们的教训似乎比量子力学本身的公设更为根本。

如果我们暂且（但不会太久）忘掉量子力学，只把我们能通过实验获取的信息存在一个基本上限作为基本公设，并探究由此假设产生的理论结构，会怎么样？在这项工作中，我们探讨这个问题，研究我们称之为不可及性假设的后果。

不可及性假设 $\mathcal{IH}$:
关于物理世界的一些陈述是绝对实验不可及的，也就是说，即使原则上，也不存在任何能够证实或证伪此类陈述的实验程序。

我们沿着一条与该领域以往研究部分不同的路径 [5, 6]。不假设概率论，我们从逻辑的基本语言出发 [7, 8]，构建一种我们称之为不可访问信息理论的理论。

我们从将所有合乎语法的陈述划分为两类这一思路出发，进一步发展这一观点：实验上可访问的陈述和不可访问的陈述。我们将实验上不可访问的陈述定义为：即使在理论上，也不存在任何实验方法可以评估其真值的陈述。我们不讨论诸如某事物在实验上不可访问时，所谓真或假意味着什么之类的形而上学问题。

在我们的理论中，所有实验可及的陈述都是由至少 d 个实验不可及的陈述组合而成的，其中 d 代表我们所称的实验知识的可访问性深度。例如，若 $d = 3$ 以及三个原子陈述 a, b, c 不可访问，那么陈述 $r = a \vee b$（意为“a 或 b”）便无法通过实验获取，但由 $d = 3$ 个不可访问陈述组成的陈述 $s = a \vee b \vee c$ 则可以。现在我们可以明白，为什么将 d 解释为可访问性深度：即使实验表明 s 为真，这一知识仍会让我们完全无法得知构成 s 的三个实验不可及陈述 a, b, c 的真值。这就是我们刻画实验知识根本

[1]认真接受某些陈述无法通过实验直接证实或证伪所带来的后果。

已于 Quantum 2024-08-30 接收，点击标题进行核验。根据 CC-BY 4.0 发布。

局限性的方式。

通过发展并放宽理论的初始结构，与实验不可及陈述相关联的准概率概念，以及与实验可及陈述相关联的概率概念，将自然地显现出来。在这一放宽后的情形中，可访问性深度参数 d 的取值将对我们与实验可及陈述相关联的概率向量施加一些信息论约束。令人惊讶的是，我们会发现，在众多可能的模型中，量子比特的结构是不可访问信息理论所容许的最简单模型。

我们提醒读者，我们的不可访问信息理论并非旨在推导量子力学，而是旨在研究接受某些陈述无法通过实验直接证实或证伪这一前提所带来的后果。在构造过程中，我们还将揭示如何从逻辑陈述代数的结构中推导准概率，并引入一组乘法型信息度量，称为不可访问性度量。

2 实验上可及与不可及的陈述

本节旨在介绍我们理论的框架。我们首先定义一种形式语言，它由一组符号以及一组规定这些符号如何复合的规则组成。我们将使用的符号集中的元素称为 原子陈述。原子陈述旨在表示一种陈述，用以声明系统的一项基本物理属性。重要的是，我们希望原子陈述能够编码系统的互斥属性：任何系统同时具有两种原子陈述所描述的属性都是荒谬的。这个想法稍后将形式化；但目前，我们可以举一个例子。以掷骰子为例。骰子顶面的点数是骰子的一项基本属性。此外，骰子只有一个顶面，因此谈论两个顶面是荒谬的（那就不是骰子了）。因此，像 s_1=“骰子的顶面是 6 点”或 s_2=“骰子的顶面是 1 点”这样的陈述，就是原子陈述的好例子。[2] 形式上，维度为 维度 D 的模型将由一组D 原子陈述 $\mathcal{A}(\mathcal{S}) = \{s_1, s_2, \ldots, s_D\}$ 来刻画。

[2]原子陈述的其他示例可以是“自旋的方向指向 $\vec{r}$”（对于完备方向集），或者“按下这个按钮，我在屏幕上得到一个 i ”，适用于所有可能的结果i。将被视为不可访问的陈述示例是：“量子比特的 X 和 Z 的值为 1”。

我们使用来自经典命题逻辑的一组逻辑常项，将原子陈述组合起来形成新的陈述。这些逻辑常项包括一元算子$\neg$（表示对陈述的否定），以及二元算子“或”（$\vee$）和“与”（$\wedge$）。

此外，我们引入了两个特殊元素， $\top$和 $\bot$，它们分别表示永真命题和矛盾命题的直观概念。

在我们刚才定义的所有要素集合 (A(S) $\vee, \wedge, \neg, \top, \bot$) 中，我们加入一组规则，用以规定不同陈述之间的等价关系。我们规定的第一个等价关系使用 $\bot$ 来形式化原子陈述互斥这一性质。具体而言，原子陈述的定义使得，任意两个原子陈述 s_i 和 s_j 通过“与”运算符进行复合后，所构成的陈述都等价于矛盾命题 $s_i \wedge s_j = \bot$ ，对每个 $i \neq j$ 均如此。附录 B.1 中的方程 (60)、(61)、(62) 规定了更多等价关系。例如，这些规则体现了这样一种直觉：根据幂等律 (60) 的定义，陈述 s_1 与陈述 $s_1 \vee s_1$ 应视为等价。[3] 以所有可能的方式组合原子陈述，便可得到商空间 S，它由通过组合原子陈述所形成的陈述的等价类构成；或者等价地说，它由每个陈述等价类的代表元构成。集合 S 表示我们可以表述的所有可能陈述的集合。可以看出，S 中的每个陈述都等价于仅使用二元运算符 $\vee$ 对原子陈述进行复合所得的陈述，而 $\top$ 元素等价于使用二元运算符 $\vee$ 对所有原子陈述进行复合所得的结果。由此，我们得到命题代数 (S $\vee$, $\wedge, \neg, \top, \bot$)，或者等价地说，由原子陈述集合构成的完备布尔格（见附录 B.1 和 [7, 9, 10]）。

现在，我们为格中的每个元素附上一个标签，称之为陈述的真值。这个标签可以取两个值：**T** 或**F**，分别读作“真”和“假”。如果对格中每个元素赋予真值时遵循以下规则，就称这种赋值是可接受的。首先，它必须符合真值表1所规定的规则。例如，如果我们给 s_1 赋值**T** ，给 s_2 赋值**F**，那么给 $s_{1\wedge 2} = s_1 \wedge s_2$ 赋值**F** 就是唯一可能的选择。此外，我们还要求等价陈述具有相同的真值赋值。最后，

[3]再举一个或许更复杂的例子，可以证明对于 $s_3 \wedge s_1 \vee s_2$ 原子陈述 $s_1 \vee s_3 \wedge s_1 \vee s_2 = s_1 \vee s_3 \wedge s_2$,,等价关系 ()() 成立。事实上 ()() () 根据规则 (62b)，这等价于 $s_1 \vee \bot$ 由于互斥原子陈述的定义，这又等价于 $s_1 \vee (s_1 \wedge \neg s_1)$ 根据规则 (63b)。最后，使用吸收律 (60e)即可得到所需结果。

a	b	a ∧ b
T	**T**	**T**
T	**F**	**F**
F	**T**	**F**
F	**F**	**F**

a	b	a ∨ b
T	**T**	**T**
T	**F**	**T**
F	**T**	**T**
F	**F**	**F**

a	¬ a
T	**F**
F	**T**

表1：真值表。

a	b	a ∧ b
A	**A**	**A**
A	**N**	**A/N**
N	**A**	**A/N**
N	**N**	**A/N**

a	b	a ∨ b
A	**A**	**A**
N	**A/N**	
N	**A**	**A/N**
N	**N**	**A/N**

a	¬ a
A	**A**
N	**N**

表 2：可达性表。写作 A/N 时，表示可达性值未定义且可以任选。请注意，实验上可及的陈述的复合和否定总会得到一个实验上可及的陈述。

我们希望始终将矛盾命题赋值为**F**，因为我们想表达这样的思想：原子陈述彼此互斥，因此，对于任意两个原子陈述 s_i， s_j，其中 $i \neq j$，人们预期 $\bot = s_i \wedge s_j$ 会被赋予真值**F**

只需指定每个原子陈述的真值，并要求赋值是可容许的，就足以计算格中其他所有陈述的真值。因此，可以认为真值从原子陈述出发，在格中传播。还可以进一步验证，可容许的真值赋值仅有一种：一个原子陈述为真，其余所有原子陈述均为假。

我们为每个陈述附上一个 额外的 标签，称为该陈述的**可达性值**。此标签可取两个值 **A** 和 **N**，分别读作“实验上可及”和“实验不可及”。若对格中每个元素赋予可达性值时遵循以下规则，则该赋值被视为可容许。首先，它必须符合可达性表所规定的规则2。此外，我们要求等价陈述的可达性值赋值相同。

我们预期，将实验可及的陈述组合起来或取其否定，应当得到一个实验可及的陈述。此外，我们还应预期，对陈述取否定不应改变其可及性值，否则只要考虑某个不可及陈述的否定，就能轻易地获得该陈述。在一般情况下，我们将考虑满足以下条件的格： $\top$ 始终是实验可及的。这是因为，当 $\top$实验不可及时，该格就是平凡格，其中所有陈述在实验上都不可及（见引理 1 及附录A中的证明）。我们注意到，格中元素的可及性值并不能完全由原子元素的可及性值确定。

3 逻辑不可访问信息模型

在上一节中，我们引入了讨论可访问陈述和不可访问陈述的框架以及$\mathcal{IH}$。本节将为上一节中的定义增添一些结构。我们在此非正式地陈述不可访问信息理论的主要假设。

逻辑格的假设 1

只有包含至少 d 个原子陈述的析取式才是可达的，其中 d为某个值。也就是说，我们的知识始终局限于对固定数量的陈述保持不确定性 d。我们将 d 称为模型的可访问性深度。

假设 1 背后的想法是，我们的理论应描述这样的系统：实验上可访问信息是某些实验上不可访问信息的粗粒化。这是引导我们的直觉，但从务实角度来说，我们只是在探索这样一种设想的后果：建立在这一假设之上的理论，也就是说，某些信息不可访问的理论。

可访问性深度 $\mathbb{N} \ni d \geq 1$ 是通过联结词 $\vee$形成实验上可及的陈述所需的原子陈述的最小数量。或者，将格的层级从底部开始编号（其中 $\bot$ 处于层级0）， d 是允许出现实验上可及陈述的格层级（参见附录 B.2）。

定义 1. 模型（D， d）是一个列表（L_D， A_d， V），其中 L_D 是此前引入的命题代数，或布尔格（具有 D个原子元素）， A_d 是所有符合“可访问性深度”为 d这一要求的“可达性赋值”的集合，而 V 是以此前所述方式加以约束的可容许“真值赋值”的集合。

已于 2024-08-30 接收，点击标题进行验证。根据 CC-BY 4.0 发布。

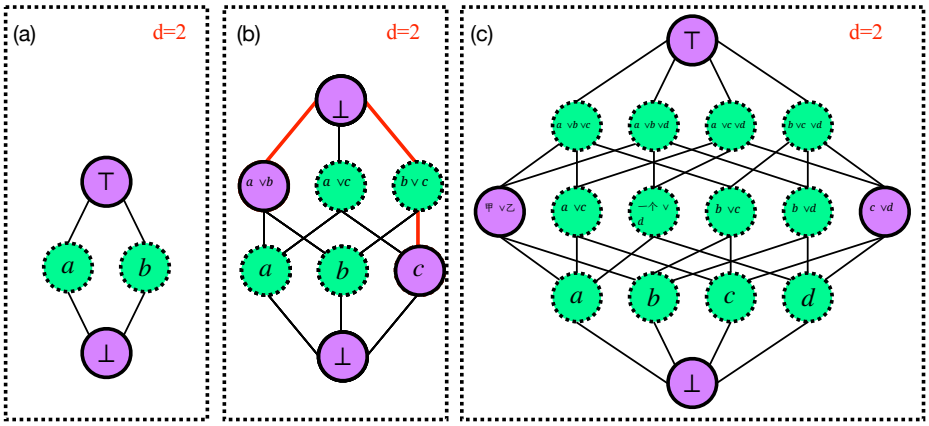

图1：三个示例的图示。陈述以彩色球表示。球的颜色代表其可及性值。紫色球对应实验上可及的陈述，绿色球对应实验上不可及的陈述。我们表示不同维度 D下可访问性深度为 $d=2$ 的模型。在面板 (a) 中表示的是例1的情形。对于 $D=2$，唯一可访问的陈述是 $\bot$ 和 $\top$，因此这是一个无用模型，因为它没有传达任何实验相关信息。在面板 (b) 中表示的是示例2的情形。对于 $D=3$，唯一可访问的陈述是 $\bot$ 和 $\top$，因此这是一个无用模型。在该面板中，证明中的一个步骤被突出显示：若选择陈述 $a \vee b$ 为可访问的，则意味着其否定 $\neg(a \vee b) = c$ 也可访问，但根据定义，c 不可能可访问，因此产生矛盾。在面板 (c) 中表示的是示例3的情形。这是第一个非平凡模型。从图中很容易看出，可访问的（紫色）陈述构成一个经典模型的子格，其 $D=2$。

定义 2。模型的配置为 (L_D, a_d, v)，其中有 $a_d \in A_d$ 和 $v \in V$。

对于每个模型，允许的构型是那些既符合可及性表规则，又符合模型可访问性深度 d 的构型。

例如，对于模型 $(D, d) = (4, 2)$，我们知道，可访问性深度 $d=2$ 的选择意味着所有原子陈述在实验上都不可及，但允许陈述 $s_i \vee s_j$，以及 $i, j \in \{1, 2, 3, 4\}$ 和 $i \neq j$ 在实验上可及。此外，如下文的一些示例所示，要求符合可及性表规则还会施加进一步的约束。

对于给定模型，我们倾向于选择能使实验可及的状态陈述数量最多的构型。我们将这些构型称为理想构型，并将其定义如下

定义 3。模型 (D, d) 的理想配置是一个允许的配置，其可及性值赋值能使层级 d 上可及陈述的数量最多。

不可访问信息理论与经典理论的区别，在于可访问陈述和不可访问陈述之间的区分。取消这一划分后，我们应当能够恢复出一种经典理论。这相当于将可访问性深度设为模型的 $d=1$，并考虑唯一的理想配置，为该理论的每一条陈述赋予可访问值。

定义 4. 维度为 m 的经典模型，是一个在理想配置中具有 $(D=m, d=1)$ 的不可访问信息模型。

我们约定用符号 m 表示经典模型的维度。

事实上，对于 $d=1$，使层级 d 上可访问陈述数量最多的可达性赋值方式，就是将所有原子陈述都设为可访问，从而还原出一种没有有效可达性标签的理论。

对于这类配置，实验上唯一可访问的陈述是 $\top$ 及其否定 $\bot$，因此它们不包含任何相关的物理信息。这是具有 $D=d$ 的模型唯一允许的配置。我们将这些模型称为无用模型。

通过以下示例可以看出，假设 1 已足够有力，能够揭示该理论的一些有趣性质。

例1. 考虑具有 $D=2$ 个原子的模型。很容易验证，唯一允许的模型是经典模型和无用模型。参见图 (a) 的 *1* 部分，我们可以看到 “a” 是

的“b” 的否定，因此“a” 和“b” 必须具有相同的可及性值。由此可知，“a” 和“b” 要么都可及（经典模型），要么“a” 和“b” 都不可及（无用模型）。

示例2。考虑一个格，其中 $D=3$ 处于理想配置中。对于 $d=1$ 它是一个经典模型；对于$d=3$ 它是一个无用模型。我们研究 $d=2$ 的情况。由于格的结构，我们有$s_2 \vee s_3 = \neg s_1$。由于 $d=2$ 我们有 s_1 是 $\boldsymbol{N}$，这意味着 $s_2 \vee s_3$ 也必须是 $\boldsymbol{N}$。对于 $\neg p_2$ 和 $\neg p_3$，我们可以用同样的推理得出：如果$d=2$ 那么不允许在层级 3 存在具有可访问陈述的构型。因此，具有三个原子元素的格实际上只能是经典模型或无用模型。图 *1* 展示了图示。

示例3。考虑一个具有 $D=4$ 的格，处于理想配置中。这是非平凡格的第一个情形。事实上，我们可以考虑一个可达深度为 $d=2$的模型。不失一般性，我们可以赋予 $s_1 \vee s_2$ 可及性值 $\boldsymbol{A}$。这意味着 $s_3 \vee s_4 = \neg(s_1 \vee s_2)$ 也为 $\boldsymbol{A}$。与此同时$s_1 \vee s_3$不能为 $\boldsymbol{A}$。事实上，如果 $s_1 \vee s_3$为 $\boldsymbol{A}$，那么我们就会得到 $s_1 = (s_1 \vee s_3) \wedge (s_1 \vee s_2)$ 也为 $\boldsymbol{A}$。这将与 $d=2$矛盾。图示见图 *1* 。

在附录 A 中，我们证明以下引理：

引理 1. 如果 $\top$ 是$\boldsymbol{N}$，那么格中的所有陈述都是$\boldsymbol{N}$。

引理2。对于每个模型(D,d)，构成可容许配置中层级 d 的两个可访问陈述的原子陈述集合不得有任何共同元素。

引理 3. 对于每个模型 (D,d)，理想配置中层级 d 处可访问的陈述为 $\frac{D}{d}$，当$\mathrm{mod}(D,d)=0$时；当$\mathrm{mod}(D,d)\neq 0$时，则为 $\lfloor \frac{D}{d} \rfloor - 1$。 —

引理 4. 对于每个模型 (D,d)，所有理想配置在原子陈述的重新标记下都是等价的。

在下文中，每当我们指定一个模型时，都会考虑其理想配置，因为这种配置能使模型中与实验相关的信息最大化。

如例 3所示，可以将一个经典系统嵌入不可访问信息模型的理想配置中。在这种情况下，我们可以说，不可访问信息模型是经典模型的膨胀。

定义5。维度为 m 的经典模型的膨胀是指任何处于理想配置中的模型 $(D,\ d)$，其层级 d 上恰有 m 个可访问陈述。

引理5. 每个具有 m 原子元素的经典模型只能膨胀为模型 $(D,\ d)$，其中，对于每个可访问性深度 d，允许的集合 D 是

$$D \in \mathcal{D}_{m,d} \coloneqq \{md\} \cup \bigcup_{i=1}^{m-1} \{(m+1)d+i\} \tag{1}$$

由引理 5 可知，经典模型无法扩展为具有任意数量原子陈述的模型 D。例如，具有$m=2$ 的经典模型无法扩展为具有 $D=5$的不可访问信息模型。

我们旨在将不可访问信息的模型解释为经典模型的细化。具体而言，我们对这样一种设想所带来的影响很感兴趣：每个经典模型都可能是更大的不可访问信息模型中的一部分。因此，从现在起，我们将重点关注经典模型的膨胀模型。下一节将对其进行刻画。

4 模型的复合

从一开始就不一定能够全面了解我们想要描述的系统。通常，人们会依次描述系统的许多不同部分，并希望通过重构过程最终获得完整的描述。因此，在还原论原则的指导下，我们可能会希望制定一条规则，将不同的模型合并为一个模型。本节将介绍如何将两个不可访信息模型拼接在一起。

两个模型（D_1, d_1）和（D_2, d_2）的复合，其原子陈述分别为 $\{r_1, r_2, \ldots, r_{D_1}\}$ 和 $\{s_1, s_2, \ldots, s_{D_2}\}$，会形成一个新的陈述格，其中 $D_1 D_2$ 原子元素为 $\{p_\alpha\}_{\alpha=1,\ldots,D_1 \cdot D_2}$，可访问性深度为 $d_1 d_2$。我们将（D_1, d_1）×（D_2, d_2）$=$（$D_1 D_2, d_1 d_2$）记作。复合格中的每个原子元素都是两个原始格的原子元素的乘积，如下所示

$$\{p_\alpha\}_{\alpha=1,\ldots,D_1 \cdot D_2} \coloneqq \{r_i \cdot s_j\}_{\substack{i=1,\ldots,D_1 \\ j=1,\ldots,D_2}}, \tag{2}$$

其中 $\alpha = (i-1)D_2 + j$。

两个模型的复合可以看作这样一个模型：其原子元素是由第一个模型中的一个元素和第二个模型中的一个元素通过二元运算符 $\wedge$连接而成的所有组合。这就是原子元素数量

