arXiv:2603.11083v1 [cs.LO] 11 Mar 2026

时序逻辑与自动机理论中的概率析取范式

ALEXANDER KUZNETSOV

俄罗斯科学院 V.A. Trapeznikov 控制科学研究所，Profsoyuznaya 街 65 号，莫斯科，117997，
俄罗斯*e-mail address*：avkuz@bk.ru

Abstract. 本文介绍了概率析取范式（PDNFs），将其作为表示和推理逻辑系统中不确定性的框架。不同于经典析取范式，PDNFs 为变量赋予实值权重，用于编码有关变量存在、缺失或被否定的概率信息。随后，我们构造了一个 PDNFs 的向量空间，从而实现代数证据组合。PDNFs 被解释为 venjunctions（时序逻辑构造）的概率分布，以及分割区间上的可积函数，其中积分决定变量概率。这一双重视角使其具备巴拿赫空间结构，并能够应用泛函分析。我们证明，在指数参数化下，PDNF 的加法与贝叶斯证据融合相一致，并推导出根据随机样本识别结果的界限。因此，该形式体系连接了逻辑、数值方法与连续概率。

1. Introduction

考虑一个包含 m 传感器的系统，并在一系列时间步中对其进行观测。在某个时刻，每个传感器都可能被触发，也可能不被触发；多个参数都可能影响其响应——既包括用户最初设定的参数（如运动、质量、照度、湿度），也包括随机因素。用户不知道特定随机因素是否存在，但确切知道哪些因素可能影响设备的运行，以及它们会如何产生影响。我们需要建立一种数学形式体系，用于在不了解随机因素影响的情况下描述这类传感器系统的运行，同时又能了解传感器对每种因素的响应方式。从本质上说，我们应实现一种复杂的传感器估计机制，类似于[SK21]。

本文提出的形式体系可以与更广泛的不确定性推理模型联系起来，例如概率布尔网络、马尔可夫逻辑网络和邓普斯特-谢弗理论。

概率布尔网络 [SDKZ02、Dro08] 模型通过为每个变量关联一个可能布尔更新函数的概率分布来描述系统。相比之下，我们的模型为每个变量关联一个实值参数，以控制该变量在逻辑合取中的随机出现。

Key words and phrases: 概率逻辑、决策理论、传感器网络、不确定性、时序逻辑。

预印本，已提交至《计算机科学中的逻辑方法》

© A.库兹涅佐夫cc 知识共享

马尔可夫逻辑网络 [RD06]为一阶逻辑公式赋予权重，其中权重表示约束的强度；一个世界的概率与其满足的公式权重之和的指数成正比。然而，我们的方法并不为整个公式赋予权重，而是为析取范式中的组成变量赋予权重。

邓普斯特-谢弗理论 [Sha76] 通过独立的质量函数量化信念与似然性，为证据组合提供了一种强大但计算量大的框架。我们的形式体系则提供了一种更简单、以代数为基础的替代方案：通过一族参数函数，直接将不确定性编码到逻辑表达式的变量中。

另一方面，我们提出了一种在某种程度上让人联想到模糊逻辑[Nil86], 的设计；不过，我们考察的不是公式中变量的真值程度，而是变量出现在公式中的概率，并且可以根据传感器的设计特征计算这一概率。

在我们的模型中，第一个层次的不确定性由映射**F**给出；对于一组固定的权重 Ξ ，它决定特定文字出现的概率（相当于一型模糊集中的隶属度）。第二个层次的不确定性源于权重 Ξ本身可能是随机的，即，可能的时序逻辑结构（venjunctions）集合上的概率分布本身也是随机的。这与二型模糊集的概念完全对应；在二型模糊集中，不确定性由“次级隶属函数” [KM01]描述。因此，我们的形式体系可解释为二型模糊逻辑思想的一种概率实现，其中随机编码器的博赫纳积分发挥着类似于二型模糊集质心的作用，对第二层次的不确定性进行平均。

此外，文章 [SDS02], 提出了一种有些相似的方法，重点研究知识库中的逻辑相关性，并提出了一种知识表示方法，用于发现所选数据特征之间的依赖关系。包含信息特征之间链接的知识库，以带有限谓词的逻辑方程形式表示。

从另一个角度看，所提出的构造可以理解为：利用一个概率公式，定义一组在某种意义上彼此相似的常规公式。我们选择能够最好地描述所关注对象的公式。例如，这对应于为一个类似于已知定理的定理构建证明。从这个意义上说，本研究延续了文章 [SK18]的工作。

最后，所提出的形式体系基于异步时序逻辑 [Vas11] ，用于表达记忆效应，并依赖于系统状态变化的时刻。时序逻辑的理论框架由 Sequention 和 venjunction 等数学工具，以及建立在这些工具之上的逻辑代数方程组成。

近期关于参数化概率自动机（pPA）的研究将参数引入转移概率，从而能够对系统性质进行组合推理[MQK25]。与此同时，知识编译领域的研究（例如关于 d-DNNF 电路的研究）旨在利用布尔电路的结构，为概率推理寻求高效的表示 [DL25]。此外，时序逻辑的扩展（例如 PTATL）使得能够验证不确定性下多智能体系统的策略性质 [JKP+25]。

这些方法分别侧重于对系统动态中的不确定性进行建模（pPA）、实现易处理的推理（d-DNNF），或验证时态性质（PTATL）；而我们的工作则采取了略有不同的方向。

所有可用所提出的形式体系完成的任务，也可以用概率正则表达式或贝叶斯网络来表述，但提出这一形式体系的主要动机是

提供一个既与普通布尔逻辑十分相似、又具有线性空间结构的框架。在我们的框架中，公式看起来像熟悉的逻辑表达式。同时，我们还可以进行加法、数乘等运算，并使用基于范数的相似度度量。本文讨论概率析取范式（该形式体系先前由作者在 [Kuz26]中构建），并研究其基本性质。我们对从概率析取范式的随机实现中完全确定其底层确定性结构所需的观测数量，得出了定量估计。此外，我们还证明了若干关于此类确定性描述可辨识性的定理，展示重复观测如何逐步消除不确定性，最终使我们能够用一份明确的可能行为清单取代概率模型。

尽管这些结果在其他概率框架中也有对应形式，但本文的主要贡献在于系统地建立离散逻辑结构与连续泛函分析之间的联系，从而为将解析方法应用于逻辑问题开辟了新的可能性。简而言之，本文的核心思想是通过可积函数对概率逻辑表达式进行编码。

2. Probabilistic Disjunctive Normal Forms and their Properties

在文章 [Kuz26], 中，作者引入了概率析取范式（PDNF），作为描述具有不确定性的系统的工具。让我们回顾上述文章中的定义，并作少量补充，然后（第 2.1 小节）从离散模型转向连续的逻辑函数模型。

定义 2.1。 我们称由 m 个变量 $x_1, \dots, x_m$构成的概率初等合取为形式 $x_1^{\xi_1} \wedge \cdots \wedge x_m^{\xi_m}$，其中 $\xi_i \in \mathbb{R}, i = 1, m$。假设 $F_1 \colon \mathbb{R} \to [0,1]^m$ 是分量单调递增映射， $F_{-1} : \mathbb{R} \to [0,1]^m$ 是分量单调递减映射，$F_0 \colon \mathbb{R} \to [0,1]^m$, $F_i = (F_i^1, \dots, F_i^m)$， $F_1^j(0) = F_{-1}^j(0) = 0$ 且 $F_{-1}^j(x) + F_0^j(x) + F_1^j(x) = 1$对所有 $x \in \mathbb{R}$, F_i^j 都连续，至多在有限个点处例外， $i = -1, 1$, ~~$j = 1, m$~~。我们定义

$$\mathbb{P}\{x_i^{\xi_i} \sim x_i\} = F_1^i(\xi_i), \tag{2.1}$$

$$\mathbb{P}\{x_i^{\xi_i} \sim \varepsilon\} = F_0^i(\xi_i), \tag{2.2}$$

$$\mathbb{P}\{x_i^{\xi_i} \sim \overline{x}_i\} = F_{-1}^i(\xi_i), \tag{2.3}$$

这里，“$x_i^{\xi_i} \sim x_i$”表示变量 x_i 出现在概率初等合取的实现中，位于 $x_i^{\xi_i}$的位置；而“ $x_i^{\xi_i} \sim \varepsilon$ ”表示变量 x_i 未出现在概率初等合取的实现中，位于 $x_i^{\xi_i}$的位置。如果$\xi_i = 0$ ，那么 x^{ξ_i} 肯定不会出现在基本合取项中。

定义 2.2. 我们将具有权重映射 $\mathbf{F} : \mathbb{R} \to \overline{([0,1]^m)}^3$的 m 个变量的概率析取范式称为形式 $Z = X_1 \vee \cdots \vee X_n$，其中 X_i、 $i = 1, n$是由 m 个变量构成的概率初等合取项，而 $\mathbf{F} = (F_{-1}, F_0, F_1)$由前一定义中的函数构成。我们将这类概率析取范式的集合记为 $\mathcal{B}_{\mathbf{F}}^m(n)$。

我们可以认为所有概率析取范式的长度都相同，因为概率析取范式可以包含任意数量的空概率初等合取项 $x_1^0 \wedge \cdots \wedge x_m^0$。概率析取范式是一种句法

构造；析取符号 $\vee$ 仅仅是有序合取项之间的分隔符。可能的语义解释将在第 3 节和第 6 节中描述。

定义 2.3。 我们将概率初等合取项 $X = x_1^{\xi_1} \wedge \cdots \wedge x_m^{\xi_m}$与数 $\alpha \in \mathbb{R}$ 的乘积称为概率初等合取项

$$\alpha \cdot X = x_1^{\alpha\xi_1} \wedge \cdots \wedge x_m^{\alpha\xi_m}.$$

我们将概率析取范式 $Z = \bigvee_{i=1}^n X_i$ 与数 $\alpha \in \mathbb{R}$ 的乘积称为以下形式

$$\alpha Z = \alpha \bigvee_{i=1}^n X_i = \bigvee_{i=1}^n (\alpha X_i). \tag{2.4}$$

显然，集合 $\mathcal{B}_{\mathbf{F}}^m(n)$ 在上述运算下构成一个线性向量空间。所有含 m 个变量的概率初等合取项的集合可表示为 $\mathbb{R}^m$；所有含 m 个变量和 n 个合取项的概率析取范式的集合可表示为矩阵集合$\mathbb{R}^{n\times m}$。

定义 2.4。 我们将概率初等合取项 $X = x_1^{\xi_1} \wedge \cdots \wedge x_m^{\xi_m}$ 与 $Y = x_1^{\eta_1} \wedge \cdots \wedge x_m^{\eta_m}$ 的和称为概率初等合取项

$$X + Y = x_1^{\xi_1+\eta_1} \wedge \cdots \wedge x_m^{\xi_m+\eta_m}.$$

我们将概率析取范式 $Z_1 = \bigvee_{i=1}^n X_i$、 $Z_2 = \bigvee_{i=1}^n Y_i$ 的和称为以下形式

$$Z_1 + Z_2 = \bigvee_{i=1}^n X_i + \bigvee_{i=1}^n Y_i = \bigvee_{i=1}^n (X_i + Y_i). \tag{2.5}$$

空间 $\mathcal{B}_{\mathbf{F}}^m(n)$ 可以使用由 $\mathbb{R}^m$ 诱导的范数进行赋范，如下：

$$\|X_i\|_1 = \|x_1^{\xi_{i1}} \wedge \cdots \wedge x_m^{\xi_{im}}\|_1 = \sum_{j=1}^m |\xi_{ij}|, \quad \left\|\bigvee_{i=1}^n X_i\right\|_1 = \sum_{i=1}^n \|X_i\|_1. \tag{2.6}$$

显然,

$$\|Z_1 + Z_2\|_1 = \sum_{i=1}^n \sum_{j=1}^m |\xi_{ij} + \eta_{ij}| \le \|Z_1\|_1 + \|Z_2\|_1$$

是一个度量。

为概率析取范式引入代数结构，其目的是对来自多个观测序列的证据聚合进行建模（见附录A）。关于此类求和的进一步说明将在第5节中给出。关于概率析取范式内部结构的更多想法包含在附录 B中。

2.1. 从离散结构过渡到连续函数。记作 $\Delta_{ij} = [j-1+(i-1)m,\ \ j+(i-1)m]$ 并定义有限黎曼可积函数 $X(t,\ \xi;\cdot) : \Delta_{tj} \to \mathbb{R}$，其支撑集为 Δ_{tj}, $t \ge 1$, $\xi = (\xi_1 \ldots,\ \xi_m) \in \mathbb{R}^m$，使得

$$\mathbb{P}\{x_j^{\xi_j} \sim x_j\} = \begin{cases} \int_{\Delta_{tj}} X(t,\xi_j;x)dx, & 1 \ge \int_{\Delta_{tj}} X(t,\xi_j;x)dx > 0, \\ 1, & \int_{\Delta_{tj}} X(t,\xi_j;x)dx > 1, \end{cases} \tag{2.7}$$

$$\mathbb{P}\{x_j^{\xi_j} \sim \overline{x}_j\} = \begin{cases} \int_{\Delta_{tj}} X(t,\xi_j;x)dx, & -1 \le \int_{\Delta_{tj}} X(t,\xi_j;x)dx < 0, \\ 1, & \int_{\Delta_{tj}} X(t,\xi_j;x)dx < -1, \end{cases} \tag{2.8}$$

$$\mathbb{P}\{x_j^{\xi_j} \sim \varepsilon\} = \begin{cases} 1 - \mathbb{P}\{x_j^{\xi_j} \sim x_j\}, & 1 \geq \int_{\Delta_{tj}} X(t, \xi_j; x)dx > 0, \\ 1 - \mathbb{P}\{x_j^{\xi_j} \sim \overline{x}_j\}, & -1 \leq \int_{\Delta_{tj}} X(t, \xi_j; x)dx < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2.9}$$

