arXiv:2302.02149v1 [cs.NE] 4 Feb 2023

神经自动机的不变量

JONE URIA-ALBIZURI、GIOVANNI SIRIO CARMANTINI、PETERBEIM GRABEN 和 SERAFIM RODRIGUES

Abstract. 神经动力系统的计算建模通常会采用神经网络和符号动力学。在一个名为 *vector symbolic architectures* 的框架中结合这两种方法的其中一种方式，会形成神经自动机。我们在该框架下探索的一个有趣研究方向，是考虑将符号动力学（例如由图灵机执行的符号动力学）映射到神经动力学，并以神经自动机表示。这种表示理论使我们能够提出诸如“大脑如何实现图灵计算？”之类的问题。具体而言，在这一表示理论中，神经自动机是通过将符号和符号串赋予数字而得到的，这种方法称为 G¨odel 编码。在这种赋值下，符号计算被表示为实相空间中的状态向量轨迹，从而可以与现实世界测量和实验数据进行统计相关分析。然而，这些赋值通常完全是任意的。因此，值得探讨这样一个问题：在这种表示下观察到的动力学有哪些方面是动力学本身固有的，又有哪些不是。本研究建立了一个形式严谨的数学框架，用于研究不同编码下神经自动机的对称性和不变量。作为一个核心概念，我们为这类系统定义了 *patterns of equality* 。我们考察了不同的宏观可观测量，例如神经网络的平均激活水平，并探究它们的不变性。我们的主要结果表明，只有定义在这些相等模式上的阶跃函数在符号重新编码下保持不变，而平均激活等则不具有不变性。我们的研究对于使用神经符号处理器对现实世界测量进行相关回归研究可能具有重要意义，有助于避免得出受特定编码影响、而非源自动力学本身的混杂结果。

计算认知神经动力学；符号动力学；神经自动机；观测量；不变量；语言处理

1. Introduction

计算认知神经动力学在很大程度上研究行为和神经生理可观测量与认知过程的神经计算模型之间的统计建模和回归分析（Gazzaniga 等人，2002；Rabinovich 等人，2012）。可通过实验测量的观测量包括反应时间（RT）、眼动（EM）、脑电图（EEG）领域中的事件相关脑电位（ERP）、脑磁图（MEG）领域中的事件相关磁场（ERF），以及功能性磁共振成像（fMRI）中的血氧水平依赖信号（BOLD）。

Date：2023年2月7日。

认知过程的计算模型通常涉及漂移扩散方法（Ratcliff，1978；Ratcliff 和 McKoon，2007）、ACT-R 等认知架构（Anderson 等人，2004）、自动机理论（Hopcroft 和 Ullman，1979）、动力系统（van Gelder，1998；Kelso，1995；Rabinovich 和 Varona，2018），尤其是神经网络（例如 Arbib（1995）；Hertz 等人（1991））。随着近年来深度学习技术的引入，神经网络日益受到欢迎（LeCun 等人（2015）；Schmidhuber（2015））。

为了对实验数据与计算模型进行统计相关分析，必须设计观测模型，将计算机模拟中的微观状态（例如模拟神经元的脉冲发放）与上述宏观可观测测量联系起来。例如，在决策领域，漂移扩散模型中的首次通过时间就是一种合适的观测模型（Ratcliff 1978，Ratcliff and McKoon 2007）。在神经电生理学领域，局部场电位（LFP）和脑电图（EEG）可以通过宏观平均场来描述，这些平均场基于神经元区室模型（beim Graben and Rodrigues 2013，Mart´ınez-Ca˜nada et al 2021，Mazzoni et al 2008）或神经场理论（beim Graben and Rodrigues 2014，Jirsa et al 2002）。针对磁共振成像（MRI）和血氧水平依赖信号，已有研究提出了专门的血流动力学观测模型（Friston et al 2000，Stephan et al 2004）。

在计算心理语言学和计算神经语言学领域（Arbib 和 Caplan 1979年，Crocker 1996年，beim Graben 和 Drenhaus 2012年，Lewis 2003），许多研究都对测量数据与模拟数据进行了统计回归分析。仅举数例：Davidson 和 Martin（2013）采用漂移扩散方法（Ratcliff 1978，Ratcliff 和 McKoon 2007），对西班牙语和巴斯克语受试者在翻译回忆实验中的速度-准确率数据建模。Lewis 和 Vasishth（2006）将英语句子在不同语言复杂度下的自定步调阅读时间，与 ACT-R 模型的预测（Anderson 等，2004）进行了相关分析。Huyck（2009）针对一项相关任务设计了由脉冲点神经元构成的赫布细胞组装网络。Stabler（2011）运用形式语言自动机模型（Hopcroft 和 Ullman 1979），阐述了阅读时间如何与自动机的工作记忆负荷相关联。类似地，Boston 等（2008）将眼动数据与概率依存语法自动机模型的预测进行了比较（Nivre 2008）。