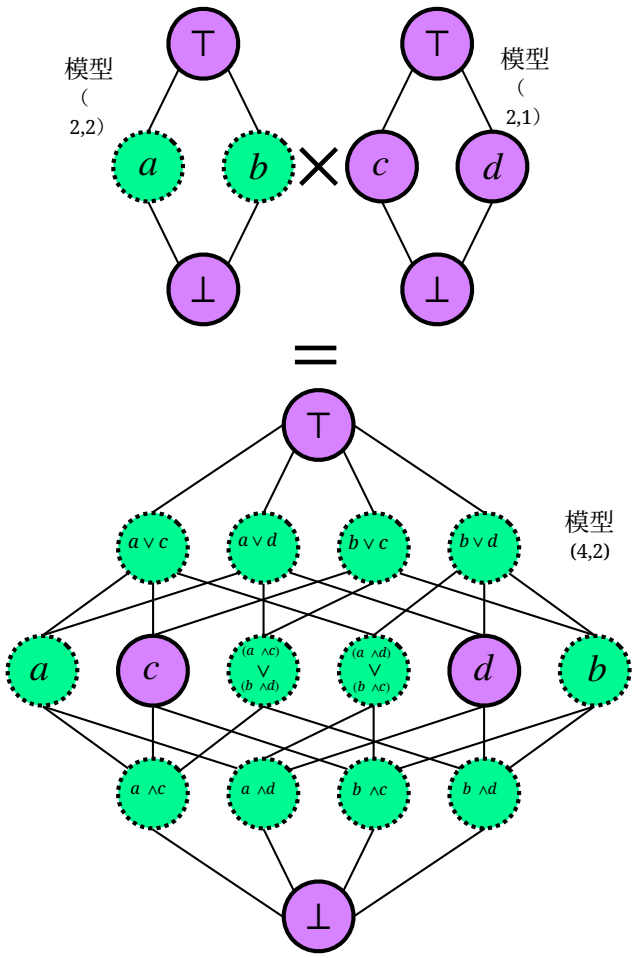

图 2：模型 (2，2) 与模型 (2，1) 的复合。

相乘的原因。与此同时，我们预期两个模型中实验可及的陈述在复合模型中仍然实验可及。例如，设第一个格中层级为 d_1 的陈述 $r_1 \vee \cdots \vee r_{d_1}$ 是实验可及的。复合这两个格时，我们预期该陈述仍然实验可及。在两个模型中分别处于层级 d_1 和 d_2 的陈述，在复合模型中处于层级 $d_1 d_2$ 。

例 4. 图2 展示了一个 $(2,\ 2)$ 模型与一个 $(2,\ 1)$模型的复合。复合模型是一个 $(2 \cdot\ 2 = 4,\ 2 \cdot\ 1 = 2)$ 模型，其可及性值继承自两个初始模型。

由于我们关注的是经典模型经膨胀后得到的不可访问信息模型，因此希望找到一个通用过程，将每个维数为 m 的经典模型映射到一个不可访问信息模型，并使该过程与模型的复合相容。我们将此过程称为膨胀过程。要求膨胀过程与模型的复合相容，意味着我们希望经典模型膨胀后的复合等价于经典模型复合后的膨胀。该过程由下述引理给出，证明见附录A。

引理 6. 存在一个唯一的膨胀过程族，以 $c \in \mathbb{N}_{>0}$ 为参数。该族中的每个膨胀过程都将每个维度为m 的经典模型映射为不可访信息模型 (m^{c+1}, m^c)。

如上一节末尾所述，我们感兴趣的是由经典模型膨胀而来的不可访信息模型。因此，我们希望对于任意选定的可访问性深度 d，都存在一个模型 $(D,\ d)$，从而存在一组 D 原子陈述，使得模型 $(D,\ d)$ 的理想配置是一个经典模型的膨胀。要求膨胀过程具有这一附加性质后，我们得到一个唯一的膨胀过程，即令 $c = 1$ 所得到的过程，因为选取 $d = 2$ 后，方程 $m^c = 2$ 在 $m,\ c \in \mathbb{N}$ 下有唯一解$m = 2,\ c = 1$。我们将其定义为标准膨胀过程。

定义 6。 一般的膨胀过程定义为以下映射

$$m \to (m^2, m), \tag{3}$$

将维度为 m 的经典模型映射到不可访问信息模型（m^2，m）。

因此，对于经典模型的膨胀，原子元素的数量必须按模型可访问性深度的平方进行缩放。图 (c) 中的面板 1可以解释为维度为 $m = 2$ 的经典模型的膨胀，也可以解释为可访问性深度为 $d = 2$ 的膨胀。

观察 1。 由于我们关注的是经典模型的膨胀，从现在起，我们将仅通过其可访问性深度d来唯一确定一个模型，因此，可访问性深度为 d 的模型是一个（d^2，d）模型。

下面我们将研究具有 d 素数的系统的性质。具有 d 非素数的系统可以看作具有 d素数的系统的唯一复合。

到目前为止，这一理论似乎没有任何实际用途。事实上，经典子格与整个格之间没有真正的信息传递。可以认为，不可访问信息理论的形式主义只是一种用于描述经典模型的过度参数化方法。

例如，考虑模型 (4,2)。要描述该模型的一个特定构型，必须指定

4 个不可访问元素的真值；随后便可用该模型表示一个维度仅为 2 的经典格的可访问信息。考虑到人们关注的是对实验上可访问信息的描述，这实际上是用四个参数描述一个二维参数空间。这种自由度体现在：不止一个不可访问真值的配置对应于同一个可访问真值配置。例如， $a \vee b$ 为真且$c \vee d$ 为假的配置，对应于两个原子元素 (a, b, c, d) 的构型 **(F**, **T**, **F**, **F)**、**(T**, **F**, **F**, **F)**。因此，就目前而言，带有不可访问信息的模型只是一种过度参数化。

此外，该理论的结构如此僵化：对于理想模型而言，不可能在不失去与可及性表规则一致性的情况下，将任何陈述的可达性从 **N** 更改为 **A** 。

在下一节中，我们将放宽组合规则，从而允许采用更丰富的结构。

5 从逻辑格到准概率

我们在二元逻辑的语境下讨论了不可访问信息的概念，其中每个陈述都被赋予一个二值真值。然而，现代科学语言并不限于二值真值赋值，而是采用概率语言。向这种现代语言的转变，历史上经历了漫长的发展过程。然而，借助 R. T. Cox[11, 12] 最先提出的思想，我们可以为这一范式转变赋予理论依据，并从形式上将概率论视为逻辑的推广。事实上，当引入「可信度」这一概念，并要求其遵循逻辑规则保持一致时，便可从命题逻辑推导出概率论规则。重要的是，概率论并非预先假定的，而是从逻辑公理和 Jaynes 所说的「常识」 [13, 14]推导而来。在这项工作中，受这些开创性研究的启发，我们沿着类似的路径，旨在将这一构造的适用范围扩展到不可访问信息理论。我们的方法不同于 R. T. Cox 和 E. T. Jaynes 的经典方法，因为我们选择将「常识」假设减少到最低限度。我们得到的是二元逻辑的一种推广，它的行为如同概率理论，但其「概率」赋值可以为负，也可以大于一。因此，我们称之为二元逻辑的推广，同时保持与命题代数结构的一致性

由此得到准概率理论。我们将这一过程称为把逻辑格模型放宽为准概率模型。完整推导见附录B.3，而我们的推导与 R. T. Cox 和 E. T. Jaynes 的经典推导之间的主要差异，则在附录B.3.3中作了重点说明。放宽为准概率模型的一项关键结果˜是构造一个函数 $Q(y \mid x)$ ，它以两个陈述 x和 y 为参数，并返回一个实数：

$$\tilde{Q}(y \mid x) = \begin{cases} 1 \text{ if } (x \wedge y = x) \wedge (x \vee y = y) \\ 0 \text{ if } x \wedge y = \bot \\ p \text{ otherwise, where } p \in \mathbb{R}, p \neq 0, p \neq 1. \end{cases} \tag{4}$$

˜函数 $Q(y \mid x)$ 在 x 于逻辑意义上蕴含 y 时返回 1，而在 x 和 y 为“不相关”的˜陈述时返回 0。要求函数 Q 符合命题复合的规则，我们发现Q 具有以下性质。

1. 求和规则：

$$\tilde{Q}(x \vee y|z) = \tilde{Q}(x|z) + \tilde{Q}(y|z) - \tilde{Q}(x \wedge y|z), \tag{5}$$

2. 乘法规则：

$$\tilde{Q}(x \wedge y|z) = \tilde{Q}(x|z)\tilde{Q}(y|x \wedge z), \tag{6}$$

3. 贝叶斯定理：

$$\tilde{Q}(y|x \wedge z) = \frac{\tilde{Q}(y|z)\tilde{Q}(x|y \wedge z)}{\tilde{Q}(x|z)}. \tag{7}$$

这表明，为了与陈述的组合规则保持一致，函数 Q 必须满足条件概率函数的性质。Q 与条件概率函数的区别在于， Q 不受必须返回区间 $[0, 1]$内数值的限制。

使用函数Q我们可以通过将与每个陈述相关联的二元真值标签 **T** 和 **F** 替换为一个实数 $Q(s_i)$， s_i 从而放宽为准概率模型，并使其满足以下性质：

- $Q(\top) = 1$,
- $Q(\bot) = 0$,
- $Q(s_i) \in \mathbb{R}$,
- $\sum_{s \in \mathcal{A}(\mathcal{S})} Q(s) = 1$,
- $Q(s_i \vee s_j) = Q(s_i) + Q(s_j) - Q(s_i \wedge s_j)$,
- $Q(s_i \wedge s_j) = Q(s_j)Q(s_i \mid s_j)$,

其中 $\mathcal{A}(\mathcal{S})$ 是陈述集 $\mathcal{S}$ 的原子元素集（见附录 B.1）。$Q(s_i)$的值是赋予陈述 s_i的准概率。由于 Q 的性质，每个复合陈述的准概率赋值都由 Q 在原子陈述 $\mathcal{S} = \{s_1, s_2, \ldots\ldots, s_{d^2}\}$上的赋值完全确定。因此，不可访问信息理论的准概率模型的一个配置，完全由向量刻画

$$\vec{q} \coloneqq \{Q(s_1), Q(s_2), \ldots, Q(s_{d^2})\}. \tag{8}$$

一般陈述 $s = s_i \vee s_j \vee \cdots \vee s_k$ 上 $Q(s)$ 的显式值，可以根据对 $\vec{q}$ 的了解轻松计算得出，其中 $s_i, s_j, \ldots, s_k$ 是原子陈述，计算如下

$$Q(s_i \vee s_j \vee \cdots \vee s_k) = Q(s_i) + Q(s_j) + \cdots + Q(s_k) \tag{9}$$

$$= q_i + q_j + \cdots + q_k. \tag{10}$$

利用这一点，可以验证 $\vec{q}$ 是一个适当的准概率向量，事实上:

$$1 = Q(\top) = \{Q(s_1), Q(s_2), \ldots, Q(s_{d^2})\} = \sum_{i=1}^{d^2} q_i. \tag{11}$$

我们得到的是准概率赋值而非概率赋值，这一事实完全符合不可访问信息理论的核心思想。实验不可及的状态并非物理上定义明确的对象，因此，由于它们无法测量，将概率分布与之关联便毫无意义。不过，正如上一节所述，我们关注的是作为经典模型膨胀的不可访问信息模型。因此，每个模型都会包含一些可访问的陈述。我们预期每个可访问陈述都能关联一个定义明确的概率。因此，为确保一致性，对于任意具有可访问性深度 d 的模型 $\tilde{}$，我们要求函数 Q 在经典子格上的限制表现为函数 $\tilde{}$，其定义为

$$\tilde{P}(y \mid x) = \begin{cases} 1 \text{ if } x \leq y \\ 0 \text{ if } x \wedge y = 1 \\ p \text{ otherwise, where } p \in (0,1), \end{cases} \tag{12}$$

使我们能够在实验上可及的陈述上恢复标准概率论，即，对于可及元素，$\tilde{P}(s) \coloneqq P(s \mid \top)$ 为正。这一限制对可容许的 $\vec{q}$ 施加了一些约束。例如，如果 s_1, s_2 是两个原子陈述，那么要求 $s_1 \vee s_2$ 在实验上可及，就相当于要求 $P(s_1 \vee s_2) = Q(s_1 \vee s_2) = q_1 + q_2 \in [0,1]$。

我们注意到，这一针对可容许的准概率赋值的限制，实际上将可达性与真值这两个概念联系了起来，而它们在逻辑格中原本是彼此分离的。从逻辑格转向准概率格时，我们放宽了真值赋值和可达性值赋值。我们将二值真值放宽为连续值，该值由格中元素的一个函数指定。同时，为了要求概率赋值与经典子格保持一致，可访问性标签（用于区分**A** 和**N**）也放宽为这样的区分：陈述由概率值还是准概率值来刻画。

在逻辑格中，真值与可达性值彼此并不相通。而在准概率格中，放宽后的真值与可达性值相互交织，形成了丰富的允许分布结构。在准概率模型中，允许构型所受的约束并非由其是否符合真值表和可达性表决定，而是由一系列不等式和等式决定，这些不等式和等式确定了为每个陈述 Q 赋值的概率或准概率性质。这些赋值通过方程(9)中的简单规则在格中传播。

我们注意到，如果将范围限定在由经典模型膨胀而来的模型上，那么在可访问性深度d这一层级上，我们总能得到一个恰当的概率向量。事实上，在层级 d 上，逻辑格中的所有可访问陈述彼此不相交，并且它们通过二元运算符 $\vee$ 组合后的结果是 $\top$，这意味着它们的概率赋值之和为 1。

进入下一节之前，我们在此重申，模型的准概率赋值完全由向量 $\vec{q}$的取值所表征。我们将 $\vec{q}$ 称为模型的状态 。

下一节将明确表征刻画允许状态 $\vec{q}$的约束。

6 准概率模型

6.1 极大表达标准模型

考虑图1面板（c）中可访问性深度为 $d = 2$ 的模型。将范围放宽至准概率格后，我们为该模型关联一个准概率向量 $\vec{q} = (q_1, q_2, q_3, q_4) = (Q(a)\ Q(b)\ Q(c)\ Q(d))$，用于描述模型的状态。按照我们的假设，在经典子格上，我们将其限制为一个概率分布。这意味着，允许的向量 $\vec{q}$ 必须满足 $Q(a \vee b) = P(a \vee b) \in [0,1]$ 和$Q(c \vee d) = P(c \vee d) \in [0,1]$。这施加了以下约束

$q_1 + q_2 \in [0,1]$ 和 $q_3 + q_4 \in [0,1]$。

在第 4 节末尾，我们看到逻辑格结构并未帮助我们对新型结构进行建模；相反，它只不过是对某个经典格的过度参数化。

对于松弛格，情况不再如此。我们确实可以将格中的其他元素限制为概率分布，而不会产生矛盾。

例5。考虑一个可访问性深度为 $d = 2$ 、关联有$\vec{q} = (Q(a)\ Q(b)\ Q(c)\ Q(d))$的模型。经典子格上、原子元素为 $a \vee b$、 $c \vee d$ 的向量$\vec{p}^{[1]} = (Q(a \vee b)\ Q(c \vee d))$ 被限制为概率向量。

在逻辑格中，我们本可以将经典子格定义在 $a \vee c$ 和 $b \vee d$ 上，而不是 定义在 $a \vee b$ 和$c \vee d$上。

在准概率模型中，我们也可以将$\vec{p}^{[2]} = (Q(a \vee c)\ Q(b \vee d))$ 限制为概率向量， $\vec{p}^{[3]} = (Q(a \vee d)\ Q(b \vee c))$ 也同样如此。选择这三个不同的概率向量会施加以下约束

$$q_i + q_j \in [0,1] \tag{13}$$

对于每个 $i, j \in \{1, 2, 3, 4\}$ 和 $i \neq j$。可以看出，这些约束并不等同于约束$q_i \in [0,1]$ 对于每个 $i \in \{1, 2, 3, 4\}$，因此 $\vec{q}$仍然是一个可能含有负分量的准概率向量。与逻辑格不同，在准概率表述中，我们可以在层级 d 上有三个可达概率向量，同时让向量 $\vec{q}$ 自由取负值，因此不会将模型约束为经典模型。该示例的图示见图 3。

沿用示例5中引入的记号，我们用方括号中的上标表示层级d 处得到的不同概率向量。我们称这些概率向量为模型的相关可达概率向量。与原始逻辑格的可访问陈述相关联的可达概率向量始终记为 $\vec{p}^{[1]}$。

在示例5 中，我们注意到，同一概率向量中的陈述组合后仍然是一个概率。例如 $p_1^{[1]} + p_2^{[1]} \in [0,1]$。来自不同概率向量的元素组合后，不一定仍是一个概率。例如 $p_1^{[1]} + p_1^{[2]} = q_1 + q_2 + q3$ 可能落在区间 $[0,1]$之外。[4]

[4]将这些性质映射到逻辑格的情形时，准概率模型的不同可达概率向量对应于不同的可达性族。在例5 中，我们会发现每个陈述都有三个可达性标签。我们将这些标签称为可达性深度1,

我们看到，对于能够被约束为概率向量一部分的陈述数量，并不存在硬性约束。对于一个具有可访问性深度 d 的一般模型，我们要问，在层级 d 上，考虑多少个概率向量 $\vec{p}^{[i]}$ 才是合理的。我们注意到，在格中要求不止一个可及概率分布，会在模型的两个性质之间产生张力：一方面，将格中的更多元素限制为概率，会提高模型的表达能力；也就是说，模型能够编码的概率分布越多，其中包含的实验相关信息就越多。同时，要求将格中更多元素限制为概率，就会对可接受的 $\vec{q}$施加更多约束；这些约束随后会传播到模型可能具有的概率分布中，进而影响模型所包含的可能的实验相关信息。我们希望自己的不可达信息理论尽可能具有表达能力，同时对 $\vec{q}$ 施加尽可能少的约束。

特别是，我们希望得到约束最少的理论，使得每个 $\vec{q}$ 都与一组唯一的可达概率向量$\{\vec{p}^{[i]}\}_i$一一对应。

为实现这一目标，我们定义最大表达标准模型。

定义 7 （最大表达标准模型（MES 模型））。一个可访问性深度为 d、其中 d 为素数的模型，是一个可访问性深度为 d 的最大表达标准（MES）模型，当且仅当它在满足以下约束的前提下，使层级 d 上的概率向量 $p^{[i]}$ 数量达到极大：层级 d 上任意两个来自不同可达概率向量的陈述，恰好共有一个原子陈述。

示例 6。考虑具有可访问性深度 $d = 2$ 的 *MES* 模型及其三个概率向量 $\vec{p}^{[1]}$、$\vec{p}^{[2]}$、$\vec{p}^{p[3]}$ （见图 3）。取 $\vec{p}^{[1]}$ 中的一个陈述，以及 $\vec{p}^{[2]}$ 或 $\vec{p}^{[3]}$中的任意其他陈述，二者的重叠至多为一个陈述。例如，取 $a \vee b$ 中的 $\vec{p}^{[1]}$ 和 $b \vee d$ 中的 $\vec{p}^{[2]}$，重叠部分仅为陈述 b。

引理 7。 一个可访问性深度为 d 的 *MES* 模型在层级 d有 $d+1$ 个概率向量。

可访问性-2、可访问性-3。 $\vec{p}^{[1]}$ 的陈述对于可访问性-1是可达的，而对于可访问性-$i \neq 1$则不可达。不同可访问性的复合不一定是可达的。复合规则不会混合可访问性：也就是说，两个可达性-1陈述的复合仍是可达性-1，而可达性-1陈述与可达性-2陈述的复合则不保证可达，因为它们分别相对于两个不同的可访问性标签可达。

定义 8（MES 模型的状态集）。对于可访问性深度为 d 的 MES 模型，MES_d 是与该模型兼容的状态集 $\vec{q}$。

$$\mathrm{MES}_d = \left\{ \vec{q} \in \mathbb{R}^{d^2} \mid \sum_{i=1}^{d^2} q_i = 1;\ p_j^{[k]} \in [0,1],\ \forall\ {}^{k=1,\ldots,d+1}_{j=1,\ldots,d} \right\} \tag{14}$$

其中$\{\vec{p}^{[k]}\}_{k=1}^{d+1}$ 是相应的概率向量。

MES 模型的定义由以下引理给出依据。

引理 8. *MES* 模型具有这样的性质：$\vec{q}$ 与层级 d上的可访问概率分布之间存在一一对应关系，且对容许态 $\vec{q}$的约束数量最少。

6.2 准概率格中的假设 1与不可访问性度量

本节探讨从逻辑格放宽到准概率模型时会遇到的一个疑问。我们决定为格的原子元素赋予一个准概率向量，使其通过传播规则转化为层级 d上相应概率向量的概率赋值。疑问在于，概率向量集合是准概率向量集合的子集。因此，如果我们无法区分赋予格的原子元素的是准概率向量还是概率向量，那么在准概率模型中，如何才能合理地表达可访问信息与不可访问信息之间的区别?