这里，$\xi_j \text{in } X(t, \xi_j; \cdot)$ 是一个参数，用于定义函数 $X(t, \xi_j; \cdot)$ 的形状。例如，如果函数 $X(t, \xi_j; \cdot) \equiv \alpha_j$ 是一个常数，$\xi_j = \alpha_j$。

定义概率初等合取的编码器如下：

$$Y(t, \xi; x) = \sum_{j=1}^{m} X(t, \xi_j; x).$$

设 $\xi : [0, n] \to \mathbb{R}^m$ 是具有某些性质的函数（例如连续函数、随机游走等）。一个概率析取范式对应于以下求和

$$Z(\xi; x) = \sum_{t=1}^{n} Y(t, \xi(t); x).$$

例如，若函数 $X(t, \xi_j; \cdot)$ 是常数，则相应的编码器

$$Z(\xi; x) = \sum_{t=1}^{n} \sum_{j=1}^{m} \xi_{tj} \mathbf{1}_{\Delta_{tj}}, \quad \xi_j(t) = \xi_{tj},$$

是一个分段线性函数，而 ξ_j 是阶跃 j 的高度（正值或负值）。

显然，用这种记法可以轻松地将经典析取范式编码为分段线性函数，其中 ξ_j 的可能取值属于集合 $\{-1, 0, 1\}$。函数$Z(\xi; \cdot)$ 可以看作一种模拟信号，它以由 ξ 确定的可靠度来编码相应的逻辑陈述。因此，由 (2.5)、(2.4) 定义的概率析取范式的求和与乘法，可以分别从信号叠加和放大的角度来理解。

因此，任意具有有限个跳跃的分段连续函数 $Z(\xi; \cdot) : [0, nm] \to \mathbb{R}$ 都编码了一个来自 $\mathcal{B}_{\mathbf{F}}^m(n)$ 的概率析取范式。我们将这类编码器的集合记为 $\mathcal{E}_{\mathbf{F}}^m(n)$。对于 $\mathcal{E}_{\mathbf{F}}^m(n)$，范数（2.6）对应于 L_1-范数

$$\|Z(\xi; x)\|_1 = \int_0^{nm} |Z(\xi; x)|dx.$$

我们还将假设集合 $\mathcal{E}_{\mathbf{F}}^m(n)$ 在 $\|\cdot\|_1$ 意义下有界，必要时甚至在 $\|\cdot\|_\infty$上确界范数意义下有界。

注 2.5. 上述编码方式的选择仅仅是为了简便。若要采用连续时间，我们应考虑编码器 $Z(\xi; \cdot, \cdot) : [0, n] \times [0, m] \to \mathbb{R}$，其中集合 $[t-1, t] \times [j-1, j]$ 对应于第 $t-$ 个概率初等合取中的第 $j-$ 个变量。在这种情况下，我们可以将概率析取范式建模为基本合取项的连续统，并将其解释为智能体知识状态的连续演化。

这类编码的概率（2.7）–（2.9）可以通过带符号积分表示

$$\iint_{[t-1,t]\times[i-1,i]} Z(\xi; \tau, x)d\tau dx, \quad t = \overline{1, n}, j = \overline{1, m}.$$

我们将这种类型的编码器集合记为 $\mathcal{E}_{\mathbf{F}}(n \times m)$。

注 2.6。 我们也可以使用编码器来编码概率 x_{ij} =真值、x_{ij} =假值，或 x_{ij} 未定义的情况。要处理三值逻辑以外的情况，我们应使用例如 $Z(\xi;\cdot,\ \cdot)$: $[0,n]\times[0,m]\to\mathbb{C}$ 并将积分的实部和虚部分别解释为逻辑值的概率。

注 2.7. 我们可以将编码器 $Z(\xi;\cdot,\ \cdot)\in\mathcal{E}_{\mathbf{F}}(n\times m)$视为一种特殊的概率元胞自动机，其元胞为 $\Delta_{tj},\ j=1,\ m$，时间为 t，可能的元胞状态为$\{-1,0,1\}$。函数 ξ: $[0,n]\to\mathbb{R}^m$ 决定了状态之间转移的概率。

3. Structural similarity between PDNFs, classical DNFs, venjunction, and asynchronous sequential logic

回顾异步时序逻辑中的逻辑—动态运算 venjunction [Vas11, p. 6] 和 sequencion [Vas11, P. 45], 的概念。Venjunction 是一种非对称逻辑—动态运算，记作 $\angle$，使得 $x\angle y$ 当且仅当 $x\wedge y=1$，并且在 x 变为 1 的时刻， y 已经为 1 时取值为 1。Sequention 被定义为一个二元函数，其参数表示为二进制元素变量序列，如下列依赖关系所示:

$$\langle x\rangle = G(\langle x_1,x_2,\dots,x_n\rangle).$$

函数的所有可能状态为 $\langle x\rangle=1$ 和 $\langle x\rangle=0$。Sequention 所取的值取决于变量的值以及这些值出现的顺序。

从这个意义上说，我们的概率析取范式可以赋予类似于基本合取项序列的语义。更准确地说，一个概率析取范式 $Z=X_1\vee\dots\vee X_n$，其中 $X_i,\ i=1,\ n$是变量 m $x_1^{\xi_{i1}}$、.....、$x_m^{\xi_{im}}$的概率初等合取项，可以看作 venjunction $Z=X_n\angle\dots\angle X_1$的类比形式。请注意，venjunctive 链中合取项的顺序是相反的，这反映了时间上的先后关系，因为 X_1 是第一次观测，而 X_n 是最后一次观测。然而，我们使用的是加权文字 $x_j^{\xi_{ij}}$，而不是布尔变量x_j; 每个概率初等合取项都可以实现为 3^m 个可能的基本合取项。因此，整个概率析取范式可以实现为 3^{mn} 个 venjunction。

形式上，令 $\mathcal{V}(m,n)$ 表示通过将 PDNF 中每个带权文字 $x_j^{\xi_{ij}}$ 替换为三个确定性文字 x_j、 x_j 或 ε 之一所能得到的所有可能 venjunction 的有限集 3^{mn}。具有权重矩阵 $\Xi=(\xi_{ij})\in\mathbb{R}^{n\times m}$ 和映射 $\mathbf{F}=(F_{-1},F_0,F_1)$ 的 PDNF $Z(\Xi)\in\mathcal{B}^m_{\mathbf{F}}(n)$ 在 $\mathcal{V}(m,n)$ 上诱导一个离散概率测度 μ_Ξ。

该测度构造如下。对于 PDNF $Z(\Xi)$ 的每个位置（$i,\ j$）（对应于第 i 个合取中的第 j 个变量），函数 F^j_{-1}， F^j_0， F^j_1 为 $x_j^{\xi_{ij}}$ 的三种可能实现分配概率。假设不同（$i,\ j$）对的赋值彼此独立——当传感器误差或随机影响在不同时间和传感器之间相互独立时，这是一种自然假设——某个特定 venjunction 的概率

$$\langle v\rangle = v_{n1}\wedge\dots\wedge v_{nm}\angle\dots\angle v_{11}\wedge\dots\wedge v_{1m}$$

可表示为以下乘积

$$\mu_\Xi(\{\langle v\rangle\}) = \prod_{i=1}^{n}\prod_{j=1}^{m}\begin{cases} F_{-1}^{j}(\xi_{ij}) & \text{if } v_{ij} = \overline{x}_j, \\ F_{0}^{j}(\xi_{ij}) & \text{if } v_{ij} = \varepsilon, \\ F_{1}^{j}(\xi_{ij}) & \text{if } v_{ij} = x_j, \end{cases}$$

其中v_{ij} 表示出现在 $\langle v\rangle$中第(i,j)个位置的文字， i 则按相反顺序排列。因此，三元组$(\mathcal{V}(m,$ $n)$ $2^{\mathcal{V}(m,n)}, \mu_\Xi)$构成一个有限概率空间。这一事实的简单证明见附录 C。

我们估计按照μ_Ξ随机生成的 venjunction $\langle v\rangle$ 在适当度量下接近同一长度的固定目标合取项 $\langle v'\rangle$ 的概率，并在附录 E 中给出此类估计的一个例子。

3.1. 编码器空间上的概率测度。 在第 6 节和附录 F 所述的传感器网络应用中，权重矩阵 Ξ 本身可能是随机的，取决于未观测到的物理量。在这种情况下，可以得到一个两阶段随机过程：首先生成随机矩阵 Ξ ，然后在给定 Ξ 的条件下，按照 μ_Ξ 抽取一个 venjunction。就来自 $\mathcal{E}_{\mathbf{F}}^{m}$ (n) 的编码器而言，我们从集合 $\mathcal{E}_{\mathbf{F}}^{m}(n)$ 中抽取一个随机元素（即一个函数），然后用所抽取的函数对 venjunction 进行编码。

Th在 $\mathcal{V}(m,\ n)$ 上的整体分布由勒贝格–斯蒂尔杰斯积分给出 al

$$\mathbb{P}(A) = \int_{\mathbb{R}^{n\times m}} \mu_\Xi(A)\, dP(\Xi), \qquad A \subseteq \mathcal{V}(m,n),$$

其中 P 表示 Ξ的分布。这种分层构造使 PDNF形式体系能够同时表达底层参数的不确定性（由 Ξ编码）以及逻辑观测的随机实现（由**F**编码）。

第 2.1 小节的构造为每个权重矩阵 $\Xi \in \mathbb{R}^{n\times m}$ 关联一个函数$Z_\Xi \in \mathcal{E}_{\mathbf{F}}^{m}(n)$，例如分段常值编码器

$$Z_\Xi(x) = Z(\xi; x) = \sum_{i=1}^{n}\sum_{j=1}^{m} \xi_{ij}\mathbf{1}_{\Delta_{ij}}(x),$$

其中 $\xi\colon\ \{1,\dots,n\} \to \mathbb{R}^m$， $\xi_j(i) = \xi_{ij}$。如果 Ξ 是服从 $\mathbb{R}^{n\times m}$上分布 P的随机变量，那么像测度 $Q = \Phi_* P$ 就在 $\mathcal{E}_{\mathbf{F}}^{m}(n)$ 上定义了一个概率分布。这里， $\Phi_* P$ 是 P 在映射 $\Phi\colon\ \Xi \mapsto Z_\Xi$，$(\Phi_* P)(A) = P($ $\Phi^{-1}(A))$ 下的推前。

对于固定的编码器 $Z \in \mathcal{E}_{\mathbf{F}}^{m}(n)$，venjunction 的概率由 (2.7)–(2.9) 中的积分得到。令 μ_Z 表示相应的概率测度，其定义在 $\mathcal{V}(m,n)$ 上。映射 $Z \mapsto \mu_Z$ 可视为从 $\mathcal{E}_{\mathbf{F}}^{m}(n)$ 到巴拿赫空间 $\mathcal{M}(\mathcal{V}(m,n))$ 的函数；该空间由定义在 $\mathcal{V}(m,n)$ 上的有限符号测度构成，并配备全变差范数。在较弱的可测性条件下（例如，若对每个 $A \subseteq \mathcal{V}(m,n)$， $Z \mapsto \mu_Z(A)$ 都是可测的），博赫纳积分 [DU77]

$$\mathbb{P}(A) = \int_{\mathcal{E}_{\mathbf{F}}^{m}(n)} \mu_Z(A)\, dQ(Z)$$

存在，并定义了 $\mathcal{V}(m,n)$ 上的一个概率测度。回顾一下，博赫纳积分将勒贝格积分推广到了取值于巴拿赫空间的函数。

因此，两阶段随机过程可以重新表述为：首先根据 Q 抽取一个随机编码器 Z，然后根据 μ_Z 抽取一个 venjunction。这一视角统一了

PDNF 形式体系的离散与连续方面，并为应用泛函分析工具铺平道路，例如博赫纳积分、函数空间中的随机过程以及遍历理论。我们将在第 6 节继续讨论这一问题。

注3.1. 为量化概率析取范式中所编码的不确定性，我们可以引入基于熵的度量。

对于固定的编码器 Z_Ξ（因而概率 $p_{tj}^{(\ell)} = F_\ell^j(\xi_{tj})$也固定），每个位置 (t, j) 的香农熵为 $H_{tj} = -\sum_{\ell\in\{-1,0,1\}} p_{tj}^{(\ell)} \log\, p_{tj}^{(\ell)}$。对所有位置求和，即得到总熵

$$H_{\mathrm{tot}}(\Xi) = \sum_{t=1}^{n}\sum_{j=1}^{m} H_{tj},$$

对于经典析取范式（所有概率均为0或1），该值消失；当三种结果等概率时，该值达到最大。

4. PDNFs as Generators of Finite Languages

固定权重矩阵为 Ξ的概率析取范式 Z 可视为字母表 $\Sigma = \{x_j, \varepsilon, \overline{x}_j\}_{j=1}^m$上有限语言的概率生成器。每个实现 v 是该字母表上长度为$\overline{n}$的词（其中每个位置对应一个时间步），所有可能词的集合为

$$L(Z) = \{\langle v\rangle \in \mathcal{V}(m, n) \mid \mu_\Xi(\{\langle v\rangle\}) > 0\}.$$