近年来，将人类语言处理与事件相关的脑动力学联系起来，已成为计算神经语言学的重要课题。从 beim Graben 等人（2000；2004）的开创性研究开始，众多研究团队也开展了类似工作（综述参见 Hale 等人（2022））。仅举几例，Hale 等人（2015）将不同的形式语言模型与参与者聆听语音时的 BOLD反应联系起来。类似地，Frank 等人（2015）利用脑电图中的不同事件相关电位成分进行此类统计建模，例如 N400（一种负向偏转，在刺激呈现后约400毫秒出现，可作为词汇语义通达的标志）。Beim Graben 和 Drenhaus（2012）将理解否定极性项（Krifka 1995）过程中随时间整合的事件相关电位，与循环神经网络的和谐观测量（Smolensky 2006）联系起来，从而将形式语言处理器实现为向量符号架构（Gayler 2006，Schlegel 等人 2021）。Rabovsky 和 McRae（2014）以及 Rabovsky 等人（2018）提出了另一种 N400事件相关电位成分神经网络模型，他们将这一标记与神经预测误差和语义更新联系起来，并将其作为观测模型。Brouwer 和 Crocker（2017）、Brouwer 等人（2017）也提出了类似观点，

以及 Brouwer 等人（2021）考察了由分层简单循环网络构成的深度神经网络（Cleeremans 等人 1989，Elman 1990）：基底层实现词汇提取，从而解释N400事件相关电位成分；上层则负责语境整合。该层级的处理失败会由另一种事件相关电位成分 P600 指示（一种正向偏转，在刺激呈现后约600毫秒出现）。由此，他们的神经计算模型实现了此前提出的提取—整合理论（Brouwer 和 Hoeks 2013，Brouwer 等人 2012）。

在 beim Graben 等人（2000；2004；2008）的研究中，他们采用了一种动力系统方法——后来被 beim Graben 和 Potthast（2009）称为认知动力学建模。这是一种三层方法：首先，以符号数据结构和算法作为认知表征与过程的模型。其次，在向量符号架构（Gayler（2006），Schlegel 等人（2021））的框架内，通过填充项—角色绑定和后续的张量积表征（Mizraji（1989）；2020，Smolensky（1990）；2006），将这些符号描述映射为向量表征。第三步，将这些线性结构用作神经网络学习的训练数据。更具体地说，符号串和形式语言处理器可以通过 G¨odel 编码映射为动力自动机（beim Graben 和 Potthast 2014，beim Graben 等人 2000；2004；2008）。最近，Carmantini 等人（2017）展示了如何以精简的方式将这些装置实现为模块化循环神经网络，并将其称为神经自动机（非确定自动机）。[1] Carmantini 等人（2017）还展示了如何将神经自动机用于神经语言学相关性研究。他们实现了一个诊断—修复分析器（Lewis 1998，Lewis 和 Vasishth 2006），用于通过交互式自动机网络处理起初存在歧义的主语关系从句和宾语关系从句（Frisch 等人 2004，Lewis 和 Vasishth 2006）。作为恰当的观测模型，他们利用所得神经网络的平均激活（Amari 1974）作为合成事件相关电位（Barr`es 等人 2013，beim Graben 等人 2008），并在尝试中获得了一个 P600 成分模型。

对于所有这些神经计算模型，符号内容都必须编码为神经激活模式。在向量符号架构中，这一过程涉及将符号映射到填充项向量，并将数据结构中可能的绑定位置映射到角色向量（beim Graben 和 Potthast，2009）。显然，这种编码完全是任意的，至少可以用所选编码的任意一种可能置换来替代。因此，问题在于神经观测模型在多大程度上对任意选定编码的置换保持不变。更为关键的是，我们必须面对这样一个问题：统计相关性分析是否只依赖于某一种特定选择的编码。只有统计模型也对重新编码保持不变时，它们才能被视为可靠的科学研究方法。

本研究旨在针对动力自动机和神经自动机及其底层移位空间这一特定情形，对不变观测模型进行严格的数学处理。本文结构如下。在第2节中，我们介绍动力系统中有关不变量的一般概念和基本定义，并在之后的第2.1节重点讨论神经动力系统这一特殊情形。在第2.2节中，我们将注意力集中于符号动力学。在

[1]请注意，神经自动机是通用计算机（尤其是图灵机）的简约实现。它们不应与深度学习方法框架中出现的神经图灵机混淆（Graves 等，2014年）。

介绍基本记号后，我们将在第2.2.1节介绍有关有根树的工具和事实，并在第2.2.2节介绍有关 G¨odel 编码的工具和事实。在第2.2.3节中，我们将这些概念与柱集联系起来，最终在第2.2.4节描述字符串的不同 G¨odel 化所对应的不变划分。随后，在第2.3节中，我们描述神经自动机的架构，以及如何从单个字符串转到带点序列。最后，在第2.3.1节中，我们描述一个由神经自动机字母表的 G¨odel 重新编码定义的对称群，并基于此前第2.2.4节所描述的不变量，定义一个在该对称性下不变的宏观可观测量。最后，在第3节中，我们将这些结果应用于一个具体实例：该神经自动机经过构造，用于模拟上下文无关文法的解析器。我们证明，给定的宏观可观测量在 G¨odel 重新编码下保持不变，而 Amari 的平均网络活动则不具备这一性质。第4节给出总结性讨论。全文所述事实的全部数学证明均收录于附录中。