为了一致地解决这个问题，我们必须回顾并记住假设 **1** 和$\mathcal{IH}$背后的直觉。我们想要描述这样一种系统：其中每个实验上可访问的陈述都由 d 实验上不可访问的陈述组成。在逻辑格中，这一事实通过标记 $\mathbf{A}/\mathbf{N}$以及可达性赋值如何依照可及性表所规定的规则沿格传播，得到了清晰的体现。以图 (4,2)中的模型为例。已知陈述 $a \vee b$ 为真，并不能给我们任何线索来推断两个原子陈述 a 或 b 中的哪一个为真。对于 d 原子元素，我们完全无知。要求层级 d 以下的所有陈述在实验上不可访问，可以理解为要求至少对 d 个元素具有完全不确定性。

我们可以用概率的语言来表述同一个要求：层级 d 的概率分布所包含的总信息，不应多于在我们对至少完全不确定 d 个原子元素（共 d^2 个）时所拥有的信息。我们如何衡量这种信息？为回答这一问题，我们将考察一些示例，以确定适用于此任务的信息度量应具备哪些关键属性。

例 7。考虑带有 $d = 2$以及特定赋值 $\vec{q} = (1,0,0,0)$ 的最大纠缠态模型。向量 $\vec{q}$ 是一个正规的准概率向量，而 $\vec{p}^{[1]} = (1,0)$、$\vec{p}^{[2]} = (1,0)$、$\vec{p}^{[3]} = (1,0)$ 是正规的概率向量。如此定义的 $A\ \vec{q}$ 意味着陈述 "a"具有确定性，从而在逻辑格中"a" 具有一个真值。这是被禁止的。

例8。考虑带有 $d = 2$ 的 *MES* 模型，以及特定赋值 $\vec{q} = (1/2, 1/2, 0, 0)$。向量 $\vec{q}$ 是一个正规准概率向量，而$\vec{p}^{[1]} = (1,0)$、$\vec{p}^{[2]} = (1/2, 1/2)$、$\vec{p}^{[3]} = (1/2, 1/2)$ 是正规概率向量。如此定义的 $A\ \vec{q}$ 在两个元素上具有完全不确定性，因此符合假设 *1* 的精神。

例9。考虑带有 $d = 2$ 的 *MES* 模型，以及特定赋值 $\vec{q}^{(2)} = (1/2, 1/2, 0, 0)$；再考虑带有 $d = 3$ 的 *MES* 模型，以及赋值 $\vec{q}^{(3)} = (1/3, 1/3, 1/3, 0, 0, 0, 0, 0, 0)$。在第一个模型中，我们在2 个元素上具有完全不确定性；在第二个模型中，我们在3 个元素上具有完全不确定性。现在考虑这两个模型的复合。复合状态由 $\vec{q}^{(2)}$ 与 $\vec{q}^{(3)}$ 复合成向量$\vec{q}^{(6)} = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6, 0, 0, \ldots, 0)$来描述。在复合模型中，我们在6个元素上具有完全不确定性。可见，不确定性是乘法性的。

在前面的示例中，我们强调了信息度量的一些理想属性。我们特别注意到，我们所寻找的信息度量在系统复合下是乘法的。在附录 D.1 中，我们引入了单参数族 不可访问性度量$\mathcal{X}_c$ ： $\Delta_D \to \mathbb{R}_+$，其中 $c \in \mathbb{R}_+$ 且 $c \neq 1$，将概率单纯形的 D维概率向量映射 Δ_D 为一个正实数。

不可访问性度量$\mathcal{X}_c(\vec{q})$ 对概率向量 $\vec{p}$ 的作用定义为

$$\mathcal{X}_c(\vec{p}) = \frac{1}{\sqrt[c-1]{\sum_{i=1}^{D} p_i^c}}. \tag{15}$$

不可访问性度量是概率向量的一个函数，用来量化我们对多少个元素完全无知，或者更准确地说，有多少个元素无法为我们的知识所获取。对于概率向量 $\vec{p} \in \Delta_D$，有如下结论：对每个 c, $1 \leq \mathcal{X}_c(\vec{p}) \leq D$。

香农熵与不可访问性度量的主要区别在于，后者具有乘法性，而香农熵具有加法性。考察示例7和8，可得对每个 c，$\mathcal{X}_c(1,\ 0,\ 0,\ 0) = 1$ 且 $\mathcal{X}_c(1/2,\ 1/2,\ 0,\ 0) = 2$。满足我们假设的 $\mathcal{X}_c$ 族的推导见附录D.1，而 $\mathcal{X}_c$ 与其他信息度量的比较见附录D.4。

在不可访问信息理论中，我们关注的是如何度量准概率向量的不可访问性。因此，在附录D.2中，我们将不可访问性度量 $\mathcal{X}$ 的定义域扩展到 MES 模型的准概率态空间。为了满足与 MES 模型结构的相容性，我们发现，在不可访问信息理论中，存在一个独立存在的唯一不可访问性度量，即作用于准概率向量的不可访问性度量 $\boldsymbol{X}$ $\vec{q}$ 如下：

$$\mathcal{X}(\vec{q}) = \frac{1}{\sum_i q_i^2}. \tag{16}$$

为 $c > 2$ 定义的所有不可访问性度量都取决于不可访问性度量 $\mathcal{X}$。

我们终于可以陈述准概率模型的假设 1 了。

准概率的假设 1

可访问性深度为 d 的 MES 模型的容许态 $\vec{q}$，其不可访问性下界为 d，即 $\mathcal{X}(\vec{q}) \geq d$。

我们已经将假设 1 表述为准概率语言中的形式，从而编码了对可访问信息可访问性深度这一概念的直观理解。最终，我们可以构建不可访问信息理论。（为方便起见，逻辑格与准概率格之间的示意性比较见附录3中的表C。）

7 不可访问信息理论

不可访问信息理论是遵循假设 1 的最大纠缠态模型。最大纠缠态模型的假设 1（见定义 6.2）限制了容许态 $\vec{q}$，使其不可访问性 $\mathcal{X}(\vec{q})$ 大于模型的可访问性深度 d。对于可访问性深度为 d 的最大纠缠态模型，不容许满足 $\vec{q}$ $\mathcal{X}(\vec{q}) \lneq d$的状态。对于每个可访问性深度 d，我们将容许态的集合记为

$$\mathcal{Q}_d = \left\{\vec{q} \in \mathrm{MES}_d \mid \mathcal{X}(\vec{q}) \geq d\right\}. \tag{17}$$

我们将使系统知识最大化的状态 $\vec{q} \in \mathcal{Q}_d$ 定义为该理论的纯态。这些状态使不可访问性最小。

定义 9（纯态）。可访问性深度为 d 的 MES 模型的不可访问信息理论的纯态，是所有满足 $\mathcal{X}(\vec{q}) = d$ 的容许态 $\vec{q}$。

由于附录 11中的引理 D.3，我们知道 $\mathcal{X}$ 是定义在 $\mathcal{Q}_d$上的拟凹函数，因此，任意纯态的凸组合都是一个不可访问信息大于或等于 d 的容许态。

8 可访问性深度为 d=2 的不可访问信息理论

本节研究不可访问信息理论中最简单的模型。我们研究一个可访问性深度为 $d = 2$ 的 MES 模型。该模型由 $D = d^2 = 4$ 个原子陈述（a，b，c，d）表征，我们为其关联一个形式为拟概率向量 $\vec{q} = (Q(a)\ Q(b)\ Q(c)\ Q(d)$ 的状态。由于我们考虑的是一个可访问性深度为 $d = 2$ 的 MES 模型，在层级 $d = 2$，根据引理 7，我们有 $d+1 = 3$ 个关联的可达概率向量。按照引理 7 证明中给出的构造，这些关联的可达概率向量为：

$$\begin{aligned}
\vec{p}^{[1]} &= (P(a \vee b), P(c \vee d)) = (q_1 + q_2, q_3 + q_4),\\
\vec{p}^{[2]} &= (P(a \vee c), P(b \vee d)) = (q_1 + q_3, q_2 + q_4),\\
\vec{p}^{[3]} &= (P(a \vee d), P(b \vee c)) = (q_1 + q_4, q_2 + q_3).
\end{aligned} \tag{18}$$

从图中对格表示的检查3 可以看出，这些概率分布如何遵循 MES 模型的定义，即在层级 d，来自不同概率向量的任意两个陈述都恰好共有一个原子陈述。显然，可访问性深度 $d = 2$ 是一个特例，因为 $d = 2$ 是唯一一个使层级 d 上的所有陈述都属于某个可及概率分布的可访问性深度。

现在我们要刻画容许态的集合 $\mathcal{Q}_2$ $\vec{q}$。$\vec{q}$ 的约束为：

- 准概率约束：$\vec{q} \in \mathbb{R}^4$ 其中 $\sum_{i=1}^4 q_i = 1$。参见第 5 节。
- 可达性约束：$q \in MES_2$ 意味着所有 $\{\vec{p}^{[i]}\}_{i=1}^{d+1}$ 都必须是概率向量。对于情况 $d = 2$ 这意味着 $(q_i + q_j) \in [0, 1]$ 对所有 $i \neq j$。参见定义 7和 8。
- 不可访问性约束：$\vec{q}$ 必须遵守假设 1，即 $\mathcal{X}_2(\vec{q}) \geq 2$。参见定义 6.2。

已于 2024-08-30接收，点击标题进行验证。以 CC-BY 4.0 协议发布。

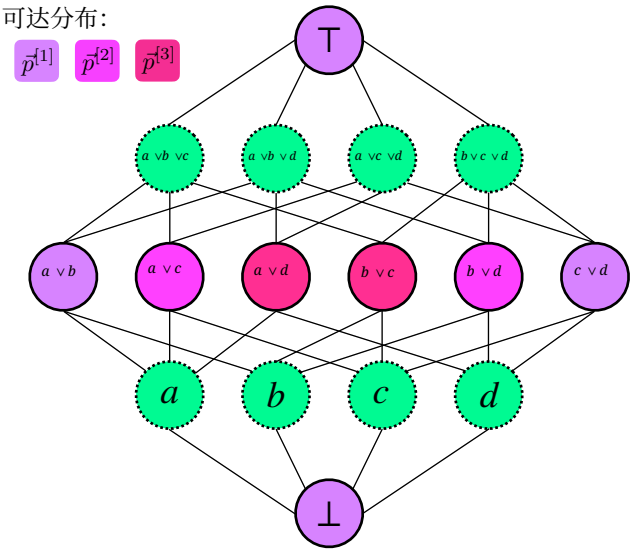

图3：可访问性深度为 $d = 2$ 的 MES 模型的格表示。层级 d 的 $d + 1 = 3$ 可访问分布包含层级 d 的整个陈述集。

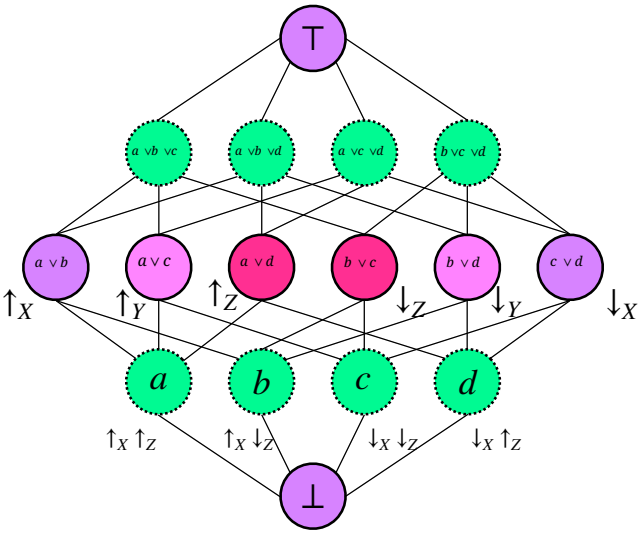

图 4：量子比特与最大纠缠态的可达性深度为 2。

因此，$\mathcal{Q}_2$ 是所有满足以下条件的 $\vec{q} \in \mathbb{R}^4$ 的集合:

$$\begin{cases} \sum_{i=1}^{4} q_i = 1, & (19) \\ (q_i + q_j) \in [0,1]\ ,\ \forall i \neq j, & (20) \\ \mathcal{X}_2(\vec{q}) \geq 2. & (21) \end{cases}$$

方程 (19) 与方程(21) 结合可推出方程 (20)，因此 $\mathcal{Q}_2$ 是所有满足以下条件的 $\vec{q} \in \mathbb{R}^4$ 的集合

$$\begin{cases} \sum_{i=1}^{4} q_i = 1 \\ \mathcal{X}_2(\vec{q}) \geq 2, \end{cases} \quad (22)$$

or

$$\begin{cases} \sum_{i=1}^{4} q_i = 1 \\ \sum_{i=1}^{4} q_i^2 \leq \frac{1}{2}, \end{cases} \quad (23)$$

这里我们将约束 $\mathcal{X}$

– $_2\ \vec{q}\ \geq 2 \quad \sum_{i=1}^{4} q_i^2 \leq ()$

1
2
改写为 。这是四维空间中一个超平面与一个实心超球的交集。该交集是三维空间中的一个实心球，因此集合 $\mathcal{Q}_2$ 是一个凸集，其极点位于球面上，也就是所有满足条件 (23) 的向量；球面上的向量是该理论的纯态，因为它们的不可访问性均为 2。

8.1 量子比特

如果我们在本节中暂时偏离不可访问

信息理论，转而讨论量子系统，这并不令人意外。类比之处逐渐显现，我们需要明确地加以讨论。

我们考虑一个量子比特系统。状态由其密度矩阵 $\rho \in \mathcal{H}_2 = \mathbb{C}^2$ 描述。在布洛赫球形式主义中，可以考虑与沿三个相互正交的轴之一测量自旋相关联的可观测量 x, y, z。对于每个轴 x, y, z，定义两个投影算符 $\{\Pi_{\uparrow}^{[\alpha]}, \Pi_{\downarrow}^{[\alpha]}\}_{\alpha=x,y,z}$，使得 $P_{\uparrow}^{[\alpha]} = \mathrm{Tr}[\Pi_{\uparrow}^{[\alpha]}\rho]$ 和 $P_{\downarrow}^{[\alpha]} = \mathrm{Tr}[\Pi_{\downarrow}^{[\alpha]}\rho]$ 分别是沿轴 α测量状态ρ得到向上或向下结果的概率。定义以下四个矩阵

$$\begin{aligned} \hat{a} &= \frac{1}{4}\left(\mathbb{I} + \sigma_x + \sigma_y + \sigma_z\right), \\ \hat{b} &= \frac{1}{4}\left(\mathbb{I} + \sigma_x - \sigma_y - \sigma_z\right), \\ \hat{c} &= \frac{1}{4}\left(\mathbb{I} - \sigma_x + \sigma_y - \sigma_z\right), \\ \hat{d} &= \frac{1}{4}\left(\mathbb{I} - \sigma_x - \sigma_y + \sigma_z\right), \end{aligned} \quad (24)$$

其中 σ_x、σ_y 和 σ_z 是泡利矩阵。可以看出，

$$\Pi_{\uparrow}^{[x]} = \hat{a} + \hat{b}, \quad \Pi_{\uparrow}^{[y]} = \hat{a} + \hat{c}, \quad \Pi_{\uparrow}^{[z]} = \hat{a} + \hat{d}, \quad (25)$$
$$\Pi_{\downarrow}^{[x]} = \hat{c} + \hat{d}, \quad \Pi_{\downarrow}^{[y]} = \hat{b} + \hat{d}, \quad \Pi_{\downarrow}^{[z]} = \hat{b} + \hat{c}. \quad (26)$$

如果我们将 $P_i^{[\alpha\wedge\beta]}{}_{,j}$ 与 $\alpha, \beta \in \{x, y, z\}$ 和 $i, j \in \{\uparrow, \downarrow\}$称为沿两个轴 α, β 分别沿方向 i 和 j同时测量量子比特所得的“概率”，

则可作如下解释

$$P_{\uparrow\uparrow}^{[x\wedge z]}=\mathrm{Tr}[\rho\hat{a}],\qquad P_{\uparrow,\downarrow}^{[x\wedge z]}=\mathrm{Tr}[\rho\hat{b}], \tag{27}$$

$$P_{\downarrow,\downarrow}^{[x\wedge z]}=\mathrm{Tr}[\rho\hat{c}],\qquad P_{\downarrow,\uparrow}^{[x\wedge z]}=\mathrm{Tr}[\rho\hat{d}]. \tag{28}$$

$$\tag{29}$$

^ ^我们注意到，$\hat{a},\hat{b},\hat{c},\hat{d}$ 是相对于希尔伯特–施密特内积而言的$\mathcal{H}_2$ 的一组正交基。因此，状态ρ 完全由向量表征

$$\vec{q}(\rho)=(\mathrm{Tr}[\rho\hat{a}],\mathrm{Tr}[\rho\hat{b}],\mathrm{Tr}[\rho\hat{c}],\mathrm{Tr}[\rho\hat{d}]), \tag{30}$$

and

$$\sum_{i=1}^{4}q(\rho)_i=\mathrm{Tr}[\rho]=1. \tag{31}$$

我们把 $P_i^{[\alpha\wedge\beta]}{}_{,j}$ 称为带引号的“概率”，因为它们实际上是准概率。对两个不同自旋的同时测量不是一个可观测量，而且 $P_i^{[\alpha\wedge\beta]}{}_{,j}$ 可以为负。

此外，我们注意到，对于量子比特的例子

$$\mathcal{X}_2(\vec{q}(\rho))=\frac{2}{Tr[\rho^2]}. \tag{32}$$

因此，对于一般的纯量子比特态 $|\psi\rangle$ ，与之关联的一般准概率向量

$$\vec{q}_{|\psi\rangle}=\left(\langle\psi|\,\hat{a}\,|\psi\rangle\,,\langle\psi|\,\hat{b}\,|\psi\rangle\,,\langle\psi|\,\hat{c}\,|\psi\rangle\,,\langle\psi|\,\hat{d}\,|\psi\rangle\right) \tag{33}$$

由一个固定的不可访问性$\mathcal{X}_2(\vec{q}_{|\psi\rangle})=2$ 表征，因为 $\mathrm{Tr}(|\psi\rangle\langle\psi|)^2=1$。对于一般的量子比特态 ρ，与之关联的准概率向量 $\vec{q}(\rho)$由一个受约束的不可访问性$\mathcal{X}_2(\vec{q}(\rho))\geq 2$表征，因为 $\mathrm{Tr}\rho^2\leq 1$。

我们看到，描述量子比特的厄米矩阵所满足的约束，即

$$\begin{cases}\mathrm{Tr}[\rho]=1\\ \mathrm{Tr}[\rho^2]\leq 1.\end{cases} \tag{34}$$

直接映射为约束(22) ，这些约束刻画了可访问性深度为 $d=2$的不可访问信息理论。

我们发现，在不可访问信息理论中，量子比特状态可以表述为可访问性深度为 $d=2$ 的 MES 模型状态。图4展示了图3 中面板的对应版本，以量子态的形式写出。特别是， $\vec{q}$ （由可访问性深度为 $d=2$ 的不可访问信息理论所允许）与描述量子比特的量子态之间存在一一对应关系。

这一结果可以在量子框架 [15]的框架下进行解释。从对量子比特的描述出发，我们可以选取任意一组由 4 个算符组成的特定框架，并将量子比特态描述为一个由（准）概率 $\vec{\mu}$构成的向量。可以看到，选取方程（24）[5], 所示的特定框架后，框架表示向量 $\vec{\mu}=\vec{q}$ 所受的约束 $\vec{\mu}$ 可自然地解释为对该向量不可访问性的约束，因此μ必须满足 X（ ≥ 2）。我们可以将这些约束与选择其他框架时得到的约束进行比较。例如，选取对称信息完备正算子值测度作为框架（如量子贝叶斯主义对量子力学的重构所做的那样），情况就有所不同。向量 $\vec{\mu}$ 将成为一个概率向量 $\vec{\mu}=\vec{p}$，因此不可访问性的界限也会变为关于 $\vec{\mu}$的另一个函数的界限（ [18]的方程（7））。在对称信息完备正算子值测度框架表示中，描述量子比特态的 $\vec{\mu}$集合所对应的界限形式，似乎没有自然的解释。引用 [18]（方程（9）之后的第二段）：“[.....] 量子态用这些术语表示时，竟变成了如此奇怪的一组！除了我们已经熟知通常形式的完整量子理论之外，还有什么能解释这一点呢？”[6]。在不可访问信息理论的形式体系中，我们无法推导出 [18], 的方程（8），但 [18]的方程（7）则成为不可访问性的下界。考虑不可访问信息理论时，部分约束可以通过逻辑推导出来，看起来也不那么奇怪[7]。

我们以最后一个类比结束本节：注意到第 2 层的可访问概率分布集合对应于量子比特的一组互相无偏算子。

9 可访问性深度为 d=3 的不可访问信息理论

该模型由 $D=d^2=9$ 个原子陈述（a，b，c，d，e，f，g，h，i）刻画，我们为其关联一个形式为准概率向量的状态 $\vec{q}$ =($Q(a)$ $Q(b)$ $Q(c)$ $Q(d)$ $Q(e)$ $Q(f)$ $Q(g)$ $Q(h)$ $Q(i)$)。由于我们考虑的是 MES 模型，在第 $d=3$ 层，我们有 $d+1=4$ 个可达概率向量。

[5]我们可以将其称为费曼–伍特斯框架，因为 Feynman 在 [16] 中讨论准概率时提到了它，而 Wootters 在 [17]中将其推广到了任意素数维度。

[6]这条陈述指的是刻画所有维度下量子状态的约束集合，而不仅仅是量子比特的情况。具体而言，在处理维度大于 2 的系统时，还必须考虑一项额外约束（ [18]中的式 (8)）。我们的研究只涉及量子比特，因此仅考虑这条陈述中所提约束的一半（较简单的一半）。

[7]我们再次提醒，关于不可达信息的理论并非对量子力学的试探性重构。

已于 Quantum 2024-08-30 接收，点击标题进行核验。以 CC-BY 4.0 许可协议发布。

在重新标记的自由范围内，这些是

$$\vec{p}^{[1]} = (P(a \vee b \vee c), P(d \vee e \vee f), P(g \vee h \vee i)), \quad (35)$$

$$\vec{p}^{[2]} = (P(a \vee d \vee g), P(b \vee e \vee h), P(c \vee f \vee i)), \quad (36)$$

$$\vec{p}^{[3]} = (P(a \vee e \vee i), P(b \vee f \vee g), P(c \vee d \vee h)), \quad (37)$$

$$\vec{p}^{[4]} = (P(a \vee f \vee h), P(b \vee d \vee i), P(c \vee e \vee g)). \quad (38)$$