因此， $L(Z)$ 是概率测度 μ_Ξ 的支撑集，其大小 $|L(Z)|$ 可从 1（确定性情形）到 3^{nm} 不等。

由于不同位置上的选择彼此独立， $L(Z)$ 是局部语言的笛卡尔积：

$$L(Z) = \prod_{i=1}^{n}\prod_{j=1}^{m} L_{ij},$$

其中，每个 L_{ij} 都是可能出现在位置 (i, j) 的文字集合，即

$$L_{ij} = \{\ell \in \{\overline{x}_j, \varepsilon, x_j\} \mid F_\ell^j(\xi_{ij}) > 0\}.$$

因此， $L(Z)$ 是一种有限正则语言，其正则表达式就是将每个位置上允许的备选项依次连接起来，即对 i, j 进行连接。

如果我们将概率析取范式视为生成合理假设的模板，那么 $L(Z)$ 正是所考虑的假设集合。目标陈述 v' 属于 $L(Z)$，当且仅当对于每个 (i, j)，文字 v'_{ij} 都受到相应权重的允许，即 $F_{v'_{ij}}^j(\xi_{ij}) > 0$。因此，权重的选择定义了一个针对所有可能 venjunction 的筛选器。

现在假设我们不知道权重，但可以从 μ_Ξ 中观测到独立抽样。如第5节所述，经过足够多次观测后，我们就能以很高的置信度识别整个集合 $L(Z)$（见命题5.3）。

一旦已知 $L(Z)$，我们就能明确列举系统所有可能的行为。如果还想知道其中哪些行为确实有用或为真，则需要进一步测试或进行逻辑推理，因为概率析取范式本身只提供候选集合。

该 语言 $L(Z)$可以用如下形式的概率正则表达式来描述

$$\bigwedge_{i=1}^{n}\bigwedge_{j=1}^{m}(\overline{x}_j \mid \varepsilon \mid x_j) \quad \text{but with some alternatives removed,}$$

其中，移除由条件 $F_\ell^j(\xi_{ij}) > 0$ 决定。

例4.1. 一个苏联故事。

让我们说明概率析取范式如何作为模板，生成具有指定概率的假设。考虑一个简单的因果叙事：“弗拉德写了一封告密信，于是我被解雇了”。在完全确定性的情境下，这可以表示为一个 venjunction

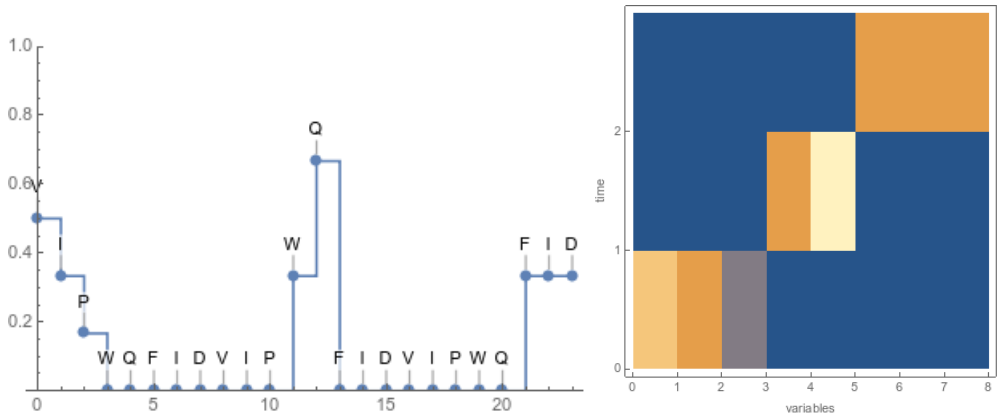
$$(F \wedge \neg J \wedge \neg D) \angle (W \wedge \neg O) \angle (V \wedge \neg I \wedge \neg P),$$

其中，变量含义如下：
▷ V：“弗拉德写了一封告密信”， I：“Ivan写了一封告密信”， P：“彼得写了一封告密信”；
▷ W：“书面告密信”， O：“口头报告”；▷ F：“我被解雇了”， J：“我被监禁了”， D：“我被降职了”。

$\angle$ 运算符表示事件按此顺序发生。然而，现实中我们可能对许多细节并不确定。假设我们知道（根据经验或先前数据）以下概率：

▷ 作者是弗拉德的概率为 $1/2$，是Ivan的概率为 $1/3$，是彼得的概率为 $1/6$；▷ 沟通方式为书面告密信的概率为 $1/3$，为口头报告的概率为$2/3$；▷ 后果是被解雇、被监禁或被降职，且每种情况的概率均为 $1/3$。

概率析取范式可以通过为每个合取中的每个变量分配权重来编码这些可能性。相应的分段线性编码器 $Z_S \in \mathcal{E}_{\mathbf{F}}^8(3)$和$U_S \in \mathcal{E}(3 \times 8)$的示例可见于图1。



(A) (B)

FIGURE 1. “苏联故事”概率析取范式的分段线性编码器 $Z_S \in \mathcal{E}_{\mathbf{F}}^8(3)$ (A) 和 $U_S \in \mathcal{E}_{\mathbf{F}}(3 \times 8)$ (B)

因此，概率析取范式提供了一整组相关叙事的紧凑表示，使我们能够探索例如序列“Ivan写了一份口头报告，而我被监禁”的可能性，或检验哪些变体与额外证据一致。

5. Probabilistic Disjunctive Normal Forms as Automata Arrays

L让我们回顾一下确定性有限自动机的定义。它是一个六元组

$$(S, \mathcal{X}, \mathcal{Y}, s^0, \delta, \lambda),$$

其中， S 是内部状态集合， $\mathcal{X}$ 是输入字母表， $\mathcal{Y}$ 是输出字母表， $s^0 \in S$ 是初始状态， δ : $S \times \mathcal{X} \to S$ 是状态转移函数， λ 是输出函数。如果输出函数的作用形式为 λ : $S \times \mathcal{X} \to \mathcal{Y}$，则该自动机称为米利机；如果输出函数的作用形式为 $\lambda : S \to \mathcal{Y}$，则该自动机称为穆尔机。

设有自动机 $A_j = (S_j, \mathcal{X}_j, \mathcal{Y}_j, s_j^0, \delta_j, \lambda_j)$， $j = \overline{1, m}$，它们在时刻 t_i, $i = \overline{1, n}$产生输出。输出字母表 $\mathcal{Y}$ 由三个符号“x”、“$\overline{x}$”和 ε （空字符串）组成。输入字母表 $\mathcal{X}$ 是 $\mathbb{R}^p$的子集。假设$\mathcal{X}_j = \cup_{i=1}^{q_j} \mathcal{X}_j^i$且 $\mathcal{X}_j^i \cap \mathcal{X}_j^k = \emptyset$, $i, k = \overline{1, q_j}$, $S_j = \{s_j^0, s_j^1, \ldots, s_j^{q_j}\}$。如果自动机 A_j接收到 $x \in \mathcal{X}_j^i$，那么无论先前处于什么状态， A_j都会转移到状态 s_j^i。状态集合还被划分为三个互不相交的子集： $S_j = S_j^{+1} \cup S_j^{-1} \cup S_j^0$,$\lambda_j(s) = x$, $s \in S_j^{+1}$, $\lambda_j(s) = \overline{x}$, $s \in S_j^{-1}$, $\lambda_j(s) = \varepsilon$, $s \in S_j^0$。令 $\mathcal{X}_j^{+1}$表示所有满足 $s_j^i \in S_j^{+1}$的$\mathcal{X}_j^i$的并集， $\mathcal{X}_j^{-1}$表示所有满足 $s_j^i \in S_j^{-1}$的 $\mathcal{X}_j^i$的并集， $\mathcal{X}_j^0$表示所有满足$s_j^i \in S_j^0$的 $\mathcal{X}_j^i$的并集。

假设所有输入都是等概率的。自动机 A_j输出“x”的概率为 $P_j^{+1} = \frac{|\mathcal{X}_j^{+1}|}{|\mathcal{X}_j|}$，输出“$\overline{x}$”的概率为 $P_j^{-1} = \frac{|\mathcal{X}_j^{-1}|}{|\mathcal{X}_j|}$，输出$\varepsilon$的概率为 $P_j^0 = \frac{|\mathcal{X}_j^0|}{|\mathcal{X}_j|}$。如果每个输入恰好对应一个状态，那么这些概率分别为 $\tilde{P}_j^{+1} = \frac{|S_j^{+1}|}{|S_j|}$ 、$\tilde{P}_j^{-1} = \frac{|S_j^{-1}|}{|S_j|}$、 $\tilde{P}_j^0 = \frac{|S_j^0|}{|S_j|}$。

如果输入未知，我们可以根据上述公式估算输出的概率，并将自动机的输出符号化地记作 x^ξ，其中 ξ 是满足某个函数 $\mathbf{F}$ 以及相应概率 P_j^{+1}、 P_j^0、P_j^{-1}下的式 (2.1)–(2.3) 的数。如果在某个时刻 $t \geq 1$有一个包含 m 个自动机（例如传感器）的系统，其输入未知，我们可以将其符号化地记作概率初等合取$x_1^{\xi_1} \wedge \cdots \wedge x_m^{\xi_m}$，其中 $\xi_1, \ldots, \xi_m$ 按上述方法计算。

现在令自动机 A_j 的新状态 $s_j(t)$ 取决于输入 $x(t) \in \mathcal{X}_j^i$ 以及先前的状态 $s_j(t-1)$。在这种情况下，我们应在任意时刻改变划分 $\mathcal{X}_j = \cup_{i=1}^{q_j} \mathcal{X}_j^i(t)$ 和 $S_j(t) = S_j^{+1}(t) \cup S_j^{-1}(t) \cup S_j^0(t)$，而概率 $P_j^{+1}, P_j^0, P_j^{-1}$ 将成为时间 t 的函数，因此，数值 $\xi_1(t) \ldots, \xi_m(t)$ 也将成为 t 的函数。我们可以将自动机在时刻 $t = \overline{1, n}$ 运行的情形，符号化地表示为来自 $\mathcal{B}_F^m(n)$ 的概率析取范式

$$x_1^{\xi_1(1)} \wedge \cdots \wedge x_m^{\xi_m(1)} \vee \cdots \vee x_1^{\xi_1(n)} \wedge \cdots \wedge x_m^{\xi_m(n)}.$$

如果不同的智能体观测传感器，且每个智能体在多个运行周期内对其进行观测，那么可以根据公式 (2.5) 聚合与各智能体观测结果相对应的概率析取范式。这种聚合能够对被观测的自动机进行综合估计，并据此提出关于其未知输入的其他假设

。这个问题也与著名的自动机辨识问题有关（例如，参见 [Wha74]）。

定义 5.1（三值结果的贝叶斯融合）。 设 $p = (p_{-1}, p_0, p_1)$ 和 $q = (q_{-1}, q_0, q_1)$ 是定义在三值集合 $\{x, \varepsilon, x\}$ 上的两个概率向量，分别表示相互独立的证据。它们的贝叶斯融合（假设先验均匀）是由下式定义的概率向量 $p * q$

$$(p * q)_\ell = \frac{p_\ell \cdot q_\ell}{\sum_{\ell' \in \{-1,0,1\}} p_{\ell'} \cdot q_{\ell'}}, \quad \ell \in \{-1, 0, 1\}.$$

定理 5.2（独立观测的组合）。 设 A 是一个确定性有限自动机，具有 m 个输出变量 x_1、…、x_m，由两个相互独立的代理 ag_1 和 ag_2 在 n 个同步时间步内进行观测。假设每个代理都使用相同的概率映射族 $\mathbf{F}$，将其观测记录为 *PDNF* $Z_k \in \mathcal{B}^m_{\mathbf{F}}(n)$ *(*$k = 1,2$*)*。此外，假设

$$F^j_\ell(\xi) = \frac{\exp(\alpha^j_\ell \xi)}{\sum_{\ell' \in \{-1,0,1\}} \exp(\alpha^j_{\ell'} \xi)}.$$

那么，*PDNF* $Z = Z_1 + Z_2$（其加法按 *(2.5)* 定义）表示来自两个代理的合并证据，即对于每个 $i = 1$、…、n 和 $j = 1$、…、m，Z 中的权重 ξ_{ij} 满足

$$F^j_\ell(\xi_{ij}) \propto F^j_\ell(\xi^{(1)}_{ij}) \cdot F^j_\ell(\xi^{(2)}_{ij}), \quad \ell \in \{-1, 0, 1\},$$