2. Invariants in dynamical systems

我们将经典的离散时间确定性动力系统考虑为其最一般形式，即有序对 $\Sigma = (X, \Phi)$，其中 $X \subset \mathbb{R}^n$ 是一个紧致豪斯多夫空间，作为其维数为 $n \in \mathbb{N}$ 的相空间，而 $\Phi : X \to X$ 是一个可逆（通常为非线性）映射（Atmanspacher 和 beim Graben，2007）。系统的流由时间迭代 Φ^t, $t \in \mathbb{Z}$ 生成，即，$t \mapsto \Phi^t$ 是动力学的单参数群，时间为 $t \in \mathbb{Z}$，满足 $\Phi^t \circ \Phi^s = \Phi^{t+s}$，其中 $t, s \in \mathbb{Z}$。相空间 $\mathbf{x} \in X$ 的元素指系统 Σ 的微观描述，因此称为微观状态。设定初始条件 $\mathbf{x}_0 \in X$ 后，系统沿轨迹 $T = \{\mathbf{x}(t) = \Phi^t(\mathbf{x}_0) | \, t \in \mathbb{Z}\}$ 确定性地演化。

有界函数 $f : X \to \mathbb{R}$ 被称为可观测量，其中 $f(\mathbf{x}) \in \mathbb{R}$ 表示在微观态 $\mathbf{x}$ 中的测量结果。函数空间 $B(X) = \{f : X \to \mathbb{R} | \, \|f\| < \infty\}$配备逐点函数加法 $(f + g)(x) = f(x) + g(x)$ 、函数乘法 $(fg)(x) = f(x)g(x)$ 以及数乘 $(\lambda f)(x) = \lambda f(x)$ （对所有 f、$g \in B(X)$、$\lambda \in \mathbb{R}$而言），被称为系统 Σ 的可观测量代数，其范数为$\| \cdot \| : B(X) \to \mathbb{R}_0^+$。将函数空间 $B(X)$ 限制为有界连续函数 $C_0(X)$，即可得到微观可观测量代数，用于描述在相空间的特定区域内能够唯一区分不同微观状态的理想测量。

相比之下，复杂的现实世界动力系统只能测量宏观性质。相应的宏观可观测量属于更大的有界函数代数[2] $B(X)$，通常被定义为所谓平均场的大尺度极限（Hepp 1972、Sewell 2002）。下文将讨论计算神经科学中的宏观平均场可观测量示例。

宏观可观测量的代数 $B(X)$ 包含阶跃函数，尤其是指示函数 χ_A，用于表示在整个 X 上不连续的真子集 $A \subset X$。由于对所有 $\mathbf{x}$、$\mathbf{y} \in A$，$\chi_A(\mathbf{x}) = \chi_A(\mathbf{y})$，微观状态 $\mathbf{x}$ 和 $\mathbf{y}$ 无法通过对 χ_A 的宏观测量加以区分。因此，

[2]事实上，此处需要相对于给定概率测度的本质有界函数代数。要恰当地处理这些概念，必须参考 *algebraic quantum theory*（Sewell 2002）。

Jauch（1964）和 Emch（1964）称它们为宏观等价。[3] 在给定的数学框架中，宏观等价的微观状态类构成一个宏观态（Emch 1964，Jauch 1964，Sewell 2002）。因此，宏观可观测量会将动力系统 Σ 的相空间划分为若干宏观态。

微观可观测量代数 $C_0(X)$ 和宏观可观测量代数 $B(X)$ 分别是带有额外代数乘积的线性空间。作为向量空间，它们允许构造线性同态 $\varphi : B(X) \to B(X)$，这些同样是向量空间。线性同态空间的一个重要子空间是线性自同构空间 Aut($B(X)$)，其中包含可逆线性同态。空间 Aut($B(X)$) 关于函数复合 $(\varphi \circ \eta)(f)$ 还构成一个群，称为代数 $B(X)$ 的自同构群。

接下来，设 G 是一个群，它在动力系统 Σ 的自同构群 Aut($B(X)$) 中具有忠实表示 α。那么，对于 $a \in G$, $\alpha_a \in$Aut($B(X)$)，映射会将可观测量 $f \in B(X)$ 映射到其变换后的 $\alpha_a(f) \in B(X)$，使得对于两个 $a, b \in G$，有 $\alpha_{a*b}(f) = (\alpha_a \circ \alpha_b)(f)$，其中“$*$”表示 G 中的群乘积。群 G 被称为动力系统 Σ 的对称性（Sewell 2002）。此外，如果 G 的表示与 Σ 的动力学可交换，

$$(\alpha_a(f \circ \Phi^t))(\mathbf{x}) = f(\Phi^t(\alpha_a^*(\mathbf{x}))) \tag{1}$$

对所有 $\mathbf{x} \in X$，群 G 称为动力学对称性（Sewell 2002）。在式（1）中，映射 $\alpha_a^* : X \to X$ 是通过以下方式将观测量提升到相空间的结果

$$f \circ \alpha_a^* = \alpha_a(f)\,. \tag{2}$$

例如，考虑宏观可观测量 χ_A，即真子集 $A \subset X$ 的示性函数。再次选择 α_a^*，使得 $\alpha_a^*(\mathbf{x}) \in A$对所有 $\mathbf{x} \in A$都成立，这样便保持了 χ_A 不变：$\chi_A(\alpha_a^*(\mathbf{x})) = \chi_A(\mathbf{x})$。

更一般地，如果满足以下条件，我们就称可观测量 $f \in B(X)$在对称性 G 下是不变量

$$f(\alpha_a^*(\mathbf{x})) = f(\mathbf{x}) \tag{3}$$

对所有 $a \in G$都成立。本研究旨在考察特定神经动力系统中的此类不变量，即动力自动机和神经自动机（Carmantini 等人 2017，beim Graben 等人 2000；2004；2008）。