在可访问性深度 $d = 3$，并非所有层级 d 的陈述都属于某个可达概率分布。事实上，在层级 $d = 3$，可以找到 $\frac{d^2!}{d!\,(d^2-d)!} = 84$个陈述，而只有 $d(d+1) = 12$ 会进入可达概率分布。与情形 $d = 2$ 类似，此处容许态集合 $\mathcal{Q}_3$ $\vec{q} \in \mathbb{R}^9$的特征如下：

$$\begin{cases} \sum_{i=1}^{9} q_i = 1, & (39) \\ (q_i + q_j + q_k) \in [0,1]\ ,\ \forall i \neq j, i \neq k, j \neq k, & (40) \\ \sum_{i=1}^{9} q_i^2 \leq \frac{1}{3}. & (41) \end{cases}$$

与 $d = 2$ 的情况不同，这里三个条件缺一不可。事实上，准概率向量

$$\tilde{q} = \left(\frac{1}{4} - \frac{1}{24}\left(3+\sqrt{33}\right), 0, 0, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, 0, \frac{1}{24}\left(3+\sqrt{33}\right)\right), \quad (42)$$

满足条件 (39) 和条件 (41)，但不满足条件 (40)。

10 展望与结论

很久以前，在量子力学发展之际，逻辑学领域曾围绕真陈述与可证明陈述之间的区别展开讨论 [19]。我们能否为所有真陈述找到证明？所有真陈述都能被证明吗？在众多学者中，库尔特·哥德尔通过提出著名的不完备性定理，为这些问题给出了答案，彻底证实了有些陈述虽为真，却无法被证明[8]。我们的研究受到这些成果的启发，在逻辑与物理学之间建立了一种引人深思的类比。我们设想，实验过程可以作为证明有关现实的陈述的一种工具。如果逻辑证明是运用逻辑工具的算法，那么实验就是运用物理世界工具的算法。正如有些陈述无法通过逻辑证明得到证明，我们设想也可能存在无法通过实验检验得到证明的陈述。

这一想法显然受到了量子力学启示的影响。事实上，自诞生以来，量子力学一直在表明，我们的认知必然存在某些局限。这里需要说明一点，“通过实验检验得到证明”这一说法可能不太容易理解，因为它似乎与关于可证伪性的论述中使用的措辞相冲突。我们所说的是这样的陈述：“做这件事 和 那件事，我就会在这个屏幕上得到这些数字。”所谓通过实验检验来证明这一陈述，指的是执行前面所说的这件事 和那件事，并核实我们是否在屏幕上得到了预言的数字。

尽管这正是本研究背后的理由，我们还是决定聚焦于较少涉及哲学的部分，单纯研究$\mathcal{IH}$的后果，并将可访问与不可访问陈述之间的区别形式化。可访问陈述是指其真值可以通过实验评估的陈述。我们决定用经典命题逻辑的语言将$\mathcal{IH}$ 形式化，主要有两个原因。第一个原因是简洁；我们认为经典命题逻辑是处理陈述时最简单的一套规则之一。第二个原因是我们熟悉大量关于从逻辑推导概率论的文献，这为我们提供了一个机会：在我们的框架内推导概率论，而不是从外部将其附加进来（例如，允许出于操作层面的理由对状态作凸组合）。我们进一步假设，可访问知识具有一个可访问性深度 d。每个可访问陈述由 d 个不可访问陈述组成。这意味着，即使我们能够评估某个陈述的真值，对于 d 个不可访问陈述，我们仍然会感到不确定。可访问性深度d 刻画了一个基本的“无知”程度。由这些基本假设可以推导出丰富的结构。令人惊讶的是，我们发现可访问性深度与模型中需要考虑的陈述总数之间存在关联。对于可访问性深度 d，我们至少需要考虑 d^2 个陈述。

基于概率论基础方面的结果，我们得以放宽并扩展不可访问信息理论的逻辑模型，将其拓展为准概率模型。我们仅以一致性准则为基础，从第一原理推导出准概率结构，说明当允许存在不可访问陈述时，准概率描述是自然的首选语言。[9] 我们的构造在一定程度上让人想起，部分关于

[8] 给定一个固定的形式系统。

[9] 在第4 节末尾，我们观察到，添加可达性标签似乎不过是过度参数化，只是为每个陈述添加一个带有相应解释的新标签。然而，正是对这种

隐变量模型，并表明，对于隐变量相关知识的建模，准概率是一个很好的工具。事实上，我们的方法引出了以下问题：为什么要为隐变量赋予概率？我们如何为这一特定限制提供依据？对于无法访问、处于隐藏状态的事物，我们无法据此建立频率或信念。在不可访问信息理论的准概率扩展中，我们发现有必要定义一种不确定性度量。为了量化不可访问程度并保留知识的可访问性深度这一概念，我们需要定义不可访问性度量。给定一个准概率向量，不可访问性度量可以量化其知识深度。

不可访问信息理论简单地定义为一种具有两个约束的准概率理论：当准概率局限于可访问陈述时，预期其表现得像规范的概率；并且允许的准概率向量具有有界的不可访问性值。除此之外别无他物（几乎如此）。

鉴于我们的前提，当我们发现不可访问信息理论与量子力学形式主义之间的联系时，并未感到意外。事实上，我们发现，量子比特的结构可以从不可访问信息最简单模型的状态结构中推导出来[10]。尽管我们认为这一结果是次要的（本工作的目的是研究不可访问信息理论本身，而不是推导量子比特的结构），但同一结果也可以被重新表述为进一步发展和研究不可访问信息理论的额外动力。

总而言之，不可访问信息理论虽然建立在简单的假设之上，却展现出丰富的结构。在发展这一理论的过程中，我们得到了有趣的结果。我们从逻辑原理推导出准概率，为将其用作工具确立了基本依据。我们仅从少数几条原理出发，推导出不可访问性度量这一族——一组乘法型信息度量，它们以许多不同的名称和形式零散地出现在文献中（见 [20, 21, 22, 23, 24]）。我们恢复了结构-

量子比特作为不可访问信息的最简单模型的结构，也就是说，每个可访问陈述始终由两个不可访问陈述复合而成的模型。这是一段愉快的旅程，但尚未抵达终点。可以探索的方向有很多。从抽象的角度看，我们注意到，不可访问信息理论似乎没有纳入“对象”这一概念。所谓“对象”，是指陈述所指涉的事物。就目前而言，该理论只讨论彼此之间由命题代数以及理论的约束联系起来的陈述。理想情况下，对“对象”的定义应能通过复合来定义陈述的“变换”。将这一概念限定在量子力学的情形下就容易理解了：对于量子比特，这些变换就是幺正变换，它们能够将关于某一方向自旋的陈述变为关于任意方向自旋的陈述。另一个方向是妥善地表述状态更新过程，即从获取新信息时开始的状态更新过程。还有一个方向是仔细研究可访问性深度高于 $d = 2$ 的模型。

致谢。首先，我要感谢 Matteo Scandi 这位最棒的“饭搭子”，我们进行了许多愉快的讨论，他帮助我在开展这项工作时始终保持热情。接下来，我要感谢 Francesco Buscemi，他在我开展这项工作的部分内容期间热情地接纳我加入他的研究小组，为这项工作的不同部分提供了启发，并与我分享了他对概率论基础问题的热忱。最后，我要感谢 Markus P. Müller。在我展示这项工作的雏形版本后，他提出了恰当的问题，由此促成了不可访问性度量的定义。我们感谢审稿人提出的诸多建议和修改意见，它们极大地改进了这项工作。“不可及性假设”这一名称由审稿人提出。本项目获得西班牙政府（FIS2020-TRANQI 和 Severo Ochoa CEX2019-000910-S）、Fundacio Cellex、Fundació Mir-Puig、加泰罗尼亚自治区政府（SGR 1381 和 CERCA 计划）的资助。周长理论物理研究所的研究由加拿大政府通过加拿大创新、科学与经济发展部，以及安大略省政府通过研究、创新与科学部提供支持。

逻辑格中可访问陈述与不可访问陈述之间的区别，对于确定如何放宽逻辑二值赋值至关重要，正如附录 B.3.3所述。

[10]我们在此指出，量子比特的结构可以（历史上也确实如此）从许多不同的原理体系出发，以多种不同方式推导出来。因此，量子比特可以被视为一种基本结构。基于同样的原因，在此阶段，能够从第一性原理构造出量子比特，可以看作是一种风格练习。

参考文献

[1] Chirag Dhara、Gonzalo de la Torre 和 Antonio Acín。“观测到的随机性可以被认证为完全内禀的吗？”《物理评论快报》**112**，100402 (2014年)。

[2] Antonio Acín 和 Lluis Masanes。“量子物理学中的认证随机性”。《自然》540，213–219（2016年）。[3] Miguel Herrero-Collantes 和 Juan Carlos Garcia-Escartin。“量子随机数生成器”。《现代物理评论》89，015004（2017年）。[4]Gabriel Senno、Thomas Strohm 和 Antonio Acín。“量子测量内禀随机性的量化”。《物理评论快报》131，130202（2023年）。[5] Robert W. Spekkens。“量子态认知观的证据：玩具理论”。《物理评论A》75，032110（2007年）。[6] Robert W. Spekkens。“准量子化：具有认知限制的经典统计理论”。载于 Giulio Chiribella 和 Robert W. Spekkens 编，《量子理论：信息基础与补论》。《物理学基础理论》。施普林格·荷兰，多德雷赫特（2016年）。[7] C. Piron。“量子物理学基础”。W. A. Benjamin 高级图书计划。（1976年）。[8] K.H. Knuth。“格上的赋值及其在信息论中的应用”。载于《2006年IEEE国际模糊系统会议》。第217–224页。（2006年）。[9] Kevin H. Knuth。“从序关系推导定律”。《美国物理联合会会议论文集》707，204–235（2004年）。[10]Kevin H. Knuth。“格对偶性：概率与熵的起源”。《神经计算》67，245–274（2005年）。[11] R. T. Cox。“概率、频率与合理预期”。《美国物理学杂志》14，1–13（1946年）。[12] Richard T. Cox。“可能推理的代数”。J. Hopkins。（1961年）。[13]E. T. Jaynes。“概率论：科学的逻辑”。剑桥大学出版社。剑桥（2003年）。[14] E. T. Jaynes。“大脑如何进行似然推理？”载于 Gary J. Erickson 和 C. Ray Smith 编，《科学与工程中的最大熵与贝叶斯方法：基础》。《物理学基础理论》。施普林格·荷兰，多德雷赫特（1988年）。[15] Christopher Ferrie 和 Joseph Emerson。“框架希尔伯特空间：悬置量子理论的准概率图像”。《新物理学杂志》11，063040（2009年）。[16] Richard P. Feynman。“负概率”。载于 Basil Hiley 和 F. David Peat 编，《量子启示：纪念大卫·玻姆文集》。劳特利奇（1987年）。[17] William K Wootters. “有限状态量子力学的维格纳函数表述”。《物理学年鉴》176, 1–21 (1987年)。[18] Marcus Appleby、Christopher A. Fuchs、Blake C. Stacey 和 Huangjun Zhu. “引入 Qplex：量子理论的全新舞台”。《欧洲物理学杂志 D》71, 197 (2017年)。[19] Ernest Nagel 和 James Roy Newman. “哥德尔证明”。Taylor & Francis. (2005年)。[20] M. O. Hill. “多样性与均匀度：一种统一记号及其后果”。《生态学》54, 427–432 (1973)。[21] Roseli Pellens 和 Philippe Grandcolas. “生物多样性保护与系统发育系统学：在灭绝危机中守护我们的进化遗产”。施普林格. (2016年)。[22] Sandu Popescu、Anthony J. Short 和 Andreas Winter. “纠缠与统计力学的基础”。《自然·物理》2, 754–758 (2006年)。[23] Joe Dunlop、Oliver Cohen 和 Anthony J. Short. “平均意义上的本征态热化”。《物理评论 E》104, 024135 (2021)。[24]Časlav Brukner 和 Anton Zeilinger. “量子测量中的操作不变信息”。《物理评论快报》83, 3354–3357 (1999)。[25] J. Aczél. “函数方程及其应用讲义”。Academic Press. (1966)。[26] Ariel Caticha. “熵物理学”。未发表。（未完成）。[27] Ariel Caticha. “理性信念的量化”。《美国物理联合会会议论文集》1193, 60 (2009年)。[28] Hans Reichenbach. “概率理论”。加利福尼亚大学出版社. (1971)。[29] Myron Tribus. “理性描述、决策与设计”。Elsevier Science & Technology. (1969)。[30]C. Ray Smith 和 Gary J. Erickson. “概率理论与结合律方程”。载于 Paul F. Fougère 编，《最大熵与贝叶斯方法》。《物理学基础理论》。施普林格·荷兰，多德雷赫特 (1990)。[31] N. H. Abel. “关于两个独立变量 x 和 y（如 f(x, y)）的函数的研究，这些函数具有如下性质：f(z, f(x,y)) 是 z、x 和 y 的对称函数。”《纯粹与应用数学杂志》1826, 11–15 (1826)。[32]János Aczél. “结合律方程再探”。《美国物理联合会会议论文集》707, 195–203 (2004年)。[33]Claudi Alsina、Frank Maurice 和 Berthold

Schweizer. “结合函数：三角模与 Copula”。世界科学出版社. (2006年)。[34] Daniel A. Klain 和 Gian-Carlo Rota. “几何概率导论”。剑桥大学出版社. (1997)。[35] L. Losonczi. “关于次数为 α” 的熵的刻画。《Metrika》**28**, 237–244 (1981)。[36] P. L. Kannappan 和 P. K. Sahoo. “关于一个与和形式非可加性相关的函数方程

定义域上的非加性信息度量。I.”. Kybernetika **22**, 268–275(1986).

[37] 帕拉尼亚潘·坎纳潘。《带有应用的函数方程与不等式》。 施普林格科学与商业媒体。（2009年）。

[38] Scott Aaronson：《量子力学是理论空间中的一座孤岛吗？》（2004年）arXiv:quant-ph/0401062。

A 证明s

为了简化语言，我们将由二元运算符 $s_i \vee s_j \vee \cdots \vee s_k$ 连接的一组陈述 $\vee$ 与其索引集合 $(i, j, \ldots, k)$ (参见第 B.1节) 对应起来。因此，陈述复合规则映射为集合的并集与交集规则，如下例所示

$$s_i \leftrightarrow (i), \tag{43}$$
$$s_i \vee s_j \leftrightarrow (i, j) \tag{44}$$
$$(s_i \vee s_j) \wedge (s_j \vee s_k) = s_j \leftrightarrow (i, j) \cap (j, k) = (j). \tag{45}$$

引理 1. 如果 $\top$ 是 $\boldsymbol{N}$，那么格的所有陈述都是 $\boldsymbol{N}$。

证明。 令 $\top$ 的可达性为$\mathbf{N}$。在格中选取一个陈述 s_i 。如果 s_i 是$\mathbf{A}$ ，那么 $\neg s_i$ 就是$\mathbf{A}$。这意味着 $\top = s_i \vee \neg s_i$ 是$\mathbf{A}$。这是一个矛盾。 □

引理2。 对于每个模型(D, d)，由可容许配置中层级 d的两个可达陈述所包含的原子陈述集合，不得有任何共同元素。

证明。考虑模型（D, d）。假设（$i_1, i_2, \ldots, i_d$）和（$j_i, j_2, \ldots, j_d$）是层级d的两个可达陈述，并且（$i_1, i_2, \ldots, i_d$）$\cap$（$j_i, j_2, \ldots, j_d$）$=$（$k_1, \ldots, k_{r \lneq d}$）（我们施加 $r \neq d$，否则这两个陈述将完全相同）。根据表2中的可访问性传播规则，层级 r 的陈述（$k_1, \ldots, k_{r \lneq d}$）必须是可达的，但模型的可达性深度 d 禁止层级 d 以下的任何陈述成为可达陈述。因此，在层级 d ，所有可达陈述都不得相交。 □

引理 3. 对于每个模型 (D, d)，理想配置中层级 d 上的可访问陈述为 $\frac{D}{d}$ ，如果 $\mathrm{mod}(D, d) = 0$；如果 $\mathrm{mod}(D, d) \neq 0$，则为 $\lfloor \frac{D}{d} \rfloor = 1$。

证明。考虑一个一般的 (D, d) 模型。原子元素集合为 $\{s_1, \ldots, s_D\}$。
根据引理 2，在层级 d 上，我们至多可以考虑 $\lfloor \frac{D}{d} \rfloor$ 个可访问陈述。在 $D = md$的情况下，若$m \in \mathbb{N}$，可访问陈述的数量最大可达 m。这类模型是由含有 $m = \frac{D}{d}$ 个原子元素的经典模型膨胀而成。

我们需要证明当 $\mathrm{mod}(D, d) \neq 0$ 时，该引理成立。首先，我们证明在层级 d上不可能有超过 $\lfloor \frac{D}{d} \rfloor$个陈述可访问。我们将证明，层级 d 上的 $\lfloor \frac{D}{d} \rfloor$ 个陈述不可能都设为可访问，因为这样做会推出某个低于 d 层级的陈述必须可访问。我们将层级 d 上两两不相交的 $\lfloor \frac{D}{d} \rfloor$个陈述设为可访问。层级 d 上所有可访问陈述的并集 $\tilde{s}$ ，是由 $d \lfloor \frac{D}{d} \rfloor = D - \mathrm{mod}(D, d)$ 个原子陈述组成的陈述。根据可访问性传播规则， $\tilde{s}$ 必定是一个可访问陈述。再次应用可访问性传播规则，可知 $\neg\tilde{s}$也必须可访问，但 $\neg\tilde{s}$ 由 $D - (D - \mathrm{mod}(D, d)) = mod(D, d) \lneq d$ 个原子陈述组成，因此根据定义不可能可访问。因此，层级 d上可访问的陈述不能超过 $\lfloor \frac{D}{d} \rfloor - 1$ 个。

现在，我们将层级 d 上 $\lfloor \frac{D}{d} \rfloor - 1$ 不相交的陈述定义为可访问陈述。层级 d 上所有可访问陈述的并集 $\tilde{s}$ 是由 $d(\lfloor \frac{D}{d} \rfloor - 1) = D - \mathrm{mod}(D, d) - d$ 个原子陈述组成的陈述。在这种情况下，$\neg\tilde{s}$ 由 $D - (D - \mathrm{mod}(D, d) - 1) = mod(D, d) + d \gneq d$ 个原子陈述组成，因此上述论证不适用。

□

引理 4。 对于每个模型 (D,d)，理想配置在对原子状态重新标记后均彼此等价。

证明。对于满足 $\mathrm{mod}(\binom{D}{d}) = 0$ 的模型 (D,d)，由引理2和引理3可知，在层级 d 选择可访问陈述集，对应于将 $\binom{D}{d}$ 中的全部原子陈述划分为若干个互不重叠的集合，每个集合的基数为 d。给定这种划分，所有其他可能的划分都可通过重新标记得到。对于满足 $\mathrm{mod}(\binom{D}{d}) \neq 0$ 的模型 (D,d)，由引理2和引理3可知，在层级 d 选择可访问陈述集，对应于将 $\lfloor \binom{D}{d} \rfloor - 1$ 中的原子陈述划分为若干个互不重叠的集合，每个集合的基数为 d，再加上剩余的 $d + \mathrm{mod}(D,d)$ 个原子陈述集。这是对原子陈述集的一个完整划分，因此，这类划分的所有其他可能形式都可通过重新标记得到。

□

引理5。 每个具有 m 原子元素的经典模型只能膨胀为模型 (D,d)，其中对于每个可访问性深度 d，允许的 D 集合为

$$D \in \mathcal{D}_{m,d} := \{md\} \cup \bigcup_{i=1}^{m-1} \{(m+1)d + i\} \tag{1}$$

证明。 这是引理 3的直接应用。 □

引理 6。 存在一个唯一的膨胀过程族，以 $c \in \mathbb{N}_{>0}$ 为参数。这些膨胀过程分别将每个维度为 m 的经典模型映射到不可访问信息模型 (m^{c+1}, m^c)。

证明。 我们希望刻画与模型的复合相容的一般膨胀过程集合。也就是说，我们希望找到一个函数：给定经典模型的维度 m，返回不可访问信息模型（D,d）的理想配置中的原子陈述数量 D 和可访问性深度 d。此外，我们希望该函数与模型的复合相容：模型复合的膨胀必须对应于膨胀复合后得到的模型。

我们知道，对于任意维度为 m 的经典模型，任何满足引理 、d 条件的二元组（D）都会确定一个不可访问信息模型，该模型在其理想配置下是经典模型 m 的一种膨胀。为了明确写出膨胀的一般过程，我们考虑一个函数 F，它以经典维度 m 为参数，并返回膨胀模型的可访问性深度 $d = F(m)$。根据引理 5，膨胀时刻中允许的原子陈述数量是集合 $\mathcal{D}_{m,\,F(m)}$ 的成员。由此可知，膨胀的一般过程将每个经典模型与一个不可访问信息模型相关联。

$$m \to (\tilde{D}, F(m)) \tag{46}$$

˜其中 $D \in \mathcal{D}_{m,\,F(m)}$。

˜现在让我们考虑维度为 m_1 和 m_2的两个通用经典模型。它们的膨胀分别为 $(D_1, F(m_1))$˜ ˜ ˜ 和 $(D_2, F(m_2))$，其中 $D_1 \in \mathcal{D}_{m_1,F(m_1)}$ 以及 $D_2 \in \mathcal{D}_{m_2,F(m_2)}$。经典模型的复合是一个维度为 $m_1 \cdot m_2$ 且具有膨胀的经典模型

$$(\tilde{D}_{12}, F(m_1 \cdot m_2)) \qquad \text{(inflation of composition)} \tag{47}$$