其中 $\xi^{(k)}_{ij}$ 是 Z_k中的权重。

此外，如果给定 A的真实状态后，各智能体的观测条件独立，那么概率向量

$$\big(F^j_{-1}(\xi_{ij}), F^j_0(\xi_{ij}), F^j_1(\xi_{ij})\big)$$

恰好是来自 Z_1 和 Z_2的相应向量的贝叶斯融合。

证明。 令我们固定一个时间步 i 和一个变量 x_j。对于智能体 k（$k = 1$，2），权重 $\xi^{(k)}_{ij}$ 通过映射 $\mathbf{F}$，编码了基于该智能体部分信息，对 x_j 在时间i 的三种可能实现的概率。根据定义 1，这些概率为

$$p^{(k)}_\ell = F^j_\ell(\xi^{(k)}_{ij}), \quad \ell \in \{-1, 0, 1\},$$

其中 $p^{(k)}_{-1} + p^{(k)}_0 + p^{(k)}_1 = 1$。

假设智能体的观测相互独立（给定自动机的真实状态），那么两个智能体对特定实现 ℓ 赋予的联合概率与乘积 $p^{(1)}_\ell \cdot p^{(2)}_\ell$ 成正比。归一化后得到

$$p^{\text{joint}}_\ell = \frac{p^{(1)}_\ell p^{(2)}_\ell}{\sum_{\ell' \in \{-1,0,1\}} p^{(1)}_{\ell'} p^{(2)}_{\ell'}}.$$

现在考虑求和 $Z = Z_1 + Z_2$ 中出现的权重 $\xi_{ij} = \xi_{ij}^{(1)} + \xi_{ij}^{(2)}$。根据 $\mathcal{B}_{\mathbf{F}}^m(n)$ 中加法的定义，我们有

$$F_\ell^j(\xi_{ij}) = F_\ell^j(\xi_{ij}^{(1)} + \xi_{ij}^{(2)}) = \frac{\exp(\alpha_\ell^j(\xi_{ij}^{(1)} + \xi_{ij}^{(2)}))}{\sum_{\ell' \in \{-1,0,1\}} \exp(\alpha_{\ell'}^j(\xi_{ij}^{(1)} + \xi_{ij}^{(2)}))} =$$

$$= \frac{\exp(\alpha_\ell^j \xi^{(1)}) \exp(\alpha_\ell^j \xi^{(2)})}{\sum_{\ell' \in \{-1,0,1\}} \exp(\alpha_{\ell'}^j \xi_{ij}^{(1)}) \exp(\alpha_{\ell'}^j \xi_{ij}^{(2)})} = \frac{p_\ell^{(1)} p_\ell^{(2)}}{\sum_{\ell' \in \{-1,0,1\}} p_{\ell'}^{(1)} p_{\ell'}^{(2)}} = p_\ell^{\text{joint}}.$$

因此，函数 $F_\ell^j(\xi_{ij})$ 与 $F_\ell^j(\xi_{ij}^{(1)})$ 和 $F_\ell^j(\xi_{ij}^{(2)})$ 的贝叶斯融合一致。 □

在此语境下，指数参数化 $F_\ell^j(\xi) \propto \exp(\alpha_\ell^j \xi)$ 很自然，因为它使空间 $\mathcal{B}_{\mathbf{F}}^m(n)$ 成为线性指数族。在这种情况下，映射 $\xi \mapsto (F_{-1}^j(\xi)\ F_0^j(\xi)\ F_1^j(\xi))$ 就是 Softmax变换，广泛应用于统计建模和机器学习。权重的加法组合对应于独立概率相乘，这是合并独立证据的标准规则。

前述定理表明，如果概率映射采用指数（softmax）形式，那么权重相加恰好对应于贝叶斯融合。从本质上说，这是一个众所周知的结果，只不过以本文的框架重新表述。关于这一问题的进一步讨论，请参见附录 D。

下面描述概率析取范式所对应的确定性结构。考虑一个具有固定但未知权重矩阵 Ξ 的概率析取范式。实际可能出现（即具有正概率）的 venjunction 构成某个子集 $\mathcal{V}_+ \subseteq \mathcal{V}(m, n)$；设 $M = |\mathcal{V}_+| \leq 3^{nm}$。假设基于物理或设计方面的考虑，我们知道所有可能出现的 venjunction 的概率存在一个统一下界：

$$\mu_\Xi(\{\langle v \rangle\}) \geq p_{\min} > 0 \qquad \text{for every } \langle v \rangle \in \mathcal{V}_+.$$

我们根据 μ_Ξ 独立抽取 venjunction。集邮者问题中的一个经典结果告诉我们：要以至少 $1 - \delta$ 的概率，至少一次看到一个大小为 M 的集合中的每个元素，只需进行

$$N \geq \frac{1}{p_{\min}} \ln \frac{M}{\delta}$$

次观测。因此，我们可以表述如下结论

命题5.3。 看到最罕见的可能 *venjunction* 的概率至少为 $p_{\min}$。我们需要

$$N \geq \frac{1}{p_{\min}} \ln \frac{3^{nm}}{\delta}$$

次观测 才能以至少 $1 -$ 的概率使每个可能的 *venjunction* 至少出现一次 δ.

证明。固定的 venjunction $v \in \mathcal{V}_+$ 在 N 次试验中未被观测到的概率至多为 $(1 - p_{\min})^N \leq e^{-p_{\min} N}$。根据并集界， M个 venjunction 中至少有一个仍未被观测到的概率不超过 $M e^{-p_{\min} N} \leq 3^{nm} e^{-p_{\min} N}$。令该界等于 δ，即可得到所需的 N。 □

因此，如果事先已知正下界 $p_{\min}$，我们就能保证在进行 N 次观测后（按上述方式选择 N），以 $1 - \delta$ 的置信度观测到每一种可能的 venjunction。当概率析取范式模型用于描述某个

物理系统时，这一点很有用，因为已知该系统的每种可能状态都具有某个特定的最小概率。

假设经过足够长的一系列独立观测后，我们已经收集到了所有实际可能出现的 venjunction 的完整集合 $\mathcal{V}_+ = \{\langle v_1\rangle, \dots, \langle v_M\rangle\}$。每个 $\langle v_r\rangle$ 都是一个特定的确定性 venjunction，即由文字组成的序列（每个文字可以是x_j、 x_j或 ε）。现在，我们希望将该集合表示为一个同时涉及一组隐变量 $h_1, \dots, h_l$（对所有时间步都相同）的单一确定性 venjunction。这些隐变量是抽象占位符；在每个合取中，它们可以以自身形式出现、以否定形式出现，或缺省，恰如普通变量 x_j一样。

对集合 $\mathcal{V}_+$ 进行编码的一种自然方式，是为每个观测到的 venjunction 引入一个隐变量组合。如果隐变量可以取自一个有限集（例如，每个 h_s 都可以出现、被否定或缺省），那么我们可以写成

$$\mathcal{V}_+ = \bigvee_{r=1}^{M} \left(\bigwedge_{j=1}^{m} x_j^{a_{rj}} \ \wedge \ \bigwedge_{s=1}^{l} h_s^{b_{rs}} \right),$$

其中，符号 $\bigvee$ 表示 venjunction，指数表示相应变量在第 r个合取中的形式（出现、否定或缺失）。系数a_{rj} 由观测到的 venjunction $\langle v_r\rangle$唯一确定，而系数 b_{rs} 可以任意选取，只要不同的 r 对应不同的系数（这样每个合取就对应唯一的隐变量组合）。例如，若令 $l = \lceil \log_3 M \rceil$，并将 b_{rs} 的每种组合解释为索引 r的三进制表示，就能得到一种有效编码。

因此，这种表示始终存在，但绝非唯一。若没有关于隐变量物理含义的额外信息，那么 $\mathcal{V}_+$ 与隐变量配置集合之间的任何双射都能产生一种确定性描述。在实践中，如果已知隐变量代表特定的未观测因素，那么这些因素对观测的影响可能会施加约束，从而选出一种特定表示。否则，构造出的确定性 venjunction 仅仅证明了概率析取范式可以通过引入额外变量扩展为确定性公式。

因此，权重 ξ_{ij} 编码了隐状态的概率，而代数运算使我们能够结合来自多个独立观测的证据，逐步降低不确定性。在极限情况下，当所有可能的 venjunction 都已被观测到时，概率析取范式可以替换为等价的确定性描述，不过这可能需要引入辅助变量。

6. Connection of PDNFs and Markov Processes

权重 ξ_{ij} 可以不是固定不变的，而是根据某个随机过程随时间演变。一种自然且重要的情况是，每个权重都遵循随机游走。在每个时间 $t > 1$，权重 $\xi_{tj} = \eta + \xi_{t-1,\ j}$ 是随机游走的一次实现，而输出概率$F_\ell^j(\xi_{tj})$ 本身也变成随机变量。

这种构造产生了一个两层随机过程（见第3.1小节）。底层随机游走 $\{\xi_j(t)\}_{t\geq 1}$ 描述了传感器 j 测量的物理量的演变。

因此，得到的 venjunction 序列 $\{\langle v(t)\rangle\}_{t\geq 1}$ 是一个隐马尔可夫模型，其中随机游走是隐藏状态，而概率析取范式则规定了发射概率。

利用第 2.1 小节中引入的概率析取范式的连续表示，我们可以考虑映射 Π_Z : $\mathcal{RW} \to \mathcal{E}^m_{\mathbf{F}}(n)$，其中 $\mathcal{RW}$ 是所考虑的随机游走集合，$\Pi_Z(\xi) \mapsto Z(\xi;\,\cdot) \in \mathcal{E}^m_{\mathbf{F}}(n)$，$\xi \in \mathcal{RW}$。$\Pi_Z(\xi)$ 是时间 t 上的编码器值随机过程，其取值形式为 $Z(\xi(t); x)$。Π_Z 的均值函数由博赫纳积分给出

$$\mathbb{E}[\Pi_Z(\xi)](x) = \mathbb{E}[Z(\xi;x)] = \int_{\xi\in\mathcal{RW}} Z(\xi;x)\, d\mathbb{P}_\xi,$$

其中 $\mathbb{P}_\xi$ 是 ξ的分布。对于概率析取范式的分段常数表示 ，

$$(\Pi_Z(\xi))(x) = \sum_{t=1}^{n}\sum_{j=1}^{m} \xi_{tj}\mathbf{1}_{\Delta_{tj}}(x),$$

这可简化为

$$\mathbb{E}[\Pi_Z(\xi)](x) = \sum_{t=1}^{n}\sum_{j=1}^{m} \mathbb{E}[\xi_{tj}]\,\mathbf{1}_{\Delta_{tj}}(x) = \sum_{t=1}^{n}\sum_{j=1}^{m} (t\mathbb{E}\eta + \mathbb{E}\xi_{0,j})\mathbf{1}_{\Delta_{tj}}(x).$$

因此，均值函数再次呈现分段常数形式，其阶梯高度由时间 t 上权重的期望值给出。如果 $\xi_{0,\,j}$ 是确定性的，$\mathbb{E}[\xi_{tj}] = \xi_{0,\,j} + t\mu$ 其中 $\mu = \mathbb{E}[\eta]$。

编码器 $\Pi_Z(\xi)$ 在点 x 处的方差定义为

$$\mathrm{Var}(\Pi_Z(\xi))(x) = \mathbb{E}\big[(\Pi_Z(\xi)(x) - \mathbb{E}[\Pi_Z(\xi)(x)])^2\big].$$

利用分段常数表示

$$(\Pi_Z(\xi))(x) = \sum_{t=1}^{n}\sum_{j=1}^{m} \xi_{tj}\mathbf{1}_{\Delta_{tj}}(x),$$

并注意到不同的配对 $(t,\ j)$ 对应的示性函数具有不相交支集，我们得到

$$\mathrm{Var}(\Pi_Z(\xi))(x) = \sum_{t=1}^{n}\sum_{j=1}^{m} \mathrm{Var}(\xi_{tj})\,\mathbf{1}_{\Delta_{tj}}(x).$$

若 ξ_{tj} 遵循具有独立同分布增量 η_s 的随机游走 $\xi_{tj} = \xi_{0j} + \sum_{s=1}^{t}\eta_s$，增量方差为 $\sigma^2 = \mathrm{Var}(\eta_s)$，且初始值为 $\xi_{0,\,j}$ （可能是方差为 $\mathrm{Var}(\xi_{0,\,j})$ 的随机变量），则

$$\mathrm{Var}(\xi_{tj}) = \mathrm{Var}(\xi_{0,j}) + t\sigma^2.$$

代入后，我们得到

$$\mathrm{Var}(\Pi_Z(\xi))(x) = \sum_{t=1}^{n}\sum_{j=1}^{m} \big(\mathrm{Var}(\xi_{0,j}) + t\sigma^2\big)\mathbf{1}_{\Delta_{tj}}(x).$$