2.1. 神经动力学。神经动力系统本质上是循环神经网络，由大量的模型神经元（或单元） $n \in \mathbb{N}$组成，这些神经元通过复杂图相互连接（Arbib 1995，Hertz 等人（1991），LeCun 等人（2015），Schmidhuber（2015））。经过适当的归一化后，单元的活动（例如其脉冲率）可以用单位区间 $[0,1] \subset \mathbb{R}$ 中的一个实数表示。于是，整个网络的微观态便成为 n 维超立方体中的一个向量，$\mathbf{x} \in X = [0,1]^n \subset \mathbb{R}^n$。微观观测量是各个坐标轴上的投影算子，

$$f_i(\mathbf{x}) = x_i$$

[3]参见 beim Graben 和 Atmanspacher（2006；2009）所使用的相关概念“认识等价”。

对于 $1 \leq i \leq n$。在离散时间下，网络动力学通常由一个非线性差分方程给出

$$\mathbf{x}(t+1) = \Phi_{\mathbf{W}}(\mathbf{x}(t)) \,. \tag{4}$$

这里，$\mathbf{x}(t) \in X$ 是网络在时间 t 的激活向量（微观态），而$\Phi_{\mathbf{W}}$ 是由突触权重矩阵 $\mathbf{W} \in \mathbb{R}^{n^2}$参数化的非线性映射。通常假设映射 $\Phi_{\mathbf{W}}$ 具有如下形式：

$$\Phi_{\mathbf{W}}(\mathbf{x}) = \mathbf{F}(\mathbf{W} \cdot \mathbf{x}) \,, \tag{5}$$

其中，$\mathbf{F} = (F_i)_{1 \leq i \leq n}$ 是非线性压缩函数：$X \to X$ 作为网络的激活函数。对于 $F_i = \Theta$ （其中 Θ 表示海维赛德阶跃函数），方程（4、5）描述了一个由麦卡洛克–皮茨神经元组成的网络（McCulloch 和 Pitts，1943）。另一种常用的激活函数是逻辑斯蒂函数

$$F_i(x) = \frac{1}{1 + \mathrm{e}^{-x_i}} \,,$$

用于描述放电率模型（例如参见 beim Graben（2008））。将方程（5）替换为映射

$$\Phi_{\mathbf{W}}(\mathbf{x}) = (1 - \Delta t)\mathbf{x} + \Delta t \, \mathbf{F}(\mathbf{W} \cdot \mathbf{x}) \tag{6}$$

可得到一个离散时间漏积分网络（beim Graben 和 Rodrigues，2013；beim Graben 等人，2009；Wilson 和 Cowan，1972）。使用欧拉方法进行数值模拟时，时间步长取为 $\Delta t < 1$。

为了对神经网络仿真与神经生理学实验的实验数据进行相关分析，需要将高维神经激活空间 $X \subset \mathbb{R}^n$ 映射到一个维度低得多的观测空间；该空间由 $p \in \mathbb{N}$ 宏观可观测量 f_k 张成：$X \to \mathbb{R}$ $(1 \leq k \leq p)$。这种投影的标准方法是主成分分析（PCA）（Elman，1991）。如果 PCA 仅限于第一个主轴，那么得到的标量变量可以视为神经网络整体活动的度量。在计算神经语言学领域，beim Graben 等人（2008）利用了 PCA 投影。

另一个重要的标量可观测量（例如，Beim Graben 和 Drenhaus（2012）将其用作神经元观测模型）是 Smolensky 的和谐度（Smolensky，1986）

$$H(t) = \mathbf{x}(t)^{+} \cdot \mathbf{W} \cdot \mathbf{x}(t) \tag{7}$$

其中 $\mathbf{x}^{+}$ 为转置激活状态向量，$\mathbf{W}$ 为上文所述的突触权重矩阵。

Brouwer 等人（2017）提出，实际微观态与其动力学前驱状态之间的“不相似性”，即

$$D(t) = 1 - \frac{\mathbf{x}(t) \cdot \mathbf{x}(t-1)}{\|\mathbf{x}(t)\| \|\mathbf{x}(t-1)\|} \tag{8}$$

可作为一种合适的神经元观测模型。

然而，在本研究中，我们使用 Amari 的平均网络活动（Amari 1974）

$$A(t) = \frac{1}{n} \sum_i x_i(t) \tag{9}$$

作为神经网络的时变“合成事件相关电位”（Barr`es 等人，2013；Carmantini 等人，2017）。

2.2. 符号动力学。符号动力学源自一个离散时间但空间连续的动力系统 Σ，其相空间 X 被划分为有限个 m 两两不相交且完全覆盖该空间的子集 X（Lind 和 Marcus，1995）。因此

$$\mathcal{P} = \{A_k \subset X | A_k \cap A_j = \emptyset \text{ for } k \neq j\,, \quad \bigcup_{k=1}^{m} A_k = X\}\,.$$

这种划分可以由一个值域有限的宏观可观测量诱导。将划分集合 A_k 的索引 k 作为一个特殊符号 s_t，赋给状态 x(*t*)（当 x(*t*) $\in A_k$ 时），系统的轨迹便映射为一个双向无限的符号序列。相应地，动力学的流映射 Φ 可通过左移位 σ 表示为 $\sigma(s_t) = s_{t+1}$。

遵循 beim Graben 等人（2004；2008）以及 Carmantini 等人（2017）的做法，符号指的是有限集合 A 中的一个特殊元素，我们称该集合为字母表。由符号 $w \in \mathrm{A}^l$ 组成的序列称为长度为 l 的词，记作 $l = |w|$。所有可能长度 w 均为有限长度 $|w| \geq 0$ 的词所构成的集合，也称为 A 上的词汇表，记作 A^*（对于 $|w| = 0$，$w = \epsilon$ 表示“空词”）。