˜其中 $D_{12} \in \mathcal{D}_{m_1 \cdot m_2, F(m_1 \cdot m_2)}$。
每个经典模型的膨胀的复合就是该模型

$$(\tilde{D}_1 \cdot \tilde{D}_2, F(m_1) \cdot F(m_2)) \qquad \text{(composition of inflations)} \tag{48}$$

为了要求膨胀的一般过程与复合相容，我们希望模型（47）与模型（48）相同。我们从令两个复合模型的可达深度相等开始，由此得到，对于函数 F 必须满足

$$F(m_1 \cdot m_2) = F(m_1)F(m_2). \tag{49}$$

由 [25] 我们知道，该函数方程的解为 $F(x) = x^c$，其中参数 c 独立于 m。由于模型的可达深度是一个大于1 的整数，我们希望 F 返回一个大于1的整数值，因此 $c \in \mathbb{N}_{>0}$。于是，模型（47）和（48）可写为

$$(\tilde{D}_{12}, (m_1 \cdot m_2)^c), \qquad \text{(inflation of composition)}, \tag{50}$$
$$(\tilde{D}_1 \cdot \tilde{D}_2, m_1^c \cdot m_2^c), \qquad \text{(composition of inflations)}, \tag{51}$$

$\tilde{D}_1 \in \mathbb{D}_{m_1, m_1^c}$ 和 $\tilde{D}_2 \in \mathbb{D}_{m_2, m_2^c}$。现在我们必须找到元素 $\tilde{D}_{12} \in \mathbb{D}_{m_1 \cdot m_2, (m_1 \cdot m_2)^c}$，其中 $\tilde{D}_{12} \in \mathbb{D}_{m_1 \cdot m_2, (m_1 \cdot m_2)^c}$，$\tilde{D}_1 \in \mathbb{D}_{m_1, m_1^c}$ 和 $\tilde{D}_2 \in \mathbb{D}_{m_2, m_2^c}$，使得

$$\tilde{D}_{12} = \tilde{D}_1 \cdot \tilde{D}_2, \tag{52}$$

独立于 m。由于 c 必须独立于 m，否则膨胀过程就不是一般性的，因此唯一的解是让一般的膨胀过程从 $D = dm$ 中选出 $\mathcal{D}_{m,\, d}$。因此我们有 $\tilde{D}_{12} = m_1 \cdot m_2 \cdot (m_1 \cdot m_2)^c$，$\tilde{D}_1 = m_1 \cdot m_1^c$，以及 $\tilde{D}_2 = m_2 \cdot m_2^c$。于是，模型（47）和（48）可以写成

$$((m_1 \cdot m_2) \cdot (m_1 \cdot m_2)^c, (m_1 \cdot m_2)^c), \qquad \text{(inflation of composition)}, \tag{53}$$
$$((m_1 \cdot m_1^c) \cdot (m_2 \cdot m_2^c), m_1^c \cdot m_2^c), \qquad \text{(composition of inflations)}. \tag{54}$$

我们得到了一族与复合相容、并由正整数 c 参数化的一般膨胀过程；因此，对于任意固定的 c，该膨胀过程会将每个维度为 m 的经典模型关联到一个不可访问信息模型，具体如下

$$m \to (m^{c+1}, m^c). \tag{55}$$

□

引理 7. 可访问性深度为 d 的最大纠缠态模型在层级 d 有 $d+1$ 个概率向量。

证明。 证明的关键在于说明：对于一个由 d^2 个元素构成的集合 $\mathcal{L} = \{s_0, \ldots, s_{d^2-1}\}$，其中 d 为素数，可以选取至多 $d(d+1)$ 个集合 $\{\ell_i\}_{i=0}^{d(d+1)-1}$，每个集合包含 d 个元素，使得任意两个集合的重叠至多为 1 个元素，$|\ell_i \cap \ell_j| \leq 1, \forall i \neq j$。

我们通过构造来证明该陈述。
我们通过将 $\mathcal{L}$ 划分为 d 个互不重叠、每个含有 d 个元素的集合，选出最初的 d 个集合 $\{\ell_i\}_{i=0}^{d-1}$。我们选取集合 $\ell_0 = \{s_0, \ldots, s_{d-1}\}$、$\ell_1 = \{s_d, \ldots, s_{2d-1}\}$、…、$\ell_{d-1} = \{s_{d(d-1)}, \ldots, s_{d^2-1}\}$。其余 d^2 个集合$\{\ell_i\}_{i=d}^{d(d+1)-1}$ 按照以下规则构造，该规则决定元素 s_i 属于哪个集合 ℓ。考虑以下方程

$$x = \mathrm{mod}\left(\left\lfloor \frac{j}{d} \right\rfloor y + \mathrm{mod}(j, d), d\right). \tag{56}$$

对于每个 $j \in \{0, 1, \ldots, d^2-1\}$ 和每个 $y \in \{0, 1, \ldots, d-1\}$，都存在唯一一个满足该方程的 $x \in \{0, 1, \ldots, d-1\}$。换言之，对于 j 的每个取值，都存在 d 个数对 (x, y)，其中 $\in \{0, 1, \ldots, d-1\}$，$\times \{0, 1, \ldots, d-1\}$，且满足方程 (56)。对于每个 j，陈述 s_j 属于 d 集合 $\ell_{d(x+1)+y}$，其中 x 和 y 是方程 (56) 的解。

我们注意到，对于每个 $x \in \{0, 1, \ldots, d-1\}$，集合 $\{\ell_{d(x+1)+y}\}_{y=0}^{d-1}$ 构成 $\mathcal{L}$ 的一个互不重叠的划分，因此对应于一个可及概率分布。

要证明这个构造有效，我们注意任意两个集合 $g = \left\lfloor \frac{m}{d} \right\rfloor$ 当且仅当 $\left\lfloor \frac{m}{d} \right\rfloor = \left\lfloor \frac{n}{d} \right\rfloor$ 且 $g = d(x+1) + y$，其中 (x, y) 是该方程组的唯一解。

$$\begin{cases} x = \mathrm{mod}\left(\left(\left\lfloor \frac{m}{d} \right\rfloor y + \mathrm{mod}(m, d)\right), d\right) \\ x = \mathrm{mod}\left(\left(\left\lfloor \frac{n}{d} \right\rfloor y + \mathrm{mod}(n, d)\right), d\right) \end{cases}, \tag{57}$$

如果 $\left\lfloor \frac{m}{d} \right\rfloor \neq \left\lfloor \frac{n}{d} \right\rfloor$。这证明该集合是极大的，因为添加任何新集合都会至少与其重叠 2 个元素；这也证明，从$\{\ell_i\}_{i=0}^{d(d+1)-1}$ 中选出的任意两个集合的重叠至多有 1 个元素。

□

引理 8. 对于 *MES* 模型，$\vec{q}$ 与层级为 d且对容许态的约束数量最少的可及概率分布之间存在一一对应关系 $\vec{q}$。

证明。我们先从一个计数论证开始。在可访问层级上有 $d+1$ 个可访问概率分布，因此有 $d(d+1)$ 个变量和 $d+1$ 个约束（每个 $d+1$ 个可访问概率分布的元素之和必须为 1），总计有 $(d-1)(d+1)$ 个 $=d^2-1$ 自由变量。在原子层级上有一个准概率向量 $\vec{q}$，因此有 d^2 个变量和一个约束 constraint $\sum_{i=1}^{d^2} q_i = 1$，总计有 d^2-1 个自由变量。现在让我们来看 $\vec{q}$ 与 $\{\vec{p}^{[i]}\}_{i=1}^{d^2}$ 之间的一一映射。按照引理 7 的构造，我们考虑原子陈述 $\mathrm{L} = \{s_0, \ldots, s_{d^2-1}\}$，其中 d 为素数，以及集合 $\{\ell_i\}_{i=0}^{d(d+1)-1}$。对于每个 $i \in \{0, 1, \ldots, d^2-1\}$，我们考虑包含陈述 s_i 的 d 集合 $\{\ell_{dx+y}\}_{(x,y)\in\mathcal{I}(i)}$。集合 $\mathrm{I}(i)$ 明确定义如下

$$\mathcal{I}(i) = \{(x,y) \in \{0,1,\ldots,d-1\} \times \{0,1,\ldots,d-1\} | x = \mathrm{mod}(\left\lfloor \frac{i}{d} \right\rfloor y + \mathrm{mod}(i,d), d)\}. \tag{58}$$

对于每个集合$\{\ell_{dx+y}\}_{(x,y)\in\mathcal{I}(i)}$，都对应一个可访问概率值$\{p_{dx+y}\}_{(x,y)\in\mathcal{I}(i)}$。将这些可访问概率值相加，可得

$$\sum_{(x,y)\in\mathcal{I}(i)} p_{dx+y} = dq_i + 1, \tag{59}$$

其中，$q_i = Q(s_i)$ 是与陈述 s_i 相关的不可访问准概率。因此，可以从层级 d 的可访问概率分布重构向量 $\vec{q}$。反过来显然也成立：由向量 $\vec{q}$ 可以重构层级 d 的可访问概率分布。通过注意到，只移除一个可访问概率分布（即选取一个 x 并舍弃所有概率值 $\{p_{dx+y}\}_{y=0}^{d-1}$），就无法再从可访问概率分布中恢复 $\vec{q}$，可以验证所施加的约束数量已是最少。于是，对于一般的最大纠缠态$_d$ 模型，我们有从单个准概率分布 $\vec{q}$ 到一组 $d+1$ 可访问概率分布 $\{\vec{p}^{[i]}\}_{i=1}^{d+1}$ 的可逆映射。

□

B 陈述与准概率

B.1 陈述代数

我们在日常生活中使用的语言，很大程度上依赖于联结词 $\vee$（读作“或”）和 $\wedge$（读作“且”），以及一元否定算子 $\neg$。这些联结词在它们于陈述集上生成的代数中，编码了我们理解和感知信息的机制结构的一部分。研究如何使用运算 $\vee$、$\wedge$ 或 $\neg$ 组合和操作陈述，或许能让我们对这一机制有更深入的了解。

我们提出一个示例，其中只需考虑两个陈述 s_1 和 s_2。为便于阐述，我们考虑这两个特定的陈述。

$$\begin{aligned} s_1 &= \text{"The apple is blue"}, \\ s_2 &= \text{"The dog is red"}. \end{aligned}$$

一个有趣的问题是，仅使用 $\vee$ 、 $\wedge$ 和 $\neg$ ，并可自由使用 s_1 和 s_2 ，能够得到多少个新的陈述。这个问题很容易回答。我们可以得到以下陈述：

$$
\begin{aligned}
&s_3 := \neg s_1 = \text{"苹果不是蓝色的"}, \quad s_4 := \neg s_2 = \text{"狗不是红色的"},\\
&s_5 := s_1 \vee s_2 = \text{"苹果是蓝色的，或者狗是红色的"}, \quad s_6 :\\
&= s_1 \vee \neg s_2 = \text{"苹果是蓝色的，或者狗不是红色的"}, \quad s_7 :\\
&= \neg s_1 \vee s_2 = \text{"苹果不是蓝色的，或者狗是红色的"}, \quad s_8 :\\
&= \neg s_1 \vee \neg s_2 = \text{"苹果不是蓝色的，或者狗不是红色的"}, \quad s_9 :\\
&= s_1 \wedge s_2 = \text{"苹果是蓝色的，并且狗是红色的"}, \quad s_{10} := s_1 \wedge \neg s_2 = \text{"}\\
&\text{苹果是蓝色的，并且狗不是红色的"}, \quad s_{11} := \neg s_1 \wedge s_2 = \text{"苹果不是}\\
&\text{蓝色的，并且狗是红色的"}, \quad s_{12} := \neg s_1 \wedge \neg s_2 = \text{"苹果不是蓝色的，}\\
&\text{并且狗不是红色的"},
\end{aligned}
$$

$s_{13} := (s_1 \wedge s_2) \vee (\neg s_1 \wedge \neg s_2) =$ "（苹果是蓝色的，而且狗是红色的）或（苹果不是蓝色的，或者狗不是红色的）"， $s_{14} :$
$= (s_1 \wedge \neg s_2) \vee (\neg s_1 \wedge s_2) =$ "（苹果是蓝色的，而且狗不是红色的）或（苹果不是蓝色的，或者狗是红色的）"，

$$
\begin{aligned}
&s_{15} := s_1 \vee \neg s_1 = \text{"苹果是蓝色的，或者苹果不是蓝色的"},\\
&s_{16} := s_1 \wedge \neg s_1 = \text{"苹果是蓝色的，而且苹果不是蓝色的"}。
\end{aligned}
$$

我们能写出陈述的任何新组合吗？当然可以，例如我们可以想到

$$s_{17} := s_5 \vee s_1 = \text{"The apple is blue or the dog is red or the dog is red"},$$

但很快我们就会意识到 这个陈述实际上等同于 s_5 的陈述。我们可以想到这个陈述 nt

$$s_{18} := s_2 \vee \neg s_2 = \text{"The dog is red or the dog is not red"}$$

作为一个新的陈述，但很快我们就能说服自己，这个陈述所传达的信息与陈述 s_{15} 所传达的信息相同，或者与那个复杂得多的陈述所传达的信息相同

$$\tilde{s}_{18} = ((s_1 \wedge s_2) \vee (\neg s_1 \wedge \neg s_2)) \vee ((s_1 \wedge \neg s_2) \vee (\neg s_1 \wedge s_2)).$$

我们可以尝试对任何其他陈述进行这个游戏，从而认识到，这 16 个陈述 s_1、.....、 s_{16} 就是我们能够构成的全部陈述。

为了不只是凭借直觉来进行等价陈述的识别过程，我们现在将刚才用于组合陈述的规则形式化。

我们首先定义一个部分人为构造的演练场，用于简化处理。在我们的全部工作中，我们都将考虑可能陈述总数有限的情形。我们认为这是一个相当人为的限制，因为凭直觉思考时，人们会将陈述集视为无限的。我们将要考虑的每个陈述集都关于运算 $\wedge$、 $\vee$ 和 $\neg$闭合。我们考虑构成布尔格结构的陈述。布尔格是一个分配的有补格。陈述格是一种代数结构（$\mathcal{S}$ $\wedge$， $\vee$），由一个陈述集 $\mathcal{S}$ 以及两个二元运算 $\wedge$、 $\vee$ 组成，并满足以下规则

$$
\begin{aligned}
&x \vee x = x, \quad x \wedge x = x, && \textit{(Idempotency)} && (60a)\\
&x \vee y = y \vee x, \quad x \wedge y = y \wedge x, && \textit{(Commutativity)} && (60b)\\
&x \vee (y \vee z) = (x \vee y) \vee z, && \textit{("or" associativity)} && (60c)\\
&x \wedge (y \wedge z) = (x \wedge y) \wedge z, && \textit{("and" associativity)} && (60d)\\
&x \vee (x \wedge y) = x \wedge (x \vee y) = x. && \textit{(Absorption)} && (60e)
\end{aligned}
$$

符号” $=$ ”用于表示陈述之间的等价关系。如果一个格是完备格（$\mathcal{S}$ $\wedge, \vee, \top, \perp$），那么它具有一个“最大元” $\top$ 和一个“最小元” $\perp$，使得对于所有陈述 $s \in \mathcal{S}$

$$
\begin{aligned}
&s \vee \top = \top, && (61a)\\
&s \wedge \perp = \perp . && (61b)
\end{aligned}
$$

最大元和最小元是某种特殊元素，它们不属于陈述集，却存在于每个完备格中。在一个完备格（$\mathcal{S}$ $\wedge, \vee, \top, \bot$）中，可能存在陈述 $s_i \in \mathcal{S}$，使得不存在任何 $s_j, s_k \in \mathcal{S}$，满足 $s_j \neq s_k$，且 $s_i = s_j \vee s_k$。这些称为原子陈述。我们将 $\mathcal{S}$ 中的原子陈述集记为$\mathcal{A}(\mathcal{S})$。最大元可以看作与通过二元运算符 $\vee$, $\top = \bigvee_{s \in \mathcal{S}} s$组合$\mathcal{S}$ 中的所有陈述所得的陈述等价，而最小元可以看作与通过二元运算符$\wedge$, $\bot = s_i \wedge s_j$ 组合任意两个不同原子陈述所得的陈述等价，其中 $s_i \neq s_j$ 是原子陈述。 $\top$ 和 $\bot$ 的“最大元”与“最小元”这一命名直接源自下一节介绍的格的图示。陈述格（$\mathcal{S}$ $\wedge, \vee, \top, \bot$）满足以下规则时称为结合律

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \qquad \textit{(Distributivity of } \wedge \textit{ over } \vee) \tag{62a}$$
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z). \qquad \textit{(Distributivity of } \vee \textit{ over } \wedge) \tag{62b}$$

如果对每个陈述 $s \in \mathcal{S}$，都存在唯一的陈述 $\neg s \in \mathcal{S}$ （ s 的否定），使得

$$x \vee \neg x = \top, \qquad \textit{("or" complementation)} \tag{63a}$$
$$x \wedge \neg x = \bot . \qquad \textit{("and" complementation)} \tag{63b}$$

这些规则定义了 陈述之间的逻辑等价关系，而无需依赖真值的概念 ue.

通常，在偏序集的语境下，格被定义为每个元素都有唯一上确界和下确界的偏序集。我们可以从陈述代数的格结构转到定义如下陈述间偏序关系的偏序集的格结构：对于任意两个陈述 $s_i, s_j \in \mathcal{S}$

$$s_i \leq s_j \iff s_i \wedge s_j = s_i \text{ and } s_i \vee s_j = s_j. \tag{64}$$

此外，我们可以 [8], 也可以方便地将陈述视为集合。在这种表述下，二元运算符 $\wedge$ 和 $\vee$ 分别对应集合的交集 $\cap$ 和并集 $\cup$ 。分配格是按集合包含关系排序的集合格 $\subseteq$。

可以用哈斯图直观地表示代数结构。由于哈斯图简单且直观，我们大量使用其提供的图示语言，并在附录 B.2中简要介绍如何使用哈斯图。

通常，为了方便记号，我们会用$\mathcal{S}$ 代替$\mathcal{A}(\mathcal{S})$。例如，写作$\mathcal{S}_3 = \{s_1, s_2, s_3\}$时，我们隐含地表示 $s_i \wedge s_j = \bot \quad \forall i, j \in [1, 3], i \neq j$，并且考虑所有元素$\{\bot, s_1, s_2, s_3, s_1 \wedge s_2, s_1 \wedge s_3, s_2 \wedge s_3, \top\}$。

我们在本节最后指出，初始示例中引入的陈述$\{s_1, \ldots, s_{16}\}$ 确实构成一个布尔格。

B.2 格的图示记号

本文采用哈斯图的图示记号。参见图 5，陈述表示为填充圆圈，圈内标有陈述名称（图版 (*a*)）。对于不包含陈述 $\top$ 或 $\bot$的任意一组圆圈，有如下图示规则。若陈述 s 的圆圈通过一条从下方连至陈述集 $\{s_1, s_2, \ldots, s_N\}$的线段相连，则表示 $s = s_1 \vee s_2 \vee \cdots \vee s_N$ （图版 (*b*)）。若陈述 s的圆圈通过一条从上方连至陈述集 $\{s_1, s_2, \ldots, s_n\}$的线段相连，则表示 $s = s_1 \vee s_2 \vee \ldots s_n$ （图版 (*c*)）。

此外，对于任意有限命题集 $\mathcal{S}$，其原子元素为 $\{s_1$、 s_2、……、 $s_D\}$，都可以在图示中引入 $\bot$ 和 $\top$ 元素，如图 6 所示，其中我们使用记号 $\neg s_i$ 表示除 s_i之外所有原子陈述的并集（例如 $\neg s_D = s_1 \vee s_2 \vee \cdots \vee s_{D-1}$）。

使用这种图示记号，我们可以如图 $\mathcal{S}_3 = \{a$、 b、 $c\}$ 所示，表示有限命题集 7中的所有陈述。

对于从 D 原子陈述 $\mathcal{S}_D = \{s_1$， s_2，……， $s_D\}$ 开始得到的一般陈述集，其图示如下。原子陈述列在包含 $\bot$元素那一行的上一行。格中从底部开始的每一行都标有一个从零开始的整数，我们称这个数为格的层级。例如，在层级 0，我们只有 $\bot$ 元素。在层级 1，我们有