对于对称随机游走 $(\mu = 0)$，均值保持不变，而方差不断增大，这反映了不确定性随时间增加。对于带漂移随机游走，均值和方差均随时间线性增长。

本理论的中等应用示例将在附录 F 中给出。我们还可以证明，在一般情况下，PDNF 编码器的均值存在。

定理 6.1。 设 $\mathcal{RW}$ 是随机游走的轨迹集合，其测度为 $\mathbb{P}_\xi$，且对于所有 t、j，一阶矩 $\mathbb{E}[|\xi_{tj}|] < \infty$ 均有限。

我们假设 $\mathcal{E}^m_{\mathbf{F}}(n) \subseteq L_1([0, nm])$ 是一个范数为 $\|f\|_1 = \int_0^{nm} |f(x)|\,dx$ 的巴拿赫空间。

此外，假设编码器构造满足如下形式的界

$$\|\Pi_Z(\xi)\|_1 \le C \sum_{t=1}^{n} \sum_{j=1}^{m} |\xi_{tj}| \tag{6.1}$$

其中某个常数为 $C > 0$ （分段线性表示满足这一条件；更一般地，如果对 ξ 的依赖在 L_1 范数下是 *Lipschitz* 连续的，也满足这一条件），且 $\Pi_Z(\xi)$ 在固定的 Z下关于 ξ连续。

博赫纳积分

$$\mathbb{E}[\Phi] = \int_{\mathcal{RW}} \Pi_Z(\xi)\, d\mathbb{P}_\xi$$

存在。

证明。 设 $(\Omega\ \mathcal{F})$ 是一个可测空间， X 是一个巴拿赫空间。若函数 $s\colon \Omega \to X$ 可写成可测集的示性函数的有限线性组合，则称其为简单函数：

$$s(\omega) = \sum_{k=1}^{n} \mathbf{1}_{A_k}(\omega)\, x_k,$$

其中 $A_k \in \mathcal{F}$ 是两两不相交的可测集， $x_k \in X$ 是固定向量。

我们需要证明 Π_Z 的强可测性，即存在一列简单函数 $\Pi_k\colon \mathcal{RW} \to \mathcal{E}^m_{\mathbf{F}}(n)$，使得对于 $\mathbb{P}_\xi$-几乎处处的 ξ，有 $\|\Pi_k(\xi) - \Pi_Z(\xi)\|_1 \to 0$。

此外，还需证明范数的可积性：

$$\int_{\mathcal{RW}} \|\Pi_Z(\xi)\|_1\, d\mathbb{P}_\xi < \infty.$$

强可测性。 映射 $\xi \mapsto \Pi_Z(\xi)$ 是随机游走轨迹 ξ （这是一个可测函数）与第 2.1 小节中编码器 $Z(\xi;\cdot)$ 的构造的复合。该编码器关于 ξ 连续，因此是博雷尔可测的。此外，$\mathcal{E}^m_{\mathbf{F}}(n) \subset L_1$ 是可分的，因此 Π_Z 的像取值于可分子集。由佩蒂斯定理 [DU77, 第 II 章，定理 1]，Π_Z 强可测：它是简单函数列的逐点极限（例如，可在巴拿赫空间中用分段常值函数逼近连续映射）。

范数的可积性。根据假设 (6.1)，我们有

$$\int_{\mathcal{RW}} \|\Pi_Z(\xi)\|_1\, d\mathbb{P}_\xi \le C \sum_{t=1}^{n} \sum_{j=1}^{m} \mathbb{E}[|\xi_{tj}|]. \tag{6.2}$$

如果随机游走的一阶矩有限（例如，增量有界），那么对于所有 t, j， $\mathbb{E}[|\xi_{tj}|] < \infty$，且总和 (6.2) 有限。 □

通过博赫纳积分得到的平均编码器函数 $\mathbb{E}[\Pi_Z(\xi)](x)$，根据其在每个子区间 Δ_{tj} 上的取值，给出传感器j在时间 t 的期望权重 $\mathbb{E}[\xi_{tj}]$ 。这些期望权重刻画了隐藏参数随机演化下系统最稳定的行为

。如果期望权重显著为正，系统倾向于输出x_j；如果显著为负，则倾向于输出 x_j；如果接近零，则输出最有可能是 ε。因此，均值函数可用于决策：例如，当期望权重超过某个阈值、表明关键事件发生概率较高时，可以触发警报或采取预防措施。在传感器网络的背景下，我们可以据此基于传感器的长期平均行为作出决策。从第 4 节的假设生成视角来看，该均值对应于所有可能选项中的一个“中心”陈述（即在平均行为方面最能代表整个可能 venjunction 集合的陈述）。

由于编码器值 $Z(\xi;\cdot)$ 被解释为变量出现（正向）或其否定出现（负向）的倾向，对这些值按照分布 $\mathbb{P}_\xi$取平均可能会导致信息丢失。分别考察正部和负部可以得到更有意义的特征。令 $Z_+(\xi;\cdot) = \max(Z,0)$，且 $Z_-(\xi;\cdot) = \min(Z,0)$，于是 $Z(\xi;\cdot) = Z_+(\xi;\cdot) + Z_-(\xi;\cdot)$。期望 $\mathbb{E}[Z_+(\xi;\cdot)]$ 和 $\mathbb{E}[Z_-(\xi;\cdot)]$（通过对相应函数计算博赫纳积分得到）分别描述变量正向和负向出现倾向的平均值。二者的比值或差值可用作衡量PDNF不对称性的指标。

7. Conclusion

本文将概率析取范式（PDNF）引入为一种结合布尔逻辑与概率建模的形式体系。我们的主要贡献是一种新颖的连续表示：每个PDNF都由分割区间上的分段连续函数编码，其中对子区间的积分决定了PDNF文字的概率。这一视角将PDNF集合转化为巴拿赫空间，使我们能够将随机PDNF视为取值于函数空间的随机过程。我们建立了权重按随机游走演变的PDNF与隐马尔可夫模型之间的联系，并计算出相应编码器过程的均值和方差的显式表达式。

此外，我们还建立了若干次要结果：
▷ 我们证明，每个概率析取范式都会在时序逻辑陈述（venjunction）集合上诱导一个概率测度，并且在温和条件下，随机编码器的期望作为博赫纳积分存在。▷ 我们表明，在指数参数化下，概率析取范式的加法对应于独立证据的贝叶斯融合，并刻画了指数族是唯一具有这一性质的族。▷ 我们得到了识别所有可能 venjunction 所需观测数量的定量界。▷ 我们还讨论了其在自动机辨识以及生成合理的逻辑陈述族方面的应用。

所提出的形式体系连接了离散逻辑、概率论与泛函分析。未来工作可以探索连续时间版本、更一般函数空间的使用，以及在统计学习与验证中的应用。

References

[DL25] Vincent Derkinderen 与 Jean-Marie Lagniez。电路感知的 d-DNNF 编译。载于 James Kwok, edit或, *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence,*

IJCAI-25，第 4454–4462 页。国际人工智能联合会议组织，2025 年 8 月。主会场。doi:10.24963/ijcai.2025/496。[Dro08] Barbara Drossel。 *Random Boolean Networks*，第 3 章，第 69–110 页。John Wiley & Sons, Ltd，2008 年。URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9783527626359.ch3, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527626359.ch3，doi:10. 1002/9783527626359.ch3。[DU77] Joseph Diestel 和 Jerry J. Uhl。 *Vector Measures*，《 *Mathematical Surveys*》第 15 卷。美国数学会，罗德岛州普罗维登斯，1977 年。[JKP+25] Wojciech Jamroga、Marta Kwiatkowska、Wojciech Penczek、Laure Petrucci 和 Teofil Sidoruk。概率定时 ATL。载于 *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems* 25, 1051 1059, 2025 ，AAMAS ' 页 –，南卡罗来纳州里奇兰。国际自主智能体与多智能体系统基金会。[KM01] Nilesh N. Karnik 和 Jerry M. Mendel。类型- 2 *Informa-* 模糊集的质心。第（ ）*tion Sciences* 132 1 195 220, 2001 卷（ ）：–。URL: https://www.sciencedirect.com/science/article/pii/ S002002550100069X, doi:10.1016/S0020-0255(01)00069-X。[Kuz26] A.V. Kuznetsov。关于命题值随机变量。 *Boletín de la Sociedad Matemática Mexicana* 32 8 8, 2026 10 1007 40590 025 00840 7 ， （ ）：。doi: 。/s - - -。[MQK25] Hannah Mertens、Tim Quatmann 和 Joost-Pieter Katoen。参数化概率自动机的组合推理，2025 年。URL: https://arxiv.org/abs/2506.08525，arXiv:2506.08525。[Nil86] Nils J. Nilsson。概率逻辑。 *Artificial Intelligence*，28(1):71–87，1986 年。URL: https://www. sciencedirect.com/science/article/pii/0004370286900317, doi:10.1016/0004-3702(86)90031-7。[RD06] Matthew Richardson 和 Pedro Domingos。马尔可夫逻辑网络。 *Machine Learning*，62(1-2):107–136，2006 年 2 月。doi:10.1007/s10994-006-5833-1。[SDKZ02] Ilya Shmulevich、Edward R. Dougherty、Seungchan Kim 和 Wei Zhang。概率布尔网络：一种基于规则的基因调控网络不确定性模型。 *Bioinformatics*，18(2):261–274，2002 年 2 月。doi:10.1093/bioinformatics/18.2.261。[SDS02] D. E. Sitnikov、B. D' Cruz 和 P. E. Sitnikova。一种基于逻辑方程组合与操作的知识表示和发现方法。 *WIT Transactions on Information and Communication Technologies* 28 2002 10 2495 020031 ， ， 。doi: 。/DATA 。[Sha76] Glenn Shafer。 *A Mathematical Theory of Evidence*。普林斯顿大学出版社，新泽西州普林斯顿，1976 年。URL: https://www.jstor.org/stable/j.ctv10vm1qb。[SK18] Andrew Schumann 和 Alexander V. Kuznetsov。侧向抑制和侧向激活神经科学条件下的数学基础。 *International Journal of Parallel Emergent and Distributed Systems* 33 3 237 256, 2018 ， ， （ ）：–。 10 1080 17445760 2018 arXiv:https://doi.org/ 。/ 。 。1439490，doi:10.1080/17445760.2018.1439490。[SK21] V.A. Sergeev 和 N.A. Korgin。将综合评级机制识别作为离散数据分析的一种方法。 *IFAC-PapersOnLine*，54(13):134–139，2021 年。第 20 届 IFAC 技术、文化与国际稳定会议 TECIS 2021。URL: https://www.sciencedirect. com/science/article/pii/S240589632101870X，doi:10.1016/j.ifacol.2021.10.433。[Vas11] Vadim Vasyukevich。 *Asynchronous Operators of Sequential Logic: Venjunction & Sequention.* , 《》第 卷。 *Digital Circuit Analysis and Design* 124 *Lecture Notes in Electrical Engineering* Springer 柏林、海德堡，2011 年。doi:10.1007/978-3-642-21611-4。[Wha74] R. M. Wharton。近似语言识别。 *Information and Control* 26 236 255, , ：–。URL: 1974 0019995874913692, 10 1016 https://www.sciencedirect.com/science/article/pii/S doi: 。/ S0019-9958(74)91369-2。

Appendix A. Discussion on the PDNFs Algebra

将概率析取范式视为一系列针对特征 $i \in \{1, \ldots, m\}$ 的观测（例如随时间采集的传感器读数）。可将概率析取范式 Z_1 中的指数 ξ_i 解读为累积权重或参数，用于概括特征 i 在序列 Z_1 中的影响；而 η_i 则表示特征 i 在序列 Z_2 中对应的权重。如果我们合并序列Z_1 和 Z_2 中的证据——假设它们属于同一观测框架——那么每个特征的综合影响自然可用总和 $\xi_i + \eta_i$ 来建模，这与.....相对应

它恰好对应于基本合取项所定义的加法（并通过 (2.5) 扩展到概率析取范式）。因此，范数 (2.6) 提供了概率析取范式绝对权重总和的度量。两个概率析取范式之间的距离较小（以 $\|Z_1 - Z_2\|_1$ 衡量），表明相应的观测序列赋予各特征的净权重相近，意味着它们在统计上相似。

考虑一个固定位置（$i,\ j$）以及一列相互独立的概率析取范式 $Z_1,\ Z_2$、.....，其在该位置的权重为 $\xi_{ij}^{(k)}$。假设所有这些权重具有相同的符号，例如，对于每个 k 都有 $\xi_{ij}^{(k)} \geq 0$（或全部为 ≤ 0）。构造累积 PDNF $Z^{(N)} = Z_1 + \cdots + Z_N$；其在位置（$i,\ j$）处的权重为 $\xi_{ij}^{(N)} = \sum_{k=1}^{N} \xi_{ij}^{(k)}$。

如果权重非负，并且存在一个常数 $c > 0$，使得其中无穷多个满足 $\xi_{ij}^{(k)} \geq c$，那么当 $N \to \infty$ 时，$\xi_{ij}^{(N)} \to +\infty$。例如，当每次观测中的权重都有一个正的下界，或者获得高于某个正阈值的权重的概率为正时，该条件成立（在后一种情况下，几乎必然会发生无穷多次观测）。