2.2.1. 有根树。我们可以将字母表 A 上所有词的集合想象成一棵正则根树，T，其中每个顶点都标有由该字母表构成的一个词，并与之对应。假设 A 对某个 $m \in \mathbb{N}$有 m 个字母。也就是说，$\mathrm{A} = \{a_1, \ldots, a_m\}$。那么，树 T按如下方式归纳构造：

(i) 树的根是一个标有空词 ϵ的顶点。(ii) 假设我们已经构造了第 n步的顶点，那么按如下方式构造第 $n+1$ 步的顶点。假设第 n步有 k 个顶点，分别标有词语 w_1、……、w_k。那么

- 对每个 $i = 1$、……、k 以及每个 $a_j \in$A，我们添加一个标记为 $w_i a_j$的新顶点。
- 对每个 $i = 1$、……、k 以及 $j = 1$、……、m，我们添加一条从 w_i 到$w_i a_j$的边。

此构造生成一棵正则根树。按照前述构造，通常第一步是将根放置在最上方的顶点处。随后，根通过边连接起来，其中每条边都与一个长度为 1 的词相关联，也就是说，与 A 中的每个符号相关联。接着迭代地，将这些标有 A 中字母的边分别连接到以该字母开头的任意长度为 2 的词，依此类推。假设 $\mathrm{A} = \{a_1, \ldots, a_m\}$，此构造便得到一棵如图 1 所示的无限树。

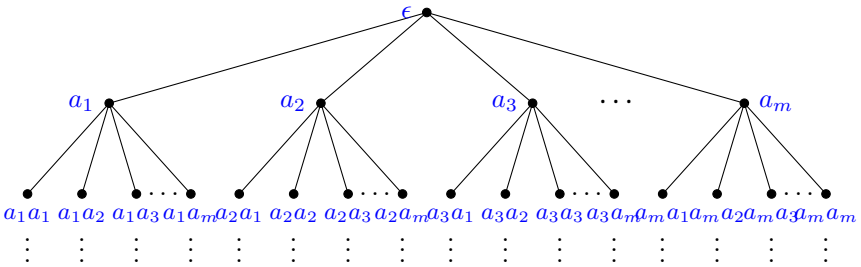

FIGURE 1. 词汇表 A^* 作为一棵根树。

树的每个顶点都对应一个由字母表 A 中符号组成的词。也就是说，树的顶点集合为 A^*。另一方面，每条从根出发的无限射线都对应一个由 A 中符号组成的无限序列，并对应于树的边界。我们将此边界记作 ∂T，如前所述，将其视为一个集合时，它等于 $\mathrm{A}^{\mathbb{N}}$。

树的构造是唯一的，但取决于我们对 A 中符号所选定的具体排序。因此，原则上，如果 $\gamma\colon \mathrm{A} \to \{0,\dots,m-1\}$ 是字母表的一种特定排序（即一个双射），其中元素 a 在 $\gamma(a) = i-1$ 时记作 a_i，那么这棵树应记作 T_γ，因为它取决于字母表的这种特定排序。

我们用 T 表示字母表 $\{0,1,\dots m-1\}$ 上的正则根树，其自然序由 $\mathbb{N}$ 诱导（见图 2）。

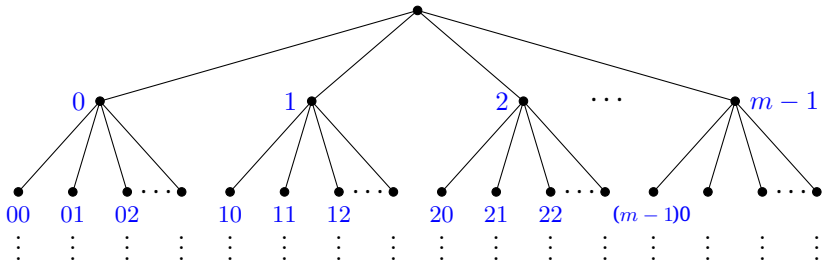

FIGURE 2. 字母表 $\{0,1,\dots,m-1\}$ 上的正则根树 T。

今后，我们将用 M 表示字母表 $\{0,\dots,m-1\}$；如前所述，将用 T 表示在 $\mathbb{N}$ 的通常排序下与字母表 M 对应的树。

当我们说该构造在符号重新排序意义下是唯一的时，我们的意思是，这两棵树作为图是同构的；其中，图的同构是保持关联关系的顶点之间的双射。事实上，对于任意双射 $\gamma\colon \mathrm{A} \to \mathrm{M}$，树 T_γ 作为图都与 T 同构。

引理 2.1。设 $\gamma : \mathrm{A} \to \mathrm{M}$ 是字母表 A 的一个排序。那么 T_γ 和 T 是同构的。

由于同构具有传递性，该引理表明，对于任意两个基数相同的字母表 A_1 和 A_2，以及这两个字母表的任意两个排序 γ_1 和 γ_2，它们对应的树 T_{γ_1} 和 T_{γ_2} 作为图是同构的。