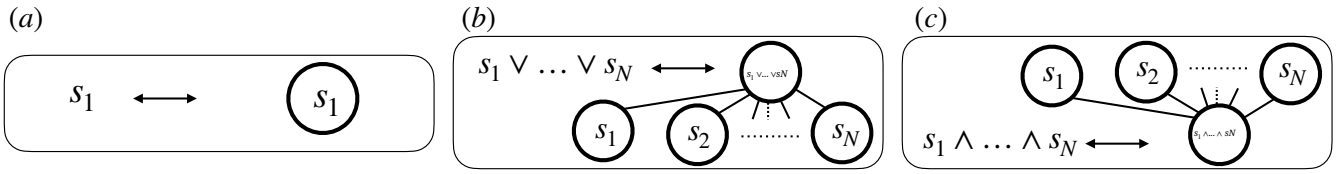

图 5：陈述用实心圆表示，圆内标有该陈述的名称（图（a））。对于不包含陈述 $\top$ 或 $\bot$ 的任意圆集合，适用以下图示规则。若陈述 s 的圆通过一条从下方连接到陈述集 $\{s_1,\ s_2,\ \ldots\ldots,\ s_N\}$ 的线相连，则表示 $s = s_1 \vee s_2 \vee \ldots\ldots \vee s_N$（图（$b$））。若陈述 s 的圆通过一条从上方连接到陈述集 $\{s_1,\ s_2,\ \ldots\ldots,\ s_n\}$ 的线相连，则表示 $s = s_1 \vee s_2 \vee \ldots s_n$（图（$c$））。

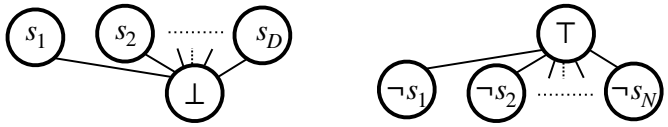

图 6：最大元和最小元的图示。

所有原子陈述；在层级 $1 < l < D$，我们有由 l 个不同原子陈述取并集得到的陈述。在层级 D，我们有 $\top$ 元素。

我们以陈述的布尔格的图示结束本节$\{s_1, \ldots, s_{16}\}$，该格在第B.1节中引入。

B.3 陈述格中的准概率

B.3.1 陈述格中的概率

从陈述格推导概率论的关键在于蕴涵这一概念。通常，人们会在日常语言中将蕴涵理解为与陈述的真值相关的概念。例如，人们会说，陈述 a 蕴涵陈述 b，如果 a 为真足以使 b 为真。我们可以将蕴涵这一概念从真值中抽象出来，只将它与陈述所包含的信息这一概念联系起来；更确切地说，将它与两个陈述之间存在的关系联系起来。我们称陈述 s_i 蕴涵另一个陈述 s_j，记作 $s_i \to s_j$，当且仅当 $s_i \vee s_j = s_j$ 和 $s_i \wedge s_j = s_i$。我们可以探究这个定义，看看它如何说得通。从诗意的角度来说，只要在所考虑的陈述集$\mathcal{S}$ 的有限词汇表中，陈述 s_i 是陈述 s_j 的提喻，就可以说陈述 s_i 蕴涵陈述 s_j。

在其著作中，[11, 12] R.T.Cox 意识到，将蕴含的概念推广为陈述之间的蕴含程度这一概念，便可得到概率论。与这些蕴含程度相关联的实数，若视为表示我们相信一个陈述蕴含另一个陈述的程度，就等同于概率。

R. T. Cox 最初提出了从逻辑推导概率论的方法和概念。[12] 此后，许多其他作者重新审视、阐明并进一步发展了这些方法和概念。E. T. Jaynes 对这一主题作了精彩阐述，见短篇著作 [14]，或见著作 [13] 的前几章。在那里，概率被推导为一种唯一的推理方法：具备常识的机器人（这是一个多少有些时代错置的概念）会采用这种方法。其他阐述见 A. Caticha 的许多著作（此处引用其未发表的 [26]，其中汇集了他的许多已发表作品，以及 [27]），也可参见 K. H. Knuth [9, 10] 和 H. Reichenbach [28] 的著作，谨举数例。下一节中，我们将放宽这些作者提出的一些要求，最终推导出准概率，而非概率。我们即将介绍的准概率推导，绝不像上述著作中的概率推导那样完整详尽；不过，我们可以对这一主题作全面论述，并着重指出它与概率推导的差异。我们将把这项分析及相关讨论留待日后展开；本文仅介绍当前工作所需的最基本工具和一致性规则。

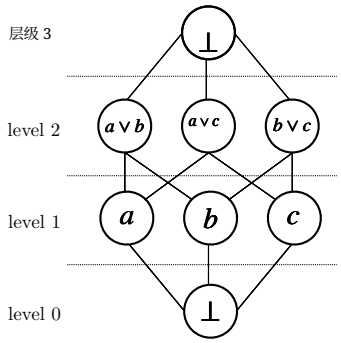

图7：有限命题集 $\mathcal{S}_3 = \{a、b、c\}$ 的图示。原子陈述收集在包含 $\perp$ 元素的那一行上方的第一行中。我们从底部开始，用从零开始的整数为格的每一行编号，并将这个编号称为格的层级。在层级 0，我们只有 $\perp$ 元素。在层级 1，我们有所有原子陈述；在层级 $1 < l < 3$，我们有由 l 个不同原子陈述的并集得到的陈述。在层级 3，我们有 $\top$ 元素。请注意，$s_1 \vee s_2$ 有两条从下方进入、分别连接到 s_1 和 s_2 的线，这意味着它是 s_1 和 s_2的并集。还请注意，s_2 有两条从上方进入、分别连接到 $s_1 \vee s_2$ 和 $s_2 \vee s_3$的线，这意味着 s_2 是 $s_1 \vee s_2$ 和 $s_2 \vee s_3$的交集。一般而言，向上对应于取并集，向下对应于取交集。

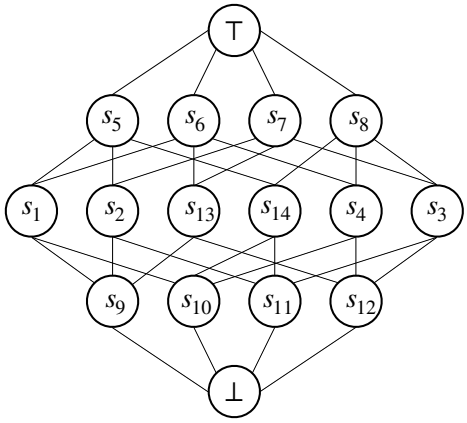

图8：第B.1节示例中陈述的图示。

B.3.2 陈述代数中的准概率

仿照 Jaynes 对可信度的定义 [13]以及 Knuth 对包含度的定义 [10], ，我们将“关系量”函数$\mathcal{R}$ 定义为一个以一对陈述（包括$\top$）为参数并返回实数的函数。函数$\mathcal{R}$ 旨在以某种方式量化两个陈述之间的关系。此外，我们要求 $\mathcal{R}(a, b)$ 在 b 蕴含 a 时返回 1，并在 $a \wedge b = \perp$（即 b 蕴含 $\neg a$）时返回 0。如此定义的函数相当一般，但通过进一步提出要求，我们便能用一些特定性质刻画$\mathcal{R}$，使其成为“条件拟概率函数”。

这些进一步的要求分为两组。首先需要牢记的是，$\mathcal{R}$ 的自变量构成一种具有特定性质的形式和代数。我们希望$\mathcal{R}$ 与这些性质相一致。关于$\mathcal{R}$ 的第二类假设，就是我们将要称为“常识性假设”的内容，这一术语借鉴了 Jaynes 使用的类似说法。常识性假设是一些定性性质，我们期望$\mathcal{R}$ 满足这些性质，以便体现我们对“可信度”的理解；这些性质也用于概率论的推导[11]。在讨论$\mathcal{R}$ 时，我们力求尽可能少地采用常识性假设，从而刻画出$\mathcal{R}$ 的一种更一般的形式，而这种形式几乎仅受与格结构相一致这一要求的约束。

[11]考克斯定理以及如何从命题逻辑推导出概率，长期以来一直备受争议。不同作者使用各种附加假设来得出相同的结果（概率论规则）。常识性假设是指任何足以推导出这些规则的附加假设集合。

已于 2024-08-30 接收，点击标题进行验证。根据 CC-BY 4.0 许可发布。

原子陈述的求和规则 例如，我们可能会问，当$\mathcal{R}$ 的一个自变量是陈述的复合时，它会如何表现。考虑原子陈述 a, b，使得 $a \wedge b = \bot$，以及第三个陈述 t，使得 $a \vee t = t$ 和 $b \vee t = t$。我们将 $\mathcal{R}(a \vee b, t)$ 视为以下自变量的函数：

$$\mathcal{R}(a \vee b, t) = \tilde{F}\left[\mathcal{R}(a,t), \mathcal{R}(b,t), \mathcal{R}(a, b \wedge t), \mathcal{R}(b, a \wedge t), \mathcal{R}(a \wedge b, t)\right], \tag{65}$$

对于一个尚未指定的函数 $\tilde{F}$，其定义域为 $\mathbb{R}^5$，陪域属于 $\mathbb{R}$。选择这些特定的参数是我们所作的一项附加假设。这是一项在文献中经常出现的假设；对某些作者而言，这项假设似乎如此自然，以至于他们甚至不把它视为一项假设。[12]。在本文中，我们并不认为这项假设显而易见到无需论证。事实上，例如，我们没有考虑 $\mathcal{R}(a \vee b, t)$ 对 $\mathcal{R}(\neg a, t)$的依赖关系。然而，这是一项范围广泛的假设，赋予我们相当大的自由度，并且符合文献中的标准方法。[13]

在我们的情形中，我们有 $a \wedge b = \bot$，因此

$$\mathcal{R}(a \vee b, t) = \tilde{F}\left[\mathcal{R}(a,t), \mathcal{R}(b,t), 0, 0, 0\right], \tag{66}$$

从而实际上将上述函数关系简化为

$$\mathcal{R}(a \vee b, t) = F\left[\mathcal{R}(a,t), \mathcal{R}(b,t)\right], \tag{67}$$

其中 F 是一个尚未指定的函数，其定义域属于 $\mathbb{R}^2$，陪域属于 $\mathbb{R}$。

现在我们考虑第四个陈述 c，使得 $a \wedge b = a \wedge c = b \wedge c = \bot$。由于陈述满足结合律（性质 (60c)），我们有：

$$a \vee (b \vee c) = (a \vee b) \vee c, \tag{68}$$

因此，我们应当预期如下：

$$\mathcal{R}(a \vee (b \vee c), t) = \mathcal{R}((a \vee b) \vee c, t). \tag{69}$$

使用 (67) 我们得到等式

$$F\left[\mathcal{R}(a,t), F\left[\mathcal{R}(b,t), \mathcal{R}(c,t)\right]\right] = F\left[F\left[\mathcal{R}(a,t), \mathcal{R}(b,t)\right], \mathcal{R}(c,t)\right]. \tag{70}$$

重命名 $x = \mathcal{R}(a,t)$、 $y = \mathcal{R}(b,t)$以及 $z = \mathcal{R}(c,t)$ 后，我们可以将方程 (70) 改写为

$$F[x, F[y,z]] = F[F[x,y], z]. \tag{71}$$

这个函数方程称为结合律方程。由 Abel 于 [31] 提出，随后由 Aczel 进行了深入研究 [25, 32], 结合律方程是概率论推导的核心。事实上，如果对 F 作出一些常识性假设 [25, 32, 33](参见第 B.3.3节)，则已知函数方程 (71) 有唯一解（除了自变量乘以一个标量的情况），其形式为

$$F(x,y) = f^{(-1)}\left(f(x) + f(y)\right), \tag{72}$$

其中， f 是一个定义域为 $\mathbb{R}$ 、陪域为 $\mathbb{R}_+$ 的可逆函数，而 $f^{(-1)}(x)$是 $f(x)$的逆函数。这个唯一解是概率论特征刻画的核心。在我们的推导中，我们并未要求对一般的 F 作任何常识性假设。尽管如此，只需将（72）代入（71）即可验证，对于任意可逆的 $f : \mathbb{R} \to \mathbb{R}$，方程（72）都是（71）的一个解。在这种情况下， F的函数形式并不存在唯一性定理。当函数$\mathrm{R}(x,y)$被限制在可达子格中的元素所构成的定义域内时，施加常识性假设便有了意义；这意味着 x 和 y 是可访问陈述。因此，在可达子格中，函数方程（71）的解形式确实是唯一的，即（72）所示的形式，并且我们预计能够恢复概率论原理。

[12]例如，Knuth 在 [9] 中仅将 $\mathcal{R}(a,t), \mathcal{R}(b,t)$ 视为可能的参数，并在其公式 (2) 之前这样说明这一选择：“由于 $a \wedge b = \bot$，这种包含度只能是 $\phi(a,t)$ 和 $\phi(b,t)$ 的函数”，其中他称作包含度的 ϕ就是我们的 $\mathcal{R}$。在 [26] 第 2 章中，作者用以下句子强调，相较于我们的选择，采用较小的参数集合仍具有足够的普适性：“*...]* 的参数涵盖了 a 和 b 在 c 语境中的全部四种可能的信念度，而不是任何可能存在问题的子集。”，其中 Caticha 的 c 对应于我们的 t。

[13]严格遵循 Cox 原著原则的概率论推导，通常从乘法规则的推导开始，而不是求和规则。在这一框架中，人们对与我们的函数 F 类似的函数的自变量作出了相应的假设。然而，在此语境下，这一假设被认为远非显而易见，并引发了广泛争论 [29, 30, 26, 13]。

在格的其余部分中， F 的解并不唯一。我们将假设 F 具有与 (72)中相同的函数形式；这一选择的依据留待今后的工作说明。此假设限制并不太强，并使一般解能够包含一个 f ，它是一个从 $\mathbb{R}$ 到 $\mathbb{R}$的任意可逆函数。

总之，我们发现 F 有如下形式的解(72)，其中 f 是一个陪域限制为 $\mathbb{R}_+$ 的可逆函数，在 $\mathcal{R}(a,t)$ 上取值，其中 a 和 t 是可访问陈述。此外，我们固定 f 使得$f(1) = 1$ 且 $f(0) = 0$。

我们将重新参数化 $\mathcal{R}(x,t) = f(\mathcal{R}(x,t))$ 视为新的"关系量"函数，对于一般命题 a 和 b，使得 $a \wedge b = \bot$ 该方程 (67) 可以改写为以下形式：

$$\tilde{\mathcal{R}}(a \vee b, t) = \tilde{\mathcal{R}}(a,t) + \tilde{\mathcal{R}}(b,t) \tag{73}$$

与陪域为 $\mathbb{R}$的$\mathcal{R}$ 。鉴于函数 f 可逆，且其定义满足 $f(1) = 1$ 和$f(0) = 0$，因此$\mathcal{R}$ 和$\mathcal{R}$ 可以互换使用，因为二者均可解释为"关系量"，其中一个只是另一个的重新参数化。我们选择使用$\mathcal{R}$ 及与这一选择相关的规则。

选择 取代$\mathcal{R}$ ，就目前而言，是任意的。理想情况下，这一选择应由一个额外的、更具"直观性"的假设加以论证，这项工作留待未来完成。这种方法将能够基于一组合理的假设，充分地推导出令人满意的准概率规则。

由方程 (73) 可见，"关系量"遵循类似于概率论中求和规则的规则。

求和规则 现在，我们将求和规则扩展到非原子命题。我们遵循 [9] 来进行此操作。考虑两个陈述 x 和 y ，并将它们写成原子陈述的复合

$$\begin{aligned} x &= (p_1 \vee p_2 \vee \cdots \vee p_n) \vee (r_1 \vee r_2 \vee \cdots \vee r_k), \\ y &= (r_1 \vee r_2 \vee \cdots \vee r_k) \vee (w_1 \vee w_2 \vee \cdots \vee w_m). \end{aligned} \tag{74}$$

显然

$$\begin{aligned} x \wedge y &= (r_1 \vee r_2 \vee \cdots \vee r_k), \\ x \vee y &= (p_1 \vee p_2 \vee \cdots \vee p_n) \vee (r_1 \vee r_2 \vee \cdots \vee r_k) \vee (w_1 \vee w_2 \vee \cdots \vee w_m). \end{aligned} \tag{75}$$

反复对不相交项使用求和规则 (73) ，我们可以写成

$$\tilde{\mathcal{R}}(x \vee y, t) = \sum_{i=1}^{n} \tilde{\mathcal{R}}(p_i, t) + \sum_{i=1}^{k} \tilde{\mathcal{R}}(r_i, t) + \sum_{i=1}^{m} \tilde{\mathcal{R}}(w_i, t). \tag{76}$$

添加和移除 $\sum_{i=1}^{k} \tilde{\mathcal{R}}(r_i, t)$，并适当合并各项，可得

$$\tilde{\mathcal{R}}(x \vee y, t) = \tilde{\mathcal{R}}(x,t) + \tilde{\mathcal{R}}(y,t) - \tilde{\mathcal{R}}(x \wedge y, t). \tag{77}$$

我们得出，$\mathcal{R}$ 遵循概率论的一般求和规则。

乘法规则 在推导 $\mathcal{R}$ 的乘法规则时，我们再次遵循 [9]。我们希望找到一个函数 G，使其与命题代数的结构相一致，从而能够写出 $\mathcal{R}(x \wedge y, t)$，而无需引用$\mathcal{R}(x \vee y, t)$。根据 [29, 30] 的分析（简要总结见 [26] 的第 2.3.1 节），我们知道该函数的参数必须是

$$\tilde{\mathcal{R}}(x \wedge y, t) = G\left[\tilde{\mathcal{R}}(x,t), \tilde{\mathcal{R}}(y, x \wedge t)\right]. \tag{78}$$

与求和规则的情形类似，我们考虑两个陈述 a和 b，使得 $a \wedge b = \bot$ 以及两个陈述 r和 s，使得 $r \wedge s = \bot$。利用求和规则和分配律 (62a) 可得

$$\tilde{\mathcal{R}}\left(a \wedge (r \vee s), t\right) = G\left(\tilde{\mathcal{R}}(a,t), \tilde{\mathcal{R}}(r \vee s, a \wedge t)\right) = G\left(\tilde{\mathcal{R}}(a,t), \tilde{\mathcal{R}}(r, a \wedge t) + \tilde{\mathcal{R}}(s, a \wedge t)\right), \tag{79}$$

$$\tilde{\mathcal{R}}\left(a \wedge (r \vee s), t\right) = G\left(\tilde{\mathcal{R}}(a,t), \tilde{\mathcal{R}}(a, r \wedge t)\right) + G\left(\tilde{\mathcal{R}}(a,t), \tilde{\mathcal{R}}(a, s \wedge t)\right). \tag{80}$$

令 (79) 等于 (80)，并代入 $u = \mathcal{R}(a,t)$、 $v = \mathcal{R}(r, a \wedge t)$以及 $w = \mathcal{R}(s, a \wedge t)$，可得函数方程

$$G(u, v + w) = G(u, v) + G(u, w). \tag{81}$$

现在我们定义 $z = v + w$，并且有

$$\frac{\partial}{\partial v} = \frac{\partial z}{\partial v}\frac{\partial}{\partial z} = \frac{\partial}{\partial z}, \tag{82}$$

$$\frac{\partial}{\partial w} = \frac{\partial z}{\partial w}\frac{\partial}{\partial z} = \frac{\partial}{\partial z}, \tag{83}$$

因此 $\frac{\partial^2}{\partial z^2} = \frac{\partial}{\partial v}\frac{\partial}{\partial w}$。将这个二阶导数应用于方程 (81) 我们得到

$$\frac{\partial^2}{\partial z^2} G(u, z) = \frac{\partial}{\partial v}\frac{\partial}{\partial w}(G(u, v) + G(u, w)) = 0, \tag{84}$$

因此， G 关于其第二个自变量是线性的，所以其最一般的形式为：

$$G(u, v) = A(u)v + B(u). \tag{85}$$

将方程 (85) 代入方程 (81) 我们看到 $B(u) = 0$ 对每个 u都成立。以类似的推理过程继续从 $\mathcal{R}((a \vee b) \wedge r)$ 出发，我们发现 G 对其第一个自变量也是线性的。最终得到

$$G(u, v) = Cuv, \tag{86}$$

其中 C 是一个常数，我们将其设为 1。我们得到 $\mathcal{R}$ 遵循概率论的乘法规则。根据乘法规则，再结合交换性 (60b)，可以得到贝叶斯定理。

准概率 现在我们终于可以定义函数 Q，它将一般陈述集的元素映射到实数。对于一般陈述 $s \in \mathcal{S}$， Q 的作用定义为

$$Q(s) = \tilde{\mathcal{R}}(s, \top). \tag{87}$$

此外，定义 $Q(s_i \mid s_j) = \mathcal{R}(s_i, s_j \wedge \top) = \mathcal{R}(s_i, s_j)$，可以轻松推导出 Q 的以下性质。：

- $Q(\top) = 1$,
- $Q(\bot) = 0$,
- $Q(s_i) \in \mathbb{R}$,
- $\sum_{s \in \mathcal{A}(\mathcal{S})} Q(s) = 1$,
- $Q(s_i \vee s_j) = Q(s_j \vee s_i) = Q(s_i) + Q(s_j) - Q(s_i \wedge s_j)$,
- $Q(s_i \wedge s_j) = Q(s_j \wedge s_i)$,
- $Q(s_i \wedge s_j) = Q(s_i)Q(s_i \mid s_j)$.