根据映射 F_ℓ^j（定义 2.1）的单调性，以及自然假设 $F_1^j(\xi) \to 1$ 和 $F_0^j(\xi)\ F_{-1}^j(\xi) \to 0$ 当 $\xi \to +\infty$ 时，我们得到

$$F_1^j(\xi_{ij}^{(N)}) \longrightarrow 1, \qquad F_0^j(\xi_{ij}^{(N)}) \longrightarrow 0, \qquad F_{-1}^j(\xi_{ij}^{(N)}) \longrightarrow 0.$$

因此，当 N 足够大时，该位置的实现几乎必然为 x_j；其行为将变为确定性的。

类似地，如果所有权重均非正，且其中有无限多个 $\xi_{ij}^{(k)} \leq -c$，那么$\xi_{ij}^{(N)} \to -\infty$ 并且，在 F_{-1}^j 具有类似渐近行为的条件下，我们得到

$$F_{-1}^j(\xi_{ij}^{(N)}) \longrightarrow 1, \qquad F_0^j(\xi_{ij}^{(N)}) \longrightarrow 0, \qquad F_1^j(\xi_{ij}^{(N)}) \longrightarrow 0,$$

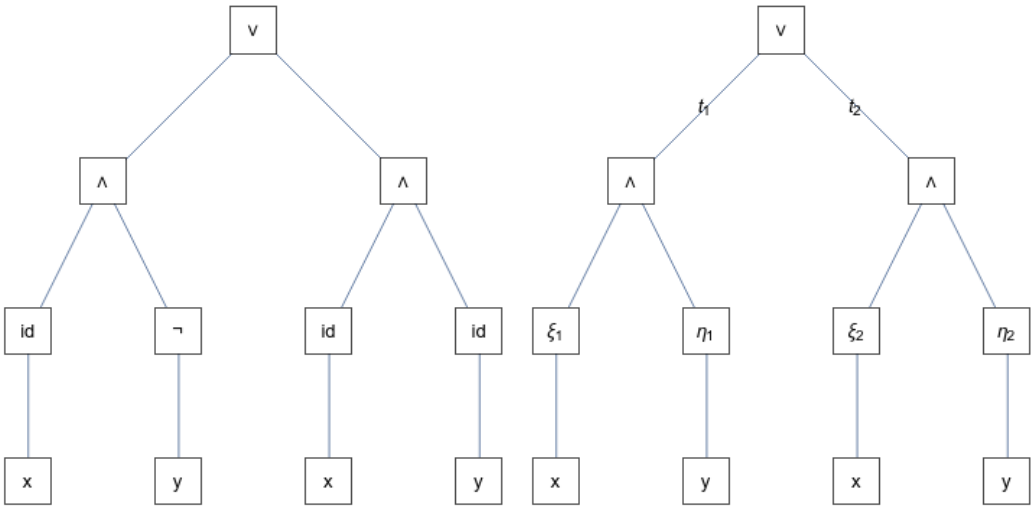
因此，输出几乎必然趋于 x_j。 $\square$

这一结果将如下直觉形式化：同一类型的证据反复出现（例如，传感器持续触发）会逐渐消除不确定性，并推动系统趋向确定性的结果。

Appendix B. Discussion on the PDNF Structure

本质上，含 m 个变量的析取范式可以表示为一棵树图（见图2a）。该图的叶节点对应于变量，根顶点表示析取符号 $\vee$，根的子节点是合取符号 $\wedge$，每个合取符号的子节点要么是否定符号 $\neg$，要么是恒等符号id（其中id(x) $= x$）。树的边没有标记；因此，同一父节点的子节点可以按任意顺序排列，而不影响逻辑含义。

概率析取范式以两种关键方式修改这一结构（见图2b）：首先，否定或恒等符号被实值权重 ξ_i所取代；其次，更为重要的是，连接根与合取符号的边被加以标记（例如，使用时间索引）。这些标记记录了例如每个合取的观测时刻，从而使各合取的顺序固定并具有语义上的重要性。



(A) (B)

FIGURE 2. 经典析取范式 $x \wedge y \vee x \wedge y$ (a)与概率析取范式$x^{\xi_1} \wedge y^{\eta_1} \vee x^{\xi_2} \wedge y^{\eta_2}$ (b)的比较

APPENDIX C. DISCUSSION ON THE PDNFS MEASURE

命题C.1. 函数 $\mu_\Xi : 2^{\mathcal{V}(m,n)} \to \mathbb{R}$ 可加地延拓到任意子集$A \subseteq \mathcal{V}(m,n)$ 为

$$\mu_\Xi(A) = \sum_{\langle v\rangle \in A} \mu_\Xi(\{\langle v\rangle\}).$$

是可测空间$(\mathcal{V}(m,n)\ 2^{\mathcal{V}(m,n)})$上的概率测度。

证明. 我们应验证概率测度的两条定义性质:
(1) 非负性与归一化: $0 \le \mu_\Xi(A) \le 1$ 对每个 $A \subseteq \mathcal{V}(m,n)$成立，且$\mu_\Xi(\mathcal{V}(m,n)) = 1$。
(2) 可数可加性（由于 $\mathcal{V}(m,n)$有限，故可归结为有限可加性）：对任意两两不交的集合 A_1, $A_2, \cdots \subseteq \mathcal{V}(m,n)$

$$\mu_\Xi\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu_\Xi(A_k).$$

非负性。根据定义 2.1，函数 F^j_{-1}、F^j_0、F^j_1、$j = 1,\ m$ 的取值均属于 $[0,1]$。因此，定义 $\mu_\Xi(\{\langle v\rangle\})$ 的乘积中的每个因子都是非负的，进而对于每个 $\langle v\rangle \in \mathcal{V}(m,\ n)$，均有 $\mu_\Xi(\{\langle v\rangle\}) \ge 0$。对于任意子集 $A \subseteq \mathcal{V}(m,\ n)$，测度 $\mu_\Xi(A)$ 是若干非负数之和，因此 $\mu_\Xi(A) \ge 0$。

归一化。我们证明整个空间的总测度等于一:

$$\mu_\Xi(\mathcal{V}(m,n)) = \sum_{\langle v\rangle \in \mathcal{V}(m,n)} \mu_\Xi(\{\langle v\rangle\}) = 1.$$

集合 $\mathcal{V}(m,n)$ 与对每个满足 $i = 1,n$、$j = 1,m$ 的数对 (i,j) 的所有 $v_{ij} \in \mathcal{C} = \{x_j, \varepsilon, x_j\}$ 选择一一对应。在不同

位置 (i,j) 处的选择相互独立的假设下，我们可以交换乘积与求和：

$$\sum_{\langle v\rangle\in\mathcal{V}(m,n)}\mu_\Xi(\{\langle v\rangle\}) = \sum_{v_{11}\in\mathcal{C}}\cdots\sum_{v_{nm}\in\mathcal{C}}\prod_{i=1}^{n}\prod_{j=1}^{m}\begin{cases}F_{-1}^j(\xi_{ij}) & \text{if } v_{ij}=\overline{x}_j,\\ F_0^j(\xi_{ij}) & \text{if } v_{ij}=\varepsilon,\\ F_1^j(\xi_{ij}) & \text{if } v_{ij}=x_j,\end{cases} =$$
$$=\prod_{i=1}^{n}\prod_{j=1}^{m}\Big(F_{-1}^j(\xi_{ij})+F_0^j(\xi_{ij})+F_1^j(\xi_{ij})\Big).$$

根据定义 2.1，对于每个 $j=1,\ \overline{m}$ 以及每个 $x\in\mathbb{R}$，我们有

$$F_{-1}^j(x)+F_0^j(x)+F_1^j(x)=1.$$

将此恒等式应用于 $x=\xi_{ij}$，得到 $F_{-1}^j(\xi_{ij})+F_0^j(\xi_{ij})+F_1^j(\xi_{ij})=1$，对每个 $(i,\ j)$ 均成立。因此

$$\sum_{\langle v\rangle\in\mathcal{V}(m,n)}\mu_\Xi(\{\langle v\rangle\})=\prod_{i=1}^{n}\prod_{j=1}^{m}1=1.$$

可加性。由于 $\mathcal{V}(m,\ n)$ 是有限的，只需证明有限可加性。设 $A_1,\ A_2,\ \ldots,\ A_K\subseteq\mathcal{V}(m,\ n)$ 两两不相交。那么

$$\mu_\Xi\left(\bigcup_{k=1}^{K}A_k\right)=\sum_{\langle v\rangle\in\bigcup_{k=1}^{K}A_k}\mu_\Xi(\{\langle v\rangle\})=\sum_{k=1}^{K}\sum_{\langle v\rangle\in A_k}\mu_\Xi(\{\langle v\rangle\})=\sum_{k=1}^{K}\mu_\Xi(A_k).$$

两个条件均满足，因此 μ_Ξ 是一个概率测度。 □

Appendix D. PDNFs and Bayesian Fusions

定理 D.1. 设 $F_\ell\colon\mathbb{R}\to(0,\ 1),\ \ \ell\in\{-1,0,1\}$ 为连续函数，满足$\sum_{\ell\in\{-1,0,1\}}F_\ell(\xi)=1$ 对所有 $\xi\in\mathbb{R}$均成立。假设对于所有 $\xi,\ \eta\in\mathbb{R}$ 都满足以下性质：

$$F_\ell(\xi+\eta)=\frac{F_\ell(\xi)F_\ell(\eta)}{\sum_{\ell'\in\{-1,0,1\}}F_{\ell'}(\xi)F_{\ell'}(\eta)},\qquad \ell\in\{-1,0,1\}.$$

那么存在常数 $\alpha_\ell\in\mathbb{R}$，使得

$$F_\ell(\xi)=\frac{\exp(\alpha_\ell\xi)}{\sum_{\ell'\in\{-1,0,1\}}\exp(\alpha_{\ell'}\xi)}.$$

证明. 定义 $g_\ell(\xi)=\ln F_\ell(\xi)$。假设变为

$$\exp\big(g_\ell(\xi+\eta)\big)=\frac{\exp\big(g_\ell(\xi)+g_\ell(\eta)\big)}{\sum_{\ell'\in\{-1,0,1\}}\exp\big(g_{\ell'}(\xi)+g_{\ell'}(\eta)\big)}.$$

取对数可得

$$g_\ell(\xi+\eta)=g_\ell(\xi)+g_\ell(\eta)-\ln\left(\sum_{\ell'\in\{-1,0,1\}}\exp\big(g_{\ell'}(\xi)+g_{\ell'}(\eta)\big)\right).$$

最后一项与 ℓ 无关；将其记为 $h(\xi,\ \eta)$。那么对于任意 ℓ，

$$g_\ell(\xi+\eta)-g_\ell(\xi)-g_\ell(\eta)=h(\xi,\eta).$$

左侧取决于 ℓ，而右侧与之无关。因此，对于任意两个指标 ℓ，ℓ'，

$$g_\ell(\xi+\eta)-g_\ell(\xi)-g_\ell(\eta)=g_{\ell'}(\xi+\eta)-g_{\ell'}(\xi)-g_{\ell'}(\eta).$$

重新整理，

$$\big(g_\ell(\xi+\eta)-g_{\ell'}(\xi+\eta)\big)-\big(g_\ell(\xi)-g_{\ell'}(\xi)\big)-\big(g_\ell(\eta)-g_{\ell'}(\eta)\big)=0.$$

定义 $\phi_{\ell,\ell'}(\xi)=g_\ell(\xi)-g_{\ell'}(\xi)$。那么上述方程可写为

$$\phi_{\ell,\ell'}(\xi+\eta)=\phi_{\ell,\ell'}(\xi)+\phi_{\ell,\ell'}(\eta).$$

因此$\phi_{\ell,\ell'}$ 满足柯西函数方程。由 F_ℓ的可测性可知，函数 g_ℓ 连续，因此$\phi_{\ell,\ell'}$ 连续。柯西方程的唯一连续解是线性函数：$\phi_{\ell,\ell'}(\xi)=c_{\ell,\ell'}\xi$ 其中某个常数为$c_{\ell,\ell'}\in\mathbb{R}$。此外，$\phi_{\ell,\ell'}=-\phi_{\ell',\ell}$，因此$c_{\ell,\ell'}=-c_{\ell',\ell}$。现在固定一个参考索引，例如 0。对于每个 ℓ，令$\alpha_\ell=c_{\ell,0}$。那么$\phi_{\ell,0}(\xi)=\alpha_\ell\xi$，即

$$g_\ell(\xi)-g_0(\xi)=\alpha_\ell\xi,$$

或者等价地，

$$g_\ell(\xi)=g_0(\xi)+\alpha_\ell\xi.$$

利用归一化条件 $\sum_{\ell\in\{-1,0,1\}}\exp(g_\ell(\xi))=1$，得到

$$\sum_{\ell\in\{-1,0,1\}}\exp\big(g_0(\xi)+\alpha_\ell\xi\big)=\exp(g_0(\xi))\sum_{\ell}\exp(\alpha_\ell\xi)=1.$$

因此 $\exp(g_0(\xi))=\big(\sum_\ell\exp(\alpha_\ell\xi)\big)^{-1}$，所以

$$F_\ell(\xi)=\exp(g_\ell(\xi))=\frac{\exp(\alpha_\ell\xi)}{\sum_{\ell'}\exp(\alpha_{\ell'}\xi)}.$$