2.2.2. 哥德尔编码。给定 $\mathrm{A}^{\mathbb{N}}$，由含有 $|\mathrm{A}| = m$ 个符号的字母表 A 上的单侧无限序列构成的空间，以及该空间中的一个序列 $s = a_1 a_2 \dots$，其中 a_k 是 s 中的第 k 个符号，并给定一个排序 $\gamma\colon \mathrm{A} \to \{0,\dots,m-1\}$，那么哥德尔化是从 $\mathrm{A}^{\mathbb{N}}$ 到 $[0,1] \subset \mathbb{R}$ 的映射，定义如下：

$$\psi_\gamma(s) := \sum_{k=1}^{\infty} \gamma(a_k) m^{-k}. \tag{10}$$

根据引理 2.1，我们知道，对于由 γ 诱导的 A 的每一种 G¨odel 化，T_γ 与 T 之间都存在图同构。由于排序的选择

字母表（换句话说， γ 的选择）是任意的，并会产生不同的 G¨odel 编码，因此我们有兴趣寻找这些不同编码的不变量。

可以用以下方式定义树的边界上的度量：给定树上的任意两条无限射线 $p = a_1a_2a_3\ldots$ 和 $q = b_1b_2b_3\ldots$，我们定义

$$d(p,q) = \begin{cases} 0 & \text{if } p = q \\ m^{-n}, & \text{if } a_i = b_i \text{ for } i = 1,\ldots,n \text{ and } a_{n+1} \neq b_{n+1} \\ 1 & \text{if } a_1 \neq b_1. \end{cases}$$

这定义了边界上的一种超度量，即满足三角不等式更强形式的度量，具体来说：

$$d(p,q) \leq \max\{d(p,r), d(r,q)\}.$$

当我们使用 G¨odel 编码对无限字符串进行编码时，会将这种超度量下彼此接近的射线映射到通常度量下区间$[0,1]$ 中彼此接近的点。

引理 2.2。设 $p = a_1a_2a_3\ldots$ 和 $q = b_1b_2b_3\ldots$ 是 A 上的两个无限字符串。那么对于任意 G¨odel 编码 ψ，我们有

$$d(p,q) \leq \frac{1}{m^n} \iff \exists k \in \{0,\ldots,m^n-1\}, \psi(p), \psi(q) \in \left[\frac{k}{m^n}, \frac{k+1}{m^n}\right).$$

请注意，该引理并不意味着区间 $[0,1]$ 上彼此接近（按照通常度量）的点来自树上彼此接近的射线。例如，如果字母表有 3 个字母，那么对于任意 $\epsilon > 0$，点 $1/3 - \epsilon$ 和 $1/3$ 可以任意接近，但它们在树上的距离为 0。事实上，它对每个 $n \in \mathbb{N}$ 都给出了区间的一种划分，并且，如果表示无限字符串的两个点根据相应 n 的划分位于同一区间内，那么它们来自两条至少共享长度为 n的公共前缀的射线。

2.2.3. 柱集。在符号动力学中，柱集（McMillan 1953）是字母表 A 上无限序列空间 $\mathbf{A}^{\mathbb{N}}$ 的一个子集，其中的序列在长度为 $l \in \mathbb{N}$ 的特定构造块上相同。因此，设 $w = \mathbf{A}^*$ 是一个长度为 l 的有限词 $a_1a_2\ldots a_l$，我们定义柱集

$$[w] = [a_1a_2\ldots a_l] = \{s \in \mathbf{A}^{\mathbb{N}} \mid s_k = a_k, \quad k = 1,\ldots,l\}. \tag{11}$$

我们也可以在图 3 所示的树上看到柱集。事实上，对于树上的每一层（这里的“层”指对应于特定固定长度词语的顶点），我们都会得到区间 $[0,1]$ 的一个划分。该层每个顶点下方的顶点都落在划分中对应的区间内。因此，从根树的视角来看，柱集对应于从该顶点向下延伸的整棵树。具体而言，词 $w \in \mathbf{A}^*$ 的柱集 $[w]$，就是从标记为 w 的顶点向下延伸的子树。

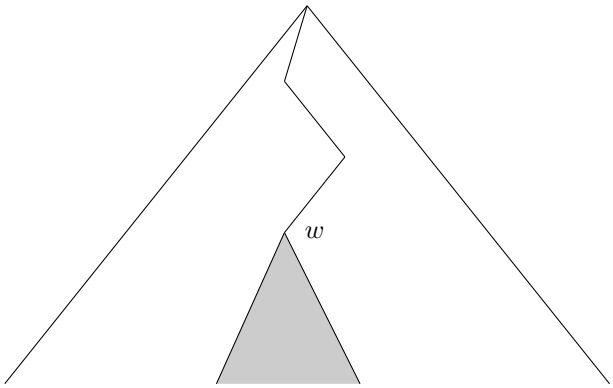

FIGURE 3. 对应于树上所见的 w 的柱集。

两个不同的 G¨odel 编码 ψ, φ 只能在其赋值 γ_1, γ_2 方面有所不同：$\mathrm{A} \to \{0, \ldots, m-1\}$。因此，如果一个置换 $\pi \in S_m$（其中 S_m 是对称群）满足以下条件，我们就称之为 G¨odel 重新编码：

$$\pi \circ \gamma_1 = \gamma_2 \, .$$

2.2.4. 不变量。我们研究的最终目标是找出 G¨odel 重新编码下的不变量。注意，按照引理 2.1 的记号，$g_{\gamma_1} : T_{\gamma_1} \to T$ 和 $g_{\gamma_2} : T_{\gamma_2} \to T$ 是两个图同构。事实上，它们诱导了 T 的一个图自同构，$g_\pi = g_{\gamma_2} \circ g_{\gamma_1}^{-1} : T \to T$。这个自同构将由 γ_1 编码的顶点映射到由 γ_2 编码的顶点。