函数 Q 是定义在分配格 [34]，上的一个 赋值，其在每个陈述上的取值都由值集 $\{Q(s)\}^{s \in \mathcal{A}(\mathcal{S})}$ 唯一确定。

函数 Q 将一个准概率分布值关联到$\mathcal{S}$中的每个陈述。我们看到，根据命题代数，可以自然地为每个陈述赋予 Q 实数，且这种赋值遵循求和规则、乘法规则和贝叶斯定理，因此其表现如同概率赋值。这里要指出的是，我们从未假设概率论，而是从与命题代数保持一致这一要求推导出了这些规则。此外，对于每个陈述格， Q的赋值都完全由 Q 在原子陈述集合上的赋值确定。因此， Q 的赋值完全由向量 $\vec{q} = \{Q(s_1)\ Q(s_2) \ldots, Q(s_D)\} \in \mathbb{R}^D$的选取确定，其中$\mathcal{A}(\mathcal{S}) = (s_1, s_2, \ldots, s_D)$是原子陈述。我们可以完全自由地选择向量 $\vec{q}$的各个元素，唯一的约束是它们的和为 1, $\sum_{i=1}^{D} q_i = 1$。如果如 Jaynes 所主张的那样，概率是对亚里士多德逻辑的推广，那么准概率语言就是对亚里士多德逻辑的进一步 推广。

B.3.3 陈述代数中的概率

在附录 B.3.1 中，我们介绍了将概率论视为对经典逻辑的松弛而推导出来这一思想。在经典逻辑中，每个陈述都对应一个二值真值。一个陈述非真即假。在概率论中，可以为每个陈述赋予一个取自连续区间的实数 $[0,1]$。

我们将为格中的每个元素赋予一个实数，视为对二值真值赋值的放宽。形式上，对于任意格 L，我们考虑集合 S_V^L，其中包含从格的元素（即陈述）到二值真值的所有函数。我们将 $V \in S_V^L$ 称为逻辑估值。每个逻辑估值都对应于对格中所有元素的不同真值赋值。我们用 $\tilde{S}_V^L$ 表示所有与可接受的真值赋值相对应的逻辑赋值的集合（见第2节），即

$$V(\top) = \mathbf{T}, \qquad V(\bot) = \mathbf{F}, \tag{88}$$

并且对于任意 $a,\ b \in L$

$$\begin{aligned}
&(V(a) = \mathbf{T} \text{ and } V(b) = \mathbf{F}) \text{ or } (V(a) = \mathbf{F} \text{ and } V(b) = \mathbf{T}) && \Longrightarrow V(a \vee b) = \mathbf{T} \text{ and } V(a \wedge b) = \mathbf{F}, \\
&(V(a) = \mathbf{F} \text{ and } V(b) = \mathbf{F}) && \Longrightarrow V(a \vee b) = \mathbf{F} \text{ and } V(a \wedge b) = \mathbf{F}, \\
&(V(a) = \mathbf{T} \text{ and } V(b) = \mathbf{T}) && \Longrightarrow V(a \vee b) = \mathbf{T} \text{ and } V(a \wedge b) = \mathbf{T}, \\
&V(a) = \mathbf{F} && \Longrightarrow V(\neg a) = \mathbf{T}, \\
&V(a) = \mathbf{T} && \Longrightarrow V(\neg a) = \mathbf{F}.
\end{aligned} \tag{89}$$

此外，对于任意格 L，我们考虑集合 $S_{\mathcal{V}}^L$，其中包含从格的元素（即陈述）到实数值的所有函数。我们将 $\mathcal{V} \in S_{\mathcal{V}}^L$ 称为实值估值。我们进一步定义 $\tilde{S}_{\mathcal{V}}^L$，即我们视为逻辑赋值 $\tilde{S}_V^L$ 放宽形式的实值估值子集，它由满足与真值表规则相容且与格的结构相容的实值估值构成。实值估值$\mathcal{V}$ 与真值表相容，当且仅当

$$\mathcal{V}(\top) = \beta, \qquad \mathcal{V}(\bot) = \alpha, \tag{90}$$

对于两个固定常数 $\alpha,\ \beta \in \mathbb{R}$ （我们假设 $\alpha < \beta$，但以下所有内容对 $\alpha > \beta$ 均成立；若要保持与逻辑估值的类比，则舍弃 $\alpha = \beta$ 的情况），以及格L中的任意两个陈述 $a,\ b$：

$$\begin{aligned}
&(\mathcal{V}(a) = \alpha \text{ and } \mathcal{V}(b) = \beta) \text{ or } (\mathcal{V}(a) = \beta \text{ and } \mathcal{V}(b) = \alpha) && \Longrightarrow \mathcal{V}(a \vee b) = \beta \text{ and } \mathcal{V}(a \wedge b) = \alpha, \\
&(\mathcal{V}(a) = \alpha \text{ and } \mathcal{V}(b) = \alpha) && \Longrightarrow \mathcal{V}(a \vee b) = \alpha \text{ and } \mathcal{V}(a \wedge b) = \alpha, \\
&(\mathcal{V}(a) = \beta \text{ and } \mathcal{V}(b) = \beta) && \Longrightarrow \mathcal{V}(a \vee b) = \beta \text{ and } \mathcal{V}(a \wedge b) = \beta, \\
&\mathcal{V}(a) = \alpha && \Longrightarrow \mathcal{V}(\neg a) = \beta, \\
&\mathcal{V}(a) = \beta && \Longrightarrow \mathcal{V}(\neg a) = \alpha.
\end{aligned} \tag{91}$$

逻辑赋值松弛这一概念，源自 α、β 与 **F**、**T**之间直观的对应关系。方程 (91) 的要求对应于实值估值在可化约为逻辑估值时与真值表保持一致的要求。

只要要求方程中列出的性质，逻辑赋值就会自动与格的结构相容(88) 和 (89)，但实值估值则不会，因为它们的陪域有所扩展。

通过要求实值赋值与格的结构相容，我们得以明确准概率的推导与概率的推导之间的一项差异。为此，我们可以将讨论限制在实值赋值，而非上一节引入的双赋值 $\mathcal{R}$ ，并用 $\mathcal{V}(x)$ 替换 $\mathcal{R}(x,t)$。

沿用 B.3.2 节中的相同方法，我们旨在确定函数 $\mathcal{F}$ 的一般形式，使其对于两个陈述 $a,\ b$ 满足 $a \wedge b = \bot$:

$$\mathcal{V}(a \vee b) = \mathcal{F}\left[\mathcal{V}(a), \mathcal{V}(b)\right], \tag{92}$$

类似于方程 (67)。按照上一节的方法，我们发现 $\mathcal{F}$ 满足

$$\mathcal{F}\left[x, \mathcal{F}[y,z]\right] = \mathcal{F}\left[\mathcal{F}[x,y], z\right], \tag{93}$$

如 (71)所示。与上一节一样，我们的目标是求解函数方程 (93)。我们假设 $\mathcal{F}$ 是从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的一般函数，并且是联合连续的。这个问题的解揭示了，在要求与格的规则相容时，实值赋值如何作用于复合陈述。这个问题的一般解并未表现出概率所具有的典型行为。因此，仅仅要求与格的结构相容，并不能唯一地将概率论确定为格上真值赋值的一种放宽形式。现在，我们对 $\mathcal{V}$ 施加一些额外的约束，从而也对 $\mathcal{F}$施加相应约束，以展示这些约束如何导出概率赋值。这些额外约束，即常识性假设，其依据在于将 $\mathcal{V}$ 解释为 可信度 ，以及对其适当行为的定性直觉。赋予 $\mathcal{V}$ 某种解释，就意味着施加额外限制，从而偏离对放宽逻辑赋值这一概念的最一般理解。可以认为，关于可信度应如何表现的直觉，源于我们对实验数据的经验，以及在概率的有限频率主义解释中对这些数据的处理[14]。由于对我们而言，不可访问陈述对于实验者而言是隐藏的，因而也不属于任何频率主义解释的范围，所以我们不要求对不可访问陈述的赋值满足这些额外要求。然而，对于实验上可及的陈述，可信度这一概念表现出某些明确的定性行为。这些额外约束或常识性假设如下：

请求1： 如果必须将 $\mathcal{V}$ 解释为“可信度”，那么显然，给一个陈述赋予高于真陈述的可信度，或低于假陈述的可信度，都是没有意义的。因此，如果 α 是 $\mathcal{V}$ 赋予假陈述的值，而 β 是 $\mathcal{V}$ 赋予真陈述的值，我们希望将 $\mathcal{V}$ 的陪域限制在闭区间 $[\alpha,\beta]$ 内。这意味着 $\mathcal{F}$ 必须是从 $[\alpha,\beta]^2$ 到 $[\alpha,\beta]$ 的函数。

请求2： 如果必须将 $\mathcal{V}$ 解释为“可信度”，那么如果 $\mathcal{V}(a)=\beta$，也就是说，陈述 a 可以被认为是真的，那么无论 $\mathcal{V}(b)$的值如何，$\mathcal{V}(a\vee b)=\mathcal{F}\left[\mathcal{V}(a),\ \mathcal{V}(b)\right]=\beta$ 也必须被认为是真的，即使 $\mathcal{V}(b)\neq\alpha$ 且 $\mathcal{V}(b)\neq\beta$。

请求3： 如果$\mathcal{V}$ 必须被解释为“可信度”，那么我们通常预期$\mathcal{V}(\alpha\vee\beta)\geq\max\{\mathcal{V}(a),\ \mathcal{V}(b)\}$，且只有在$\mathcal{V}(a)=\alpha$或$\mathcal{V}(b)=\alpha$时才取等号，因为组合陈述时，可信度应当增加。

要求 F 是满足结合律方程 (93) 的联合连续函数，再加上请求1、2、3，这对应于表示定理 2.3.4.(d)[15] [33] 的假设。（特别是，请求1对应于对 F 的定义域和陪域的规定；请求2连同请求 (91) 对应于 $G(a,a)=a$ 和 $G(u,b)=G(b,u)=b$ 中 [33] 的要求；请求3则意味着 F 没有内部幂等元。根据引理 1.3.9， [33] 这是要求 F 满足阿基米德性的等价条件。）

因此，函数$\mathcal{F}$ 可表示为以下形式

$$\mathcal{F}(x,y)=f^{-1}\left(f(x)+f(y)\right), \tag{94}$$

其中， f从 $[\alpha,\beta]$到 $\mathbb{R}^{+}$连续且严格递增，且 $f(\alpha)=0$。重要的是，给定$\mathcal{V}$后，函数 f在乘法常数意义下唯一（见 [33]中的推论 2.2.6）。这使我们可以对$\mathcal{V}$重新分级为 $\hat{}\mathcal{V}=f\circ\mathcal{V}$，从而对于每个陈述 a，都有 $\hat{}\mathcal{V}(a)\geq 0$；并且对于任意两个满足 $a\wedge b=\bot$ 的陈述 a,b，都有 $\hat{}\mathcal{V}(a\vee b)=\hat{}\mathcal{V}(a)+\hat{}\mathcal{V}(b)$。由此可见，实值估值的正值性源于附加假设，这些假设是为了使实值估值能够表现为“可信度”而作出的合理设定；“可信度”在定性上与概率相似。相比之下，若仅将实值估值视为逻辑估值的放宽，则限制较少，也允许实值估值取负值。

此外，我们指出，仅对可访问子格提出附加假设，会在 $\mathcal{V}$ 信息的定义中编码有关该格可达性赋值的信息，从而将放宽后的真值标签与可访问标签联系起来。

[14]这种解释在概率论的推导中引入了一种频率主义解释的意味。这就是为什么一些作者（如 Jaynes）试图通过引入更复杂的哲学框架，来论证关于可信度应如何表现的直觉。分析这些直觉的来源不在本附录的讨论范围之内。

[15]如果我们选择了 $\alpha>\beta$，那么我们就会应用表示定理 2.3.4.(a)，从而得到类似的结果，不同之处在于，越接近真理的可信度将对应越小的数值。

到目前为止，我们一直在讨论格上的逻辑赋值和实值赋值，但在上一节和正文中，为了从格的规则推导出准概率，我们使用了双赋值$\mathcal{R}$。由于双赋值$\mathcal{R}$ 在其第二个自变量固定为常值命题 $t \neq \perp$时表现为实值估值，因此刚才给出的论证可以直接对应到我们在B.3.2节中对准概率的推导。通过刚才的论述，我们凸显了我们从格的规则推导准概率的方法，与基于 Knuth、Skilling 和 Caticha 方法的概率论推导之间的主要区别。

要突出说明我们推导准概率的方法与使用 Cox 和 Jaynes 的方法推导概率之间的出发点差异，任务更具挑战性。

Cox 和 Jaynes 的方法从推导乘法规则开始。与我们的方法类似，Jaynes 和 Cox 都引入了双赋值，并在没有明确提出实值赋值的情况下进行推导。虽然重建乘法规则并非严格需要引入双赋值，但这样做可以简化过程。因此，我们在正文中也从双赋值入手。使用实值赋值重建乘法规则更为复杂，超出了本文的讨论范围。不过，由于 Cox 和 Jaynes 从逻辑推导概率的方法最为人熟知，我们想指出，不采用他们的常识假设，也就意味着不采用我们的常识假设。

仿照 Cox 和 Jaynes 的做法，首先要在尚未得知求和规则的情况下，求出方程 (78) 中的函数 G 。沿用我们推导求和规则时采用的推理思路，可以发现G 满足结合律函数方程。此外，还可以提出其他要求，或像推导求和规则时那样采用常识假设，最终得到 G ，它是一个从 $[0,1]^2$ 到 $[0,1]$, 的联合连续函数，具有结合律、阿基米德性，并且对每个 $x \in [0,1]$， $G(x,0) = G(x,0) = 0$以及 $G(1,1) = 1$。利用 [33] 中的定理 2.3.4.(a)和引理 2.2.1，可以得到乘法规则。Jaynes 在其推导中采用了不同的常识假设，尤其要求函数 G 对两个自变量都递增。借助引理 2.1.1 和引理 2.2.1，可以证明，我们的常识条件意味着G 必须对两个自变量都递增。因此，如果我们不采用 Jaynes 的常识条件，即 G 对两个自变量都递增，那么这就意味着不采用我们自己的常识假设。

最后，我们通过展示一个确实可能存在的例子来证明：可以构造一个具有混合可达性值的系统，使得 G 并不随两个参数同时增加。考虑一个包含 4 个原子陈述 s_1、s_2、s_3、s_4 的系统，其对应的准概率向量为 $\vec{q} = \{Q(s_1)\ Q(s_2)\ Q(s_3)\ Q(s_4)\} = \{\ ^{x}(_{x-1})\ 4\ ,\ ^{2-x}_{4}, -\ ^{x-1^2}_{4}, 3\ 4\ \}$，其中$x \in [1-\sqrt{3}, 1+\sqrt{3}]$。我们认为 $A = s_1 \vee s_2$ 可达，$B = s1 \vee s_3$ 不可达， $C = A \vee B = s_1 \vee s_2 \vee s_3$不可达。我们可以计算出 $Q(A \vee B|C) = x(x-1)$、 $Q(B|C) = x-1$ 以及 $Q(A|BC) = x$ ，并看到$Q(A \vee B|C) = G[Q(B|C)\ Q(A|BC)]$。在区间 $x \in [1-\sqrt{3}, \frac{1}{2}]$ 内，函数 G 在其两个参数都增加时却在减小。

C 逻辑格的弛豫方案

	逻辑格	准概率格
真值	为每个陈述赋予二进制值。	为每个陈述赋予实数值。
可达性	为每个陈述赋予二进制值。	对可获取概率的限制 陈述。
假设 1	不可访问原子的数量 获得一个可达 陈述所需的陈述数下界为 可访问性深度 d。	模型的状态 $\vec{q}$ 的不可访问性 下界为 可访问性深度 d。

表 3：逻辑格与准概率格的比较。

D 不可访问性度量

D.1 不可访问性度量

关于信息不可及性的主要假设告诉我们，对于给定的可访问性深度 d，我们至少对 d 个元素完全无知，或者更准确地说，有 d 个元素对于我们的知识而言是不可访问的。在本节中，我们将寻找一种能够量化这种不可访问性程度的度量。我们将这一度量称为 不可访问性度量。有了定义明确的不可访问性度量，我们就能将假设 1 形式化为对允许的准概率向量 $\vec{q}$施加的约束。

我们用希腊字母表示不可访问性度量 $\mathcal{X}$。下面我们将把$\mathcal{X}$ 刻画为定义在单纯形的概率向量上的函数 Δ_D，其中 D 是单纯形维数[16]。显然，在 Δ_D 上，不可访问元素的最大数量为 D（所有元素都不可访问，我们一无所知），而最小数量为 1 （我们可以访问每个元素，也就是说，我们掌握了最多的信息）。

我们希望找到一个满足以下要求的函数 $\mathcal{X}$： $\Delta_D \to [1, D]$ 。

请求（1）：计数性。 对于每个整数 $d \geq 1$ ，都可以为任意 Δ_D 定义具有完全不确定性的、基于 $d \leq D$ 个元素的标准状态，表示为

$$\vec{p}_{\frac{1}{d}} = (\underbrace{\frac{1}{d}, \ldots, \frac{1}{d}}_{d}, 0, \ldots, 0). \tag{95}$$

我们要求的是

$$\mathcal{X}(\vec{p}_{\frac{1}{d}}) = d, \tag{96}$$

$\vec{p}_{\frac{1}{d}}$ 中不可访问元素的数量为 d。

要求 (2)：单调性。 我们要求 $\vec{p}_{\frac{1}{d}} \in \Delta_D$ 是在 Δ_D中具有极大不可访问性的向量

$$\forall \vec{p} \neq \vec{p}_{\frac{1}{D}} \in \Delta_D, \quad \mathcal{X}(\vec{p}_{\frac{1}{D}}) > \mathcal{X}(\vec{p}). \tag{97}$$

要求 (3)：对称性。 重新标记元素不应改变不可访问元素的数量。因此，我们要求不可访问性度量对于概率向量的任意置换保持不变。用 P_π 表示与置换 π 对应的置换矩阵，我们要求

$$\mathcal{X}(P_\pi \vec{p}) = \mathcal{X}(\vec{p}), \qquad \forall \text{ permutation } \pi. \tag{98}$$

要求 (4)：乘法性。 考虑具有完全不确定性的两个配置 $\vec{p}_{\frac{1}{d_1}} \in \Delta_{D_1}$ 和$\vec{s}_{\frac{1}{d_2}} \in \Delta_{D_2}$。 如果对于 $\vec{p}_{\frac{1}{d_1}}$ ，我们对 d_1 个陈述完全不确定，而对于 $\vec{s}_{\frac{1}{d_2}}$ ，我们对 d_2 个陈述完全不确定，那么我们预期复合配置 $\vec{p}_{\frac{1}{d_1}} \otimes \vec{s}_{\frac{1}{d_2}} = \vec{r}_{\frac{1}{d_1 d_2}} \in \Delta_{D_1 D_2}$ 对 $d_1 \cdot d_2$ 个陈述完全不确定。一般而言，我们预期组合两个系统时不可访问元素的数量会相乘。因此，我们要求函数 $\mathcal{X}$具有乘法性：

$$\mathcal{X}(\vec{p} \otimes \vec{s}) = \mathcal{X}(\vec{p})\mathcal{X}(\vec{s}) \tag{99}$$

对于任意两个有限概率单纯形中的 $\vec{p}$ 和 $\vec{s}$ 。

如果能够刻画出所有符合我们要求的函数，那就再好不过了。遗憾的是，我们无法给出一般性的证明，因此需要依赖对一个假设形式的选择。尽管如此，所选的假设形式仍相当一般。

[16]为了避免在推理准概率时因缺乏直觉而受到误导，我们似乎不得不在概率向量空间上进行研究。在概率向量这一受限定义域上，不可访问性的概念似乎并无意义，因为讨论不可访问性需要使用准概率的语言。若仅限于概率向量空间，这种信息度量更适合称为不可区分性度量。然而，我们最终的目标是将这种信息度量的定义域扩展到准概率向量空间；在这一语境下，称其为不可访问性度量 是合理的，因此我们决定现在就将这种度量命名为不可访问性度量。

设想 我们要求函数$\mathcal{X}$具有如下形式

$$\mathcal{X} = (\sum_{i=1}^{D} f_i(p_i))^z, \tag{100}$$

对于指定的一组可测函数 f_i 和一个参数 $z \in \mathbb{R}$。

引理 9。 对于每个 $\vec{p} \in \Delta_{D\geq 3}$，由请求 *(1,2,3)* 表征并以 *Ansatz* 形式写出的最一般 *X*，构成如下定义的不确定性度量的 1-参数族 X_c

$$\mathcal{X}_c(\vec{p}) = \frac{1}{\sqrt[c-1]{\sum_{i=1}^{D} p_i^c}}, \tag{101}$$

其中 $c \in \mathbb{R}_+$ 和 $c \neq 1$。

证明。 根据置换下的对称性（要求（3）），我们有

$$\mathcal{X}(\vec{p}) = (\sum_{i=1}^{D} f(p_i))^z, \tag{102}$$

也就是说，只有一个单独的函数 f。

现在考虑两个概率向量 $\vec{p} \in \Delta_{D_1}$ 和 $\vec{q} \in \Delta_{D_2}$。根据乘法性要求（要求 (4)），我们有

$$\sum_{(i,j)=(1,1)}^{(D_1,D_2)} f(p_i q_j) = \sum_{i=1}^{D_1} f(p_i) \sum_{j=1}^{D_2} f(q_j). \tag{103}$$