证明至此完成。 □

该定理表明，只有指数（softmax）形式能够使概率析取范式空间中的权重相加与独立证据的贝叶斯融合完全对应。在实际应用中，也可以使用其他满足定义 2.1 条件的族$\mathbf{F}$，但此时加法组合只是近似贝叶斯融合，而非精确实现它。加法结构仍然体现了这样一种思想：较大的权重ξ_{ij} 对应更强、更明确的预测。

Appendix E. Discussion on the Probability to Generate a Venjunction Close to the Given One

让我们估计随机生成的 venjunction $\langle v\rangle$（根据μ_Ξ）在某种合适的度量下，与一个长度相同的固定目标 venjunction $\langle v'\rangle$ 接近的概率。在 $\mathcal{V}(m,\ n)$ 上，一个自然的度量是加权汉明距离。对于两个文字 $a,\ b\in\{x_j,\varepsilon,\overline{x}_j\}$（由于位置对齐，索引 j 相同），定义

$$w(a,b)=\begin{cases}0, & a=b,\\ 1, & (a=\varepsilon,\ b\in\{x_j,\overline{x}_j\})\text{ or }(b=\varepsilon,\ a\in\{x_j,\overline{x}_j\}),\\ 2, & \{a,b\}=\{x_j,\overline{x}_j\}.\end{cases}$$

对于两个 venjunction $\langle v\rangle=v_{n1}\wedge\cdots\wedge v_{nm}\angle\ldots\angle v_{11}\wedge\cdots\wedge v_{1m}$和 $\langle v'\rangle=v'_{n1}\wedge\cdots\wedge v'_{nm}\angle\ldots\angle v'_{11}\wedge\cdots\wedge v'_{1m}$，我们定义

$$d(v,v') = \sum_{i=1}^{n}\sum_{j=1}^{m} w\big(v_{ij}, v'_{ij}\big).$$

很容易验证， d 是一个度量（非负、对称且满足三角不等式），其取值为从 0 到 $2nm$ 的整数。

固定一个目标合取项 $\langle v'\rangle$ 和一个权重矩阵 Ξ。由于在不同位置（$i,\ j$）的实现相互独立（在 μ_Ξ 下），随机变量

$$\zeta_{ij} := w\big(v'_{ij}, v_{ij}\big)$$

相互独立。它们的分布取决于 v_{ij} 以及概率 $F^j_\ell(\xi_{ij})$。记

$$p^{(0)}_{ij} = \mathbb{P}_\Xi\{\zeta_{ij}=0\},\quad p^{(1)}_{ij} = \mathbb{P}_\Xi\{\zeta_{ij}=1\},\quad p^{(2)}_{ij} = \mathbb{P}_\Xi\{\zeta_{ij}=2\}.$$

这些概率很容易通过 F^j_{-1}， F^j_0， F^j_1 表示（见表1）。

v_{ij}	$p^{(0)}_{ij}$	$p^{(1)}_{ij}$	$p^{(2)}_{ij}$
x_j	F^j_1	F^j_0	F^j_{-1}
$\overline{x}_j$	F^j_{-1}	F^j_0	F^j_1
ε	F^j_0	$F^j_1 + F^j_{-1}$	0

TABLE 1. 贡献 ζ_{ij} 的概率取决于目标文字。

因此，总距离

$$S := d(v,v') = \sum_{i=1}^{n}\sum_{j=1}^{m}\zeta_{ij}$$

是取值于 $\{0,1,2\}$ 的相互独立（但分布不相同）随机变量之和。

定理 E.1. $\langle v'\rangle$ 落在半径为 ϱ 的球内的概率（其中 ϱ 是非负整数）

$$\mathbb{P}_\Xi\{d(v,v')\le\varrho\} = \sum_{k=0}^{\varrho} c_k,$$

其中 c_k 是生成函数的系数

$$G(t) = \prod_{i=1}^{n}\prod_{j=1}^{m}\big(p^{(0)}_{ij} + p^{(1)}_{ij}t + p^{(2)}_{ij}t^2\big) = \sum_{k=0}^{2nm} c_k t^k.$$

证明。 $\langle v'\rangle$ 落在半径为 ϱ的球内的概率为

$$\mathbb{P}_\Xi\{d(v,v')\le\varrho\} = \sum_{k=0}^{\varrho}\mathbb{P}_\Xi\{S=k\}.$$

定义 $c_k = \mathbb{P}_\Xi\{S=k\}$。 S 的分布可以通过概率生成函数获得

$$G(t)=\prod_{i=1}^{n}\prod_{j=1}^{m}\bigl(p_{ij}^{(0)}+p_{ij}^{(1)}t+p_{ij}^{(2)}t^{2}\bigr)=\sum_{k=0}^{2nm}c_{k}t^{k}.$$

因此 $\mathbb{P}_{\Xi}\{S\le\varrho\}=\sum_{k=0}^{\varrho}c_{k}$。 □

对于适中的 n，m，系数 c_k 可以通过逐次卷积计算。当 n 和 m 较大时，可以采用正态近似：

$$\mathbb{P}_{\Xi}\{d(v,v')\le\varrho\}\approx\Phi\!\left(\frac{\varrho+0.5-\mu_S}{\sigma_S}\right),$$

其中 Φ 是标准正态分布函数，且

$$\mu_S=\sum_{i,j}\mathbb{E}[\zeta_{ij}],\qquad\sigma_S^2=\sum_{i,j}\operatorname{Var}(\zeta_{ij})$$

with

$$\mathbb{E}[\zeta_{ij}]=p_{ij}^{(1)}+2p_{ij}^{(2)},\qquad\operatorname{Var}(\zeta_{ij})=\mathbb{E}[\zeta_{ij}^2]-(\mathbb{E}[\zeta_{ij}])^2,\quad\mathbb{E}[\zeta_{ij}^2]=p_{ij}^{(1)}+4p_{ij}^{(2)}.$$

当 μ_S 不太接近 0 或 $2nm$，且 σ_S 足够大时，这种近似是准确的。

例 E.2. 我们用一个简单的案例来说明计算过程。设有 $m=1$ 个传感器和 $n=2$ 个时间步，则集合 $\mathcal{V}(1,2)$由形如 $\langle v\rangle=v_{21}\angle v_{11}$的 venjunction 组成（请注意，顺序是反过来的，即最右边的合取项对应最早的时间）。选择目标 venjunction

$$\langle v'\rangle=x_1\angle x_1\quad(\text{so } v'_{11}=x_1,\ v'_{21}=x_1).$$

假设两个位置的概率映射相同，且由下式给出

$$F_1(\xi)=0.6,\quad F_0(\xi)=0.3,\quad F_{-1}(\xi)=0.1$$

（.....的值 ξ 无关紧要，因为我们将其固定不变）。因此，对于每个位置（i，1），我们有

$$p^{(0)}=\mathbb{P}\{\zeta_{i1}=0\}=F_1=0.6,\quad p^{(1)}=\mathbb{P}\{\zeta_{i1}=1\}=F_0=0.3,\quad p^{(2)}=\mathbb{P}\{\zeta_{i1}=2\}=F_{-1}=0.1.$$

随机变量 ζ_{11} 和 ζ_{21} 相互独立且同分布。它们共同的概率生成函数为

$$G_\zeta(t)=0.6+0.3t+0.1t^2.$$

总距离 $S=\zeta_{11}+\zeta_{21}$ 的生成函数为

$$G_S(t)=0.36+2\cdot0.6\cdot0.3\,t+\bigl(2\cdot0.6\cdot0.1+0.3^2\bigr)t^2+2\cdot0.3\cdot0.1\,t^3+0.1^2\,t^4=$$
$$=0.36+0.36\,t+0.21\,t^2+0.06\,t^3+0.01\,t^4.$$

因此，可能距离的概率为

$$\mathbb{P}\{S=0\}=0.36,\ \mathbb{P}\{S=1\}=0.36,\ \mathbb{P}\{S=2\}=0.21,\ \mathbb{P}\{S=3\}=0.06,\ \mathbb{P}\{S=4\}=0.01.$$

由此，我们得到累积概率：

ε	$\mathbb{P}\{d(v,\ v')\ \le \varepsilon\}$
0	0.36
1	0.72
2	0.93
3	0.99
4	1.00

因此，例如，随机抽取的 venjunction 与 v' 最多相差 2 的概率（即所有位置都相同，或有一个 ε 或两个 ε 等）为 93%。

正态近似。 此处 $\mu_S = 2(0.3 + 2 \cdot 0.1) = 2(0.5) = 1.0$ 和 $\sigma_S^2 = 2\big[(0.3 + 4 \cdot 0.1) - 0.5^2\big] = 2(0.7 - 0.25) = 2 \cdot 0.45 = 0.9$，因此 $\sigma_S \approx 0.949$。对于 $\varepsilon = 2$，该近似给出

$$\Phi\left(\frac{2 + 0.5 - 1.0}{0.949}\right) = \Phi(1.58) \approx 0.943,$$

该值与精确值 0.93 相当接近。

Appendix F. Control Systems Examples

示例 F.1. 一个具有不确定性的简单函数控制系统。

考虑由两个传感器组成的系统：大气压力 p 和空气温度 θ。当压力 $p(t) > p_0$时，压力传感器会在时间 t 触发；当温度 $\theta(t) > \theta_0$时，温度传感器会在时间 t 触发。此外，在空气湿度较高时，传感器也可能在时间 t 误触发，空气湿度 $\varphi(t) > \varphi_0$。如果两者 $p(t) > p_0$ 且 $\theta(t) > \theta_0$ 持续足够长的时间，我们就必须采取一些行动。

如果压力传感器在时间 t触发，我们记作 $x(t)$。如果它在时间 t未触发，我们记作 $\bar{x}(t)$。如果温度传感器在时间 t触发，我们记作 $y(t)$。如果它在时间 t未触发，我们记作 $\bar{y}(t)$。最后，如果 $\varphi(t) \le \varphi_0$, $t = 1,\ T$，我们应在以下情况下采取行动

$$Z = x(1) \wedge y(1) \vee \cdots \vee x(T) \wedge y(T), \quad T > T_0 > 1.$$

由于我们实际上不知道 $\varphi(t)$的值，因此无法确定 Z 是否对应真实情况，也无法确定它是否可用于决策。因此，我们可以改写为

$$Z(\varphi) = x(1)^{\varphi(1)} \wedge y(1)^{\varphi(1)} \vee \cdots \vee x(T)^{\varphi(T)} \wedge y(T)^{\varphi(T)}.$$

这里，$x(t)^{\varphi(t)} = x(t)$、$y(t)^{\varphi(t)} = y(t)$、$\varphi(t) \le \varphi_0$。如果 $\varphi(t) > \varphi_0$，$x(t)^{\varphi(t)}$ 可以取值$x(t)$、$\bar{x}(t)$和 $y(t)^{\varphi(t)}$ 可以取值 $y(t)$、$\bar{y}(t)$。

我们希望，如果 $p(t) > p_0 \wedge \varphi(t) > \varphi_0$，那么 $x(t)^{\varphi(t)} = x(t)$；但如果 $p(t) \le p_0 \wedge \varphi(t) > \varphi_0$，那么 $x(t)^{\varphi(t)} = \bar{x}(t)$。遗憾的是，我们没有湿度传感器，因此无法做出上述推断。相反，我们知道一些关于湿度和传感器的事实，即：

▷ 温度、压力和湿度的可能取值分别位于开区间（$\theta_{\min}, \theta_{\max}$）、（$p_{\min}, p_{\max}$）和（0,1）内，任意参数值都是等概率的。▷ 温度传感器有时会在 $\theta_x \le \theta(t) < \theta_0$, $\varphi(t) > \varphi_0$时误触发，θ_x 是误触发阈值。

▷ 压力传感器有时会在 $p_y \le p(t) < p_0, \varphi(t) > \varphi_0$时误触发， p_y 是误触发阈值。

因此，温度传感器 $x^{\varphi(t)}$ 有三种可能的输出：正确触发 $x(t)$、可疑触发 $\varepsilon(t)$，以及正确未触发 $\overline{x}(t)$。

我们可以粗略估计这些输出的概率如下：

$$\mathbb{P}\{x^{\varphi(t)} \sim x(t)\} = \frac{\theta_{\max} - \theta_0}{\theta_{\max} - \theta_{\min}}, \quad \mathbb{P}\{x^{\varphi(t)} \sim \varepsilon(t)\} = \frac{\theta_0 - \theta_x}{\theta_{\max} - \theta_{\min}},$$
$$\mathbb{P}\{x^{\varphi(t)} \sim \overline{x}(t)\} = \frac{\theta_x - \theta_{\min}}{\theta_{\max} - \theta_{\min}}. \quad \text{(F.1)}$$

压力传感器 $y^{\varphi(t)}$ 也有三种可能的输出：正确触发 $y(t)$、可疑触发 $\varepsilon(t)$，以及正确未触发 $\overline{y}(t)$。我们可以估计这些输出的概率如下：

$$\mathbb{P}\{y^{\varphi(t)} \sim y(t)\} = \frac{p_{\max} - p_0}{p_{\max} - p_{\min}}, \quad \mathbb{P}\{y^{\varphi(t)} \sim \varepsilon(t)\} = \frac{p_0 - p_y}{p_{\max} - p_{\min}},$$
$$\mathbb{P}\{y^{\varphi(t)} \sim \overline{y}(t)\} = \frac{p_y - p_{\min}}{p_{\max} - p_{\min}}. \quad \text{(F.2)}$$