正如引理 2.2 所示，G¨odel 重新编码在置换顶点后会保持柱集的大小。然而，字母表的排序方式及其对树的射线的置换方式，比单纯保持柱集大小的限制更严格。事实上，在重新排序的作用下，每个顶点只能映射到某些特定顶点，而不能映射到其他顶点。以下最核心的定义对此进行了刻画。

定义2.3. 设 $w = a_{i_1} a_{i_2} \ldots a_{i_l} \in \mathbf{M}^l$ 是一个长度为 l、按 γ 排序的字符串。我们定义整数集合 $\{1, 2, \ldots, l\}$ 的一个划分，

$$\mathcal{P}_w = \{\{j_1, \ldots, j_k\} \subset \mathbb{N} | a_{i_{j_1}} = \cdots = a_{i_{j_k}}\} \, . \tag{12}$$

对于任意词 $w \in \mathbf{M}^*$，我们称 $\mathcal{P}_w$ 为 w 的相等模式。

借助上述形式体系，我们现在可以阐述本研究的首个主要发现如下。

定理 2.4. 对于任意其他顶点 $u \in T$，存在一种 G¨odel 重新编码 π，使得 $g_\pi(w) = u$ 当且仅当

$$u \in \mathbf{M}^l \tag{13}$$

$$P_w = P_u \, . \tag{14}$$

定理 2.4指出，每个顶点都可以映射到具有相同相等模式的任意顶点，且不能映射到其他顶点。

例2.5。若 A $=\{a,b,c\}$，并且我们考虑 $w=aaabcabc\in \mathrm{A}^8$。那么我们有 $P_w=\{\{1,2,3,6\},\{4,7\},\{5,8\}\}$，由此得到所有 w 可以映射到的可能词语。可能性列表如下：

$$\begin{array}{ccc} bbbacbac & cccbacba & aaacbacb \\ bbbcabca & cccabcab & aaabcabc \end{array}$$

因此，在 $3^8=6561$ 中，我们只有 6 个可能的顶点。当然，随着我们沿着树向下深入，这一比例会逐渐降低。

就映射到 $[0,1]$ 区间的 G¨odelization 而言，我们通过一个例子来说明其含义。假设 $m=3$ 和 $l=3$，例如。那么，在图 4 中，某种颜色的柱集只能通过重新编码映射到相同颜色的柱集，不能映射到其他地方。

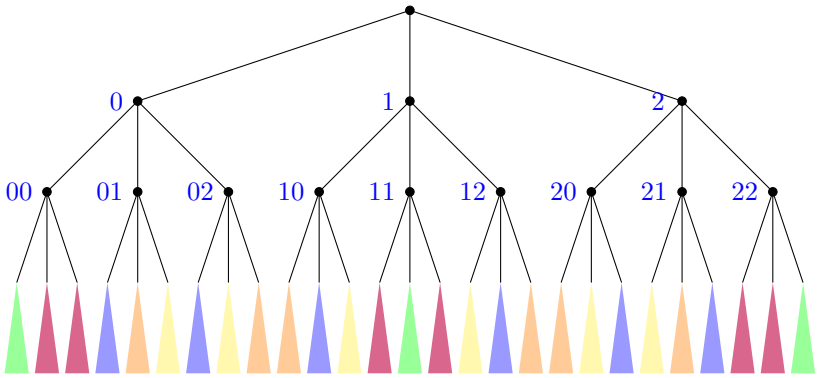

FIGURE 4. 根据相等模式对柱集进行的不变划分。

图5 展示了区间 $[0,1]$ 对应的划分：每种颜色的区间只能通过不同的赋值映射映射到相同颜色的其他区间，不能映射到其他地方。



FIGURE 5. 经过 G¨odelization 后对区间 $[0,1]$ 进行的不变划分。

2.3. 神经自动机。按照 beim Graben 等人（2004；2008）以及 Carman-tini 等人（2017）的定义，字母表 A 上的带点序列 $s\in \mathrm{A}^{\mathbb{Z}}$ 是由符号“$s=\ldots\ a_{-2}\ a_{-1}\ .\ a_0\ a_1\ a_2\ \ldots$”构成的双向无限序列，其中 $a_i\in \mathrm{A}$ 对所有指标 $i\in\mathbb{Z}$ 均成立。这里，点号“.”仅用作助记符号，表示指标 0 位于其右侧。

Carmantini 等人（2017）将点解释为一种元符号，可与两个词 $v_1,v_2\in \mathrm{A}^*$ through $v=v_1.v_2$ ˆ连接。令 A^* 表示这些带点词的集合。此外，令 $\mathbb{Z}^-=\{i\mid i<0,\ i\in\mathbb{Z}\}$ 和 $\mathbb{Z}^+=\{i\mid i\geq 0,\ i\in\mathbb{Z}\}$ 分别表示负指标和非负指标的集合。这样，我们就可以重新引入带点序列的概念，定义如下。令 $s\in \mathrm{A}^{\mathbb{Z}}$ 为一个双无限符号序列，使得 $s=w_\alpha v w_\beta$，其中 $v\in \mathrm{A}^*$ 是一个带点词 $v=v_1.v_2$