这是一个函数方程，可以利用柯西函数方程的解来求解[35, 36, 37]。我们知道，对于 $\vec{p} \in \Delta_{D_1\geq 3}$ 和 $\vec{q} \geq \Delta_{D_2\geq 2}$，最一般的可测 $f\colon\ [0,1] \to \mathbb{R}$ 满足函数方程（103）的形式如下 [35, 36, 37]:

$$f(p) = ap + b \quad \text{if } p \in [0,1] \tag{104}$$

$$f(p) = \begin{cases} 0 & \text{if } p \in [0,1) \\ 1 & \text{if } p = 1 \end{cases} \tag{105}$$

$$f(p) = \begin{cases} p^c & \text{if } p \in (0,1] \\ 0 & \text{if } p = 0 \end{cases}, \tag{106}$$

其中 $c \in \mathbb{R}$ 和 a，b 是两个常数，满足 $a + D_1 D_2 b = (a + D_1 b)\ (a + D_2 b)$。解（104）不能接受，因为常数 a 和 b 取决于 D_1 和 D_2，而我们不允许 $\mathcal{X}$ 依赖于单形的维度。解（105）也必须舍弃，因为它不满足要求（1）。因此，我们只剩下解（106）。

将（106）代入（102），得到一组连续的解$\mathcal{X}_c$，由实参数 c参数化。施加计数性条件（要求（1））相当于寻找一组值 z 和 c，使得

$$\mathcal{X}_c(\vec{p}_{\frac{1}{d}}) = \left(\sum_{i=1}^{d} \frac{1}{d^c}\right)^z = \frac{1}{d^{z(c-1)}} \coloneqq d. \tag{107}$$

我们得到 $z = \frac{1}{1-c}$和 $c \neq 1$。最后，对于解的单调性（请求 (2)），我们有 $c > 0$，由此证明得证。

□

我们将概率向量的不可访问性度量定义如下。

定义 10（不可访问性度量）。 对于任意概率单纯形 Δ_D，不可访问性度量是一个函数$\mathcal{X}_c$，其定义域为 Δ_D，陪域为 $\mathbb{R}_+$，定义为

$$\mathcal{X}_c(\vec{p}) = \frac{1}{\sqrt[c-1]{\sum_{i=1}^{D} p_i^c}}, \tag{108}$$

其中 $c \in \mathbb{R}_+$ 和 $c \neq 1$。

D.2 不可访问信息的准概率模型中的不可访问性度量扩展

不可访问性度量族 $\mathcal{X}_c$ 的推导基于这样一个假设：我们感兴趣的是测量概率向量的不可访问性大小。然而，我们的不可访问信息理论中的每个模型，都是由其对应的准概率向量 $\vec{q}$ 描述，而非由概率向量描述。此外，假设 1 指的是编码在 $\vec{q}$中的不可访问性大小。

有必要找到$\mathcal{X}_c$ 在我们的不可访问信息理论的准概率向量空间上的延拓。为此，我们将利用 $\vec{q}$ 最大纠缠态模型的这一性质，即它们与层级 d 处的可及概率分布之间存在一一对应关系（见引理 8）。

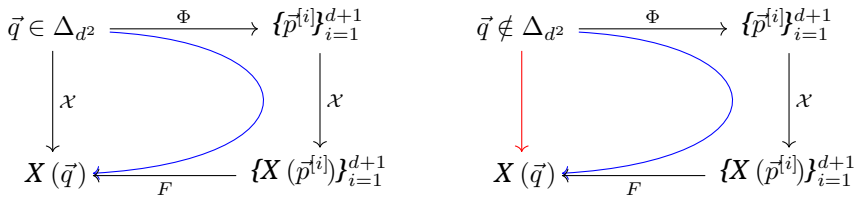

图 9：如何计算可访问性深度为 d 的 MES 模型中状态的不可访问性。当向量 $\vec{q}$ 是概率向量时，可以直接计算不可访问性 $\mathcal{X}(\vec{q})$。或者（沿蓝色箭头），可以使用一一映射 Φ 将 $\vec{q}$ 映射到与之相关的可达概率向量集合 $\{\vec{p}^{[i]}\}_{i=1}^{d+1}$，计算其中每个向量的不可访问性，再通过映射 F 取回在原始空间 $\vec{q}$ 上计算出的不可访问性值。对于可访问性深度为 d 的 MES 模型，向量 $\vec{q}$ 并不总是 Δ_{d^2} 的元素；在这些情况下，$\vec{q}$ 的不可访问性没有明确的定义（红色箭头）。然而，蓝色箭头仍然是一条可行路径，因为 Φ 总会将可访问性深度为 d 的 MES 模型中的每个 $\vec{q}$ 映射到一组合法的概率向量 $\{\vec{p}^{[i]}\}_{i=1}^{d+1}$，其不可访问性有明确定义。因此，可访问性深度为 d 的 MES 模型所允许的所有准概率向量上的不可访问性度量延拓，定义为复合 $F \circ \mathcal{X} \circ \Phi$（蓝色箭头）。

限制 $\vec{q}$ 为概率分布后，我们发现，对于 c的某些取值，可以将$_c(\vec{q})$ 的不可访问性$\mathcal{X}$表示为$\mathcal{X}_c$ 对$\{\vec{p}^{[i]}\}_{i=1,\ldots,d+1}$的不可访问性的函数，其中 d 是 MES 模型在层级 d 处所对应的可及概率分布。该过程如图9所示。我们将 Φ 称为一个可逆映射，它将可访问性深度为 d 的 MES 模型的准概率向量 $\vec{q}$映射到其对应的可及概率分布集合$\{\vec{p}^{[i]}\}_{i=1}^{d+1}$。由于根据定义，即使 $\vec{q}$取负值，$\{\vec{p}^{[i]}\}_{i=1,\ldots,d+1}$ 始终都是概率分布，因此我们总能计算它们的不可访问性。接下来，我们只需找到函数 F，在 $\vec{q} \in \Delta_D$的情况下，将针对$\{\vec{p}^{[i]}\}_{i=1,\ldots,d+1}$ 计算得到的不可访问性集合映射到 $\vec{q}$的不可访问性。我们使用该方法计算 MES 模型的准概率向量的不可访问性。我们发现，将不可访问性度量扩展到准概率向量，相当于约束参数 c 的取值，具体如下定义所述

定义 11（MES 模型的不可访问性度量）。对于任意可访问性深度为 d 的 MES 模型，不可访问性度量由函数 $\mathcal{X}$ 映射 $\vec{q}$ 至 $\mathcal{X}(\vec{q}) \in \mathbb{R}_+$ 显式表示为

$$\mathcal{X}(\vec{p}) = \sum_{i=1}^{D} \frac{1}{p_i^2}, \tag{109}$$

下面我们将详细了解如何从概率向量的一般不可访问性度量推导出 MES 模型的不可访问性。

对于最大纠缠态模型，向量 $\vec{q}$ 与层级 d 上极大可访问概率分布集合 $\{\vec{p}_{\vec{q}}^{[i]}\}_{i=1,\ldots,d+1}$ 一一对应（见引理 8）。我们希望将$\vec{q}$中编码的不可访问性程度与每个分布 $\{\vec{p}_{\vec{q}}^{[i]}\}_{i=1,\ldots,d+1}$的不可访问性程度联系起来。也就是说，我们希望找到一族函数F_c：$\mathbb{R}^{d+1} \to [1, d^2]$，使得

$$F_c\left(\{\mathcal{X}_c(\vec{p}^{[i]})\}_{i=1}^{d+1}\right) = \mathcal{X}_c(\vec{q}). \tag{110}$$

引理 10. 对于可访问性深度为 d 的最大纠缠态模型，若将 $\vec{q}$ 限定为概率向量，则对于 $c \in \{2, 3, \dots\}$，$X_c(\vec{q})$ 可使用以下递归方程由 $\{\vec{p}_{\vec{q}}^{[i]}\}_{i=1,\dots,d+1}$ 计算得出。

$$\mathcal{X}_c(\vec{q}) = \sqrt[1-c]{\frac{\sum_{i=1}^{d+1} \left(\mathcal{X}_c(\vec{p}^{[i]})\right)^{1-c} - G_c(\vec{q})}{d+2-2^{c-1}}}, \tag{111}$$

其中

$$G_c(\vec{q}) = \begin{cases} g_c(\vec{q}) & if \quad \mod(c,2) = 1 \\ g_c(\vec{q}) + \frac{1}{2}\binom{c}{\frac{c}{2}}^{1-\frac{c}{2}}\sqrt{\mathcal{X}_{\frac{c}{2}}(\vec{q})} & if \quad \mod(c,2) = 0 \end{cases}, \tag{112}$$

$$g_c(\vec{q}) = \sum_{i=1}^{\lfloor \frac{c-1}{2} \rfloor} \binom{c}{i} \sqrt[1-i]{\mathcal{X}_i(\vec{q})}\, \sqrt[1-c+i]{\mathcal{X}_{c-i}(\vec{q})}, \tag{113}$$

且 $\sqrt[0]{X_1(\vec{q})} := 1$。

证明。 定义以下量很有用

$$S_c(\vec{p}) := \sum_i p_i^c = (\mathcal{X}_c(\vec{p}))^{1-c}. \tag{114}$$

我们可以改写为

$$\mathcal{X}_c(\vec{q}) = \frac{1}{\sqrt[c-1]{S_c(\vec{q})}}, \tag{115}$$

$$\tilde{S}_c(\vec{q}) = \sum_{i=1}^{d+1} S_c(\vec{p}^{[i]}). \tag{116}$$

为了证明该引理，我们用 $\tilde{S}(\vec{q})$ 表示 $S_c(\vec{q})$。由于每个 $p_j^{[i]}$ 都是 $\vec{q}$ 中元素的和，根据二项式定理，我们有~：对于整数 c， S 可以展开为由 $\vec{q}$中元素的乘积构成的有限 和，各元素的幂为介于 0 和 c 之间的整数。这一事实使我们能够通过限制在 $c \in \{0, 2, 3, 4, \dots\}$ 并显式计算 S_c 的表达式，得到 $\vec{q} \in \Delta_d$:

$$\begin{aligned} S_1 &:= 1, \\ \tilde{S}_2 &= dS_2 + S_1, \\ \tilde{S}_3 &= (d-2)S_3 + 3S_2S_1, \\ \tilde{S}_4 &= (d-6)S_4 + 4S_3S_1 + 3S_2S_2, \\ \tilde{S}_5 &= (d-14)S_5 + 5S_4S_1 + 10S_2S_3, \\ \tilde{S}_6 &= (d-30)S_6 + 6S_5S_6 + 15S_4S_2 + 10S_3S_3, \\ \tilde{S}_c &= (d-(2^{c-1}-2))S_c + \sum_{i=1}^{\lfloor \frac{c-1}{2} \rfloor} \binom{c}{i} S_iS_{c-i} + \frac{f_c}{2}, \end{aligned} \tag{117}$$

其中

$$f_c = \begin{cases} 0 & \text{if} \quad \mod(c,2) = 1 \\ \binom{c}{c/2} S_{c/2}S_{c/2} & \text{if} \quad \mod(c,2) = 0 \end{cases}, \tag{118}$$

~且其中 S_i、 S_i 和 f_c 都是 $\vec{q}$的函数。将方程（117）代入方程（115），即可得到引理中的公式。

□

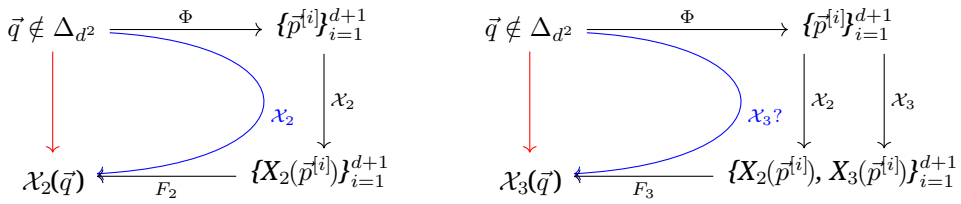

图 10：要将 $\mathcal{X}_2$ 延拓到可达性深度为 d 的 MES 模型的容许态上，我们只需定义概率单纯形上的 $\mathcal{X}_2$。要将 $\mathcal{X}_3$ 延拓到可达性深度为 d 的 MES 模型的容许态上，我们需要定义概率单纯形上的 $\mathcal{X}_3$，还需要定义 $\mathcal{X}_2$。

例如，我们有：

$$\mathcal{X}_2(\vec{q}) = \frac{d}{\sum_{i=1}^{d+1} \frac{1}{\mathcal{X}_2(\vec{p}^{[i]})} - 1}, \tag{119}$$

$$\mathcal{X}_3(\vec{q}) = \sqrt{\frac{d-2}{\sum_{i=1}^{d+1} \frac{1}{\mathcal{X}_3(\vec{p}^{[i]})^2} - \frac{3}{\mathcal{X}_2(\vec{q})}}}. \tag{120}$$

我们看到，对于 $c = 2$，我们一直在寻找的函数 F_2 确实存在，因此 $\mathcal{X}_2$ 可以延拓到可达性深度为 d 的 MES 模型所容许的所有准概率向量。对于 $c > 2$，函数 F_c 的参数必须是相关可达概率向量 $\{\mathcal{X}_{\tilde{c}}(\vec{p}^{[i]})\}_{i=1}^{d+1}$ 的可达性，该值针对所有 $\tilde{c} \leq c$ 计算得出。

这意味着，对于 $c > 2$，不可访问性度量 $\mathcal{X}_c$ 无法在未先定义不可访问性度量 $\mathcal{X}_{c-1}$的情况下定义于准概率之上；而后者又无法在未先定义不可访问性度量 $\mathcal{X}_{c-2}$ 的情况下定义，如此继续，直到达到不可访问性度量 $\mathcal{X}_2$的定义。

例如，$\mathcal{X}_3(\vec{q}) = F_3(\{\{\mathcal{X}_2(\vec{p}^{[i]})\}_{i=1}^{d+1}, \{\mathcal{X}_3(\vec{p}^{[i]})\}_{i=1}^{d+1}\})$ 正如方程中显而易见的 (120)（参见图 10）。

这一事实表明，MES 模型存在一个不可访问性层级。事实上，$\mathcal{X}_c$ 只有在考虑所有$\mathcal{X}_{\tilde{c}}$ 满足 $\tilde{c} < c$ 的情况下，才对 MES 模型的所有可容许准概率向量都有定义。图 10 给出了图示。本研究聚焦于不可访问性 $\mathcal{X}_2$，因为它是唯一能够独立存在的不可访问性，因此具有唯一性[17]。

D.3 不可访问性的性质

如前所述，不可访问性度量的下界始终为 1，上界则为单纯形维数。事实上，一个人可能无法确定其所属元素的最小数量为 1，而最大数量则是单纯形维数所支持的元素最大数量。

不可访问性既非凸也非凹，但在 MES 模型的状态集合上是准凹的，下一引理将证明这一点。

引理 11。 给定属于最大纠缠态$_d$ 的两个状态 $\vec{p}$、$\vec{q}$，且 $X_c(\vec{p}) \geq d$和 $X_c(\vec{q}) \geq d$，其中 $c \in \{2、4、6、.....\}$，则对于 $\lambda \in [0,1]$有

$$\chi_c\left(\lambda\vec{p} + (1-\lambda)\vec{q}\right) \geq \min\left(\chi_c(\vec{p}), \chi_c(\vec{q})\right). \tag{121}$$

不可访问性是 *MES* 模式的状态集合上的一个拟凹函数。 *l.*

证明。 不失一般性，考虑 $\mathcal{X}_c(\vec{p}) \leq \mathcal{X}_c(\vec{q})$。利用 $\mathcal{X}_c(\vec{q}) = (\|\vec{q}\|_c)^{\frac{c}{1-c}}$ 这一事实，我们有$\|\vec{q}\|_c \leq \|\vec{p}\|_c$和

$$\begin{aligned}\chi_c\left(\lambda\vec{p} + (1-\lambda)\vec{q}\right) &= \left(\|\lambda\vec{p} + (1-\lambda)\vec{q}\|_c\right)^{\frac{c}{1-c}} \geq \\ &\geq \left(\lambda\|\vec{p}\|_c + (1-\lambda)\|\vec{q}\|_c\right)^{\frac{c}{1-c}} \geq \\ &\geq \left(\|\vec{p}\|_c\right)^{\frac{c}{1-c}} = \mathcal{X}_c(\vec{p}).\end{aligned} \tag{122}$$

□

此外，总存在 $\vec{p}_d^{\,1}$ 的一个邻域，使得对于每个允许的 c，$\mathcal{X}_c$ 都是凹的。

[17]众所周知，2-范数是“特殊的”，并且在某种程度上无处不在，因为有许多充分理由更倾向于使用它们。关于其中一些理由的精彩论述，可以参见[38]。

D.4 不可访问性度量与其他信息度量之间的关系

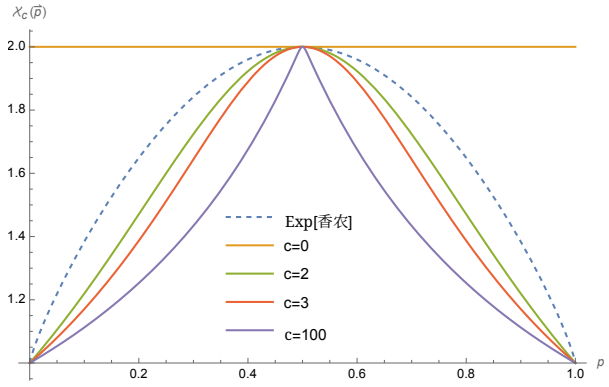

图11：不可访问性 $\mathcal{X}_c(\vec{p})$，其中 $\vec{p}=(p,\ 1-p)$。

在图 11中，我们绘制了不同参数值 c下不可访问性度量的变化情况。 $\mathcal{X}$ 与熵度量之间的主要区别在于请求 (4)。事实上，熵度量通常被要求具有可加性。我们可以通过以下关系将 $\mathcal{X}$ 与香农熵联系起来：$\log(\lim_{c\to 1}\mathcal{X}_c)$；通过以下关系将其与 Rényi 熵联系起来：$\log(S_\alpha)$；并通过以下关系将其与 Tsallis 熵联系起来： $1-(S_\alpha)^{1-c}$。

在文献中，不可访问性度量 $\mathcal{X}$ 曾用于不同语境，被称为希尔数或多样性指数 [20, 21]。在量子文献中，与 $\mathcal{X}_2$ 类似的概念是有效维度 [22, 23] 或操作不变信息 [24]。据作者所知，此前尚无人像附录 $\mathcal{X}$ D.1 中那样从逻辑请求推导出 X，而将其用作希尔数或多样性指数，也仅得到启发式论证的支持。

E 趣闻

为什么只考虑理想配置? 当然可以研究任何可能的配置，但我们决定聚焦于理想配置。我们这样做主要有两个原因

1. 我们希望最大化实验相关信息的数量；
2. 考虑图 (4,1) 中的模型配置 12。在这个模型中，经典类比的原子陈述所包含的“绝对”知识并不平衡。事实上，在不可访模型中，关于 c和 d的知识比关于 $a\vee b$的知识更为根本。一旦转向经典模型，这种区别便消失了。为了在经典模型及其向不可访信息模型的膨胀之间保持这种“对称性”，我们使用系统的可访问性深度 d 。因此，将经典模型膨胀为不可访信息模型时，我们不会给不同的陈述赋予不同的“绝对”知识。

为什么不使用复准概率? 在附录 B.3.2 的准概率推导中，没有明显的理由要求函数 R 的陪域限于实数；也可以考虑陪域为 $\mathbb{C}$ 的函数 R。这样一来，我们会得到一个函数 Q ，其行为与我们在附录 B.3.2 中推导出的函数完全相同，但它会将格中的每个陈述映射到一个复数。在这种情况下，最大纠缠态模型的状态 $\vec{q}$ 将是一个复向量。我们为什么排除这种可能性? 对于可访问性深度为 d 的最大纠缠态模型，我们希望层级 d 上的 $d(d+1)$ 个可访问陈述取值为区间 $[0,1]$ 内的实数。这意味着，对于每个可访问陈述，原子陈述的虚部之和必须恰好为零。若将状态向量 $\vec{q}$ 中每个元素的虚部记为 $\mathfrak{q}_i=\mathrm{Im}(q_i)$，那么考虑层级 d 上的 $d(d+1)$ 个可访问陈述，可得 $d(d+1)$ 个形如 $\mathfrak{q}_{i_1}+\mathfrak{q}_{i_2}+\cdots+\mathfrak{q}_{i_d}=0$ 的约束，其中 $\{i_1,\ldots,i_d\}$ 是构成一个可访问陈述的指标集合。因此，虚部必须为零。

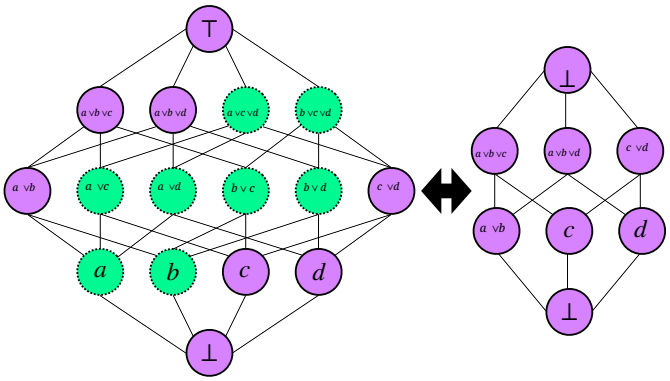

图 12：具有非理想配置的（4,1）模型的非对称膨胀。