因此，我们可以将时间 t 时传感器系统的输出视为概率析取范式 $x^{\theta_0}(t) \wedge y^{p_0}(t)$，其中映射 **F** 对应于等式 (F.1) 、 (F.2)。如果温度和压力服从非均匀分布，则应相应地修正 (F.1) 、 (F.2)。

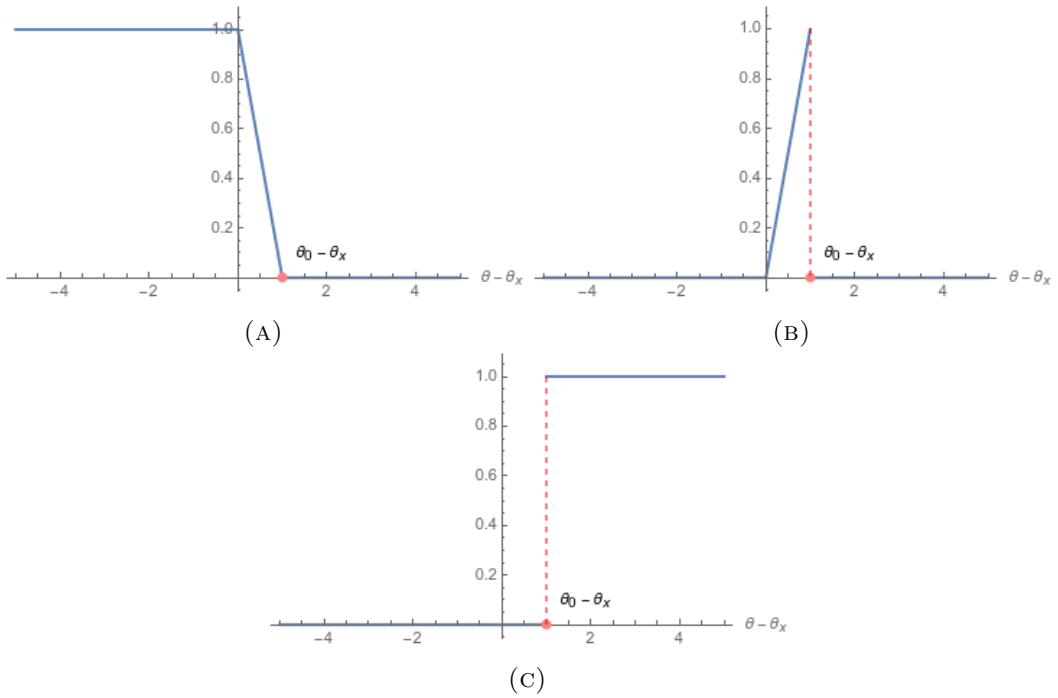


FIGURE 3. 映射 **F** 的“温度”分量 F^x_{-1}(a)、 F^x_0(b)、 F^x_1(c)

如果我们能以某种方式找出正确的温度和压力值（例如，通过无法触发的附加传感器），就可以进一步

考虑概率析取范式 $x^{\theta(t)}(t) \wedge y^{p(t)}(t)$。在确定概率时，有必要以某种方式评估这样一种情况：当 $\theta(t)$ 接近 θ_x 时， $\theta_x < \theta(t) < \theta_0$，传感器可能无法正常工作，也可能正常工作。在这种情况下，映射 $\mathbf{F}$ 应定义为

$$F^x_{-1}(\theta) = \begin{cases} 1, & \theta < \theta_x, \\ -(\theta - \theta_x)/(\theta_0 - \theta_x) + 1, & \theta_x \le \theta < \theta_0, \\ 0, & \theta \ge \theta_0, \end{cases} \tag{F.3}$$

$$F^x_0(\theta) = \begin{cases} 0, & \theta < \theta_x, \\ (\theta - \theta_x)/(\theta_0 - \theta_x), & \theta_x \le \theta < \theta_0, \\ 0, & \theta \ge \theta_0, \end{cases} \tag{F.4}$$

$$F^x_1(\theta) = \begin{cases} 0, & \theta < \theta_0, \\ 1, & \theta \ge \theta_0, \end{cases} \tag{F.5}$$

对于温度分量（F^x_{-1}, F^x_0, F^x_1）（见图 3），并以类似方式，对于压力分量（F^y_{-1}, F^y_0, F^y_1）。

应当指出，时间时刻 $t_1, \dots, t_m$的系统状态序列可以写成概率析取范式 $x^{\theta(t_1)}(t_1) \wedge y^{p(t_1)}(t_1) \vee \cdots \vee x^{\theta(t_m)}(t_m) \wedge y^{p(t_m)}(t_m)$。在实践中，人们~可以使用感知值 $\theta(t)$ 和 $\tilde{p}(t)$（由传感器或其他来源获得），而不是无法获取的真实值 $\theta(t)$ 和 $p(t)$，以通过 F^x_0和 F^y_0 评估传感器处于可疑状态的风险。

~假设所测得的值满足 $\theta(t) \ge \theta_0$ 和 $\tilde{p}(t) \ge p_0$，但在 $|\theta(t) - \theta_0| < \beta_\theta$ 和 $|\tilde{p}(t) - p_0| < \beta_p$（对于较小的正常数$\beta_\theta, \beta_p$）这一意义上接近~阈值。这种接近性增加了真实值 $\theta(t)$ $p(t)$ 实际上落在“可疑区间” $[\theta_x, \theta_0)$ 和 $[p_y, p_0)$ 之内的可能性。为防范此类测量误差，我们引入安全裕度 $\lambda_\theta = \sigma_\theta(\theta_0 - \theta_x)$、 $\lambda_p = \sigma_p(p_0 - p_x)$，其中 $\sigma_\theta, \sigma_p \in (0,1)$ 为启发式选取的参数。仅当对于所有时刻 $t_1, \dots, t_m$ 我们有

$$F^x_1\big(\tilde{\theta}(t_i) - \lambda_\theta\big) = 1, \quad F^y_1\big(\tilde{p}(t_i) - \lambda_p\big) = 1$$

并且传感器会被触发。

因此，我们应使用上述所述经过修正的映射 $\mathbf{F}_C$ 进行决策。这样，我们无需用相当繁琐的文字描述传感器系统及其决策规则，而是得到了一个令人联想到布尔代数的数学表达式。

使用文章 [Kuz26]，中所述的方法，我们之后可以根据传感器系统的实际观测重构概率析取范式，并将得到的概率析取范式与理论预测的概率析取范式进行比较。重构结果与理论结果之间的差异表明传感器中存在未知的系统误差。

例 F.2. 概率析取范式与随机游走之间的联系。

假设前一个例子中的温度 θ 和压力 p 是随机游走，例如

$$\theta(t) = 20 + \frac{1}{4}\mathcal{RW}\left(\frac{1}{2}, \frac{1}{3}\right),$$
$$p(t) = 100 + \frac{1}{2}\mathcal{RW}\left(\frac{1}{2}, \frac{1}{3}\right),$$

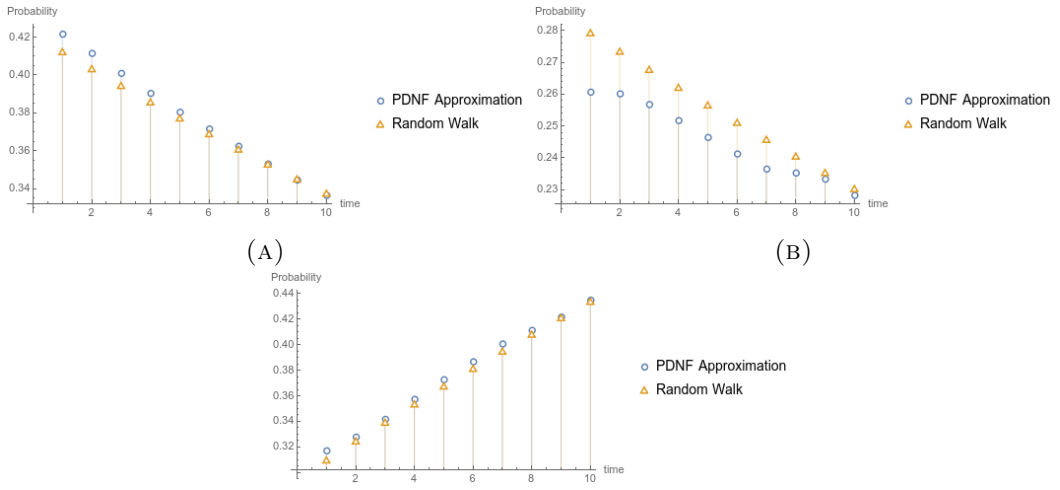
其中， $\mathcal{RW}$ 是格点随机游走，向上迈出单位步长的概率为1 2，向下迈出单位步长的概率为1 3，迈出零单位步长的概率为1 6。设时间t 属于集合$\mathbf{T} = \{30, \ldots, 39\}$, $\theta_x = 20$, $\theta_0 = 22$, $p_y = 105$, $p_0 = 106$。时间以秒为单位，温度以摄氏度为单位，压力以千帕为单位。我们使用计算机代数系统 Wolfram Mathematica 对此模型进行一个非常简单的数值模拟。

引入随机变量 $error$：当传感器误触发时，其取值为 0，否则取值为 1。每个取值的概率均为 1 2。我们比较针对每个 $t \in \mathbf{T}$ 进行 10000 次实验后计算得到的函数（F.3）–（F.5）的平均值，以及以下事件的概率

$$\chi^x_{-1}(t) = \theta(t) < \theta_x \vee \theta(t) \geq \theta_x \wedge \theta(t) < \theta_0 \wedge error = 0, \tag{F.6}$$

$$\chi^x_0(t) = \theta(t) \geq \theta_x \wedge \theta(t) < \theta_0 \wedge error = 1, \tag{F.7}$$

$$\chi^x_1(t) = \theta(t) \geq \theta_0. \tag{F.8}$$



(A) (B)

(C)

FIGURE 4. 映射分量 F^x_{-1}(A)、 F^x_0(B)、 F^x_1(C) 的平均值（圆形）与对应值 $\chi^x_{-1}(t)$、 $\chi^x_0(t)$、$\chi^x_1(t)$（三角形）的比较

我们发现（见图 4），（F.3）–（F.5）的平均值基本上是概率（F.6）–（F.8）的近似值。

对于“压力”分量，也可以得到类似的图（见图 5）；其中我们考察以下事件的概率

$$\chi^y_{-1}(t) = p(t) < p_y \vee p(t) \geq p_y \wedge p(t) < p_0 \wedge error = 0, \tag{F.9}$$

$$\chi^y_0(t) = p(t) \geq p_y \wedge p(t) < p_0 \wedge error = 1, \tag{F.10}$$

$$\chi^y_1(t) = p(t) \geq p_0. \tag{F.11}$$

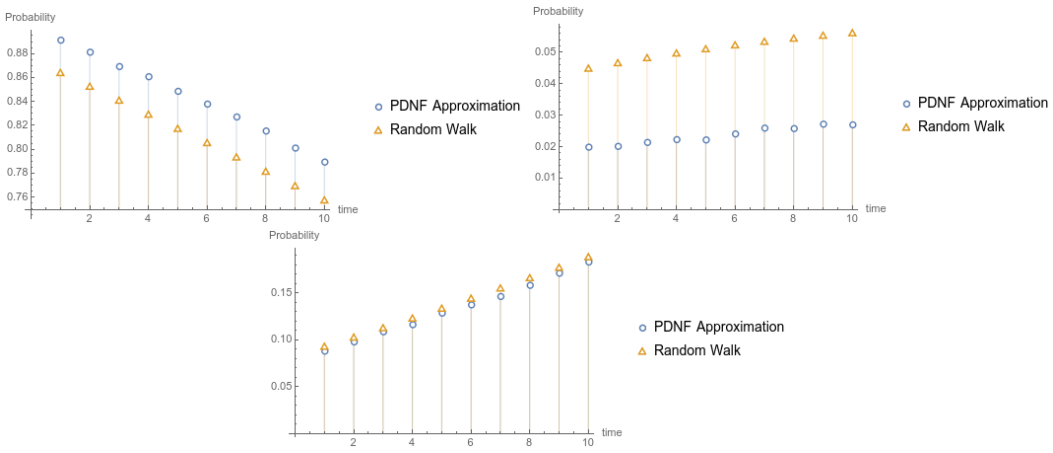


FIGURE 5. 映射分量（平均值（F^y_{-1}， F^y_0， F^y_1））（圆形）与元组（$\chi^y_{-1}(t)$ $\chi^y_0(t)$ $\chi^y_1(t)$）（三角形）的比较

本作品采用知识共享署名许可协议授权。如需查看该许可协议副本，请访问 https://creativecommons.org/licenses/by/4.0/，或致函知识共享，第二街171号，300室，旧金山，加利福尼亚州 94105，美国，或德国柏林10777，艾森纳赫街2号