并且 $w_\alpha v_1 \in \mathrm{A}^{\mathbb{Z}^-}$ 和 $v_2 w_\beta \in \mathrm{A}^{\mathbb{Z}^+}$。根据这一定义， s 的指标继承自带点词 v，因此没有明确指定。

多功能移位（VS）由 Carmantini 等人（2017）定义为一对 $M_{VS} =$(A$^{\mathbb{Z}}$, Ω)，其中 A$^{\mathbb{Z}}$ 是带点序列的空间，而 Ω : A$^{\mathbb{Z}}$ → A$^{\mathbb{Z}}$ 定义如下：

$$\Omega(s) = \sigma^{F(s)}(s \oplus G(s)) \tag{15}$$

with

$$\begin{aligned} F &: \mathbf{A}^{\mathbb{Z}} \to \mathbb{Z} \\ \oplus &: \mathbf{A}^{\mathbb{Z}} \times \mathbf{A}^{\mathbb{Z}} \to \mathbf{A}^{\mathbb{Z}} \\ G &: \mathbf{A}^{\mathbb{Z}} \to \mathbf{A}^{\mathbb{Z}}, \end{aligned} \tag{16}$$

ˆ 其中，算子“⊕”将 s 中的带点词 $v_1.v_2 \in$A* 替换为由 G指定的新带点词 $\hat{}v_1.\hat{}v_2 \in$A*，而 $F(s) = F|_{\hat{}\mathbf{A}^*}(v_1.v_2)$则确定移位步数，这与 Moore 广义移位中的情况相同（ Carmantini 等人，2017）。

非线性动力学自动机（NDA）是一个三元组 $M_{NDA} = (Y, \mathcal{P}, \Phi)$，其中$\mathcal{P}$ 是单位正方形 $Y = [0,1]^2 \subset \mathbb{R}^2$的一个矩形划分，即

$$\mathcal{P} = \{D^{(i,j)} \subset Y |\ 1 \le i \le m,\ 1 \le j \le n,\ m, n \in \mathbb{N}\}, \tag{17}$$

因此，每个单元定义为 $D^{(i,j)} = I_i \times J_j$，其中对于每个双指标 (i, j)， $I_i, J_j \subset [0,1]$ 是实区间；若 $(i, j) \neq (k, l)$，则 $D^{(i,j)} \cap D^{(k,l)} = \varnothing$，并且$\bigcup_{i,j} D^{(i,j)} = Y$。数对 $(Y,$ Φ) 是一个离散时间动力系统，其相空间为 Y，流为 Φ: $Y \to Y$ 是一个分段仿射线性映射，使得 $\Phi_{|D^{(i,j)}} := \Phi^{(i,j)}$，其中$\Phi^{(i,j)}$ 具有以下形式：

$$\Phi^{(i,j)}(\mathbf{y}) = \begin{pmatrix} a_1^{(i,j)} \\ a_2^{(i,j)} \end{pmatrix} + \begin{pmatrix} \lambda_1^{(i,j)} & 0 \\ 0 & \lambda_2^{(i,j)} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \tag{18}$$

其状态向量为 y $= (y_1, y_2)$。Carmantini 等人（2017）已证明，利用 G¨ odelization，任何多功能移位都可以映射为非线性动力学自动机。因此，可以在单位正方形 Y上再现多功能移位的活动。为此，划分（17）由所谓的依赖域（ DoD）给出。依赖域是一个二元组（l, r） $\in \mathbb{N} \times \mathbb{N}$，它定义了带点序列中点左侧和右侧字符串的长度；这些字符串与多功能移位作用于相空间有关。多功能移位的动力学完全取决于每次迭代时依赖域上的字符串形式。于是，若依赖域为（l, r），且字母表 A 的大小为 m，则单位正方形的划分由 m^r 个区间构成，这些区间位于 y_1 轴上；另有 m^l 个区间位于 y_2 轴上。它们对应于非确定自动机（NDA）根据多功能移位定义的单元。最后，神经自动机（NA）是通过模块化循环神经网络实现 NDA 的一种方式（Carmantini 等人，2017）。

该神经自动机包含一个相空间 $X = [0,1]^n$，其中底层 NDA 的二维子空间 $Y = [0,1]^2$ 仅由属于机器配置层（MCL）的两个神经元张成。其余部分$X \setminus Y$ 则由分支选择层（BSL）和线性变换层（LTL）的神经元张成，这两层共同实现分段仿射映射（18）。由于 NDA 是根据多功能移位定义的，划分中的每个矩形均由 DoD 给出，而 NDA 在每个矩形上的作用取决于所选字母表 A 的特定 G¨odel 编码。我们

关注的是此类自动机在字母表不同 G¨odel 编码下的不变宏观可观测量。

由于我们现在关注字母表 A 上的带点序列，因此，我们将不再像图 5 中那样有区间 $[0,1]$ 的不变划分，而是会有单位正方形 $Y=[0,1]^2$ 的不变划分。也就是说，我们会得到一个由矩形组成的划分，其中每个矩形表示机器在动力学的某一步可能处于的位置，也可能表示机器不会处于的位置。划分中的每种颜色都表示：在不同的 G¨odel 编码下，具有特定左右长度的某个带点序列可能出现的所有位置。

例如，假设我们的字母表在带点序列两侧各有 $m=3$ 个字母，并且我们考察点左侧长度为 $l=2$ 、点右侧长度为 $r=3$ 的词语，那么划分将如图 6 所示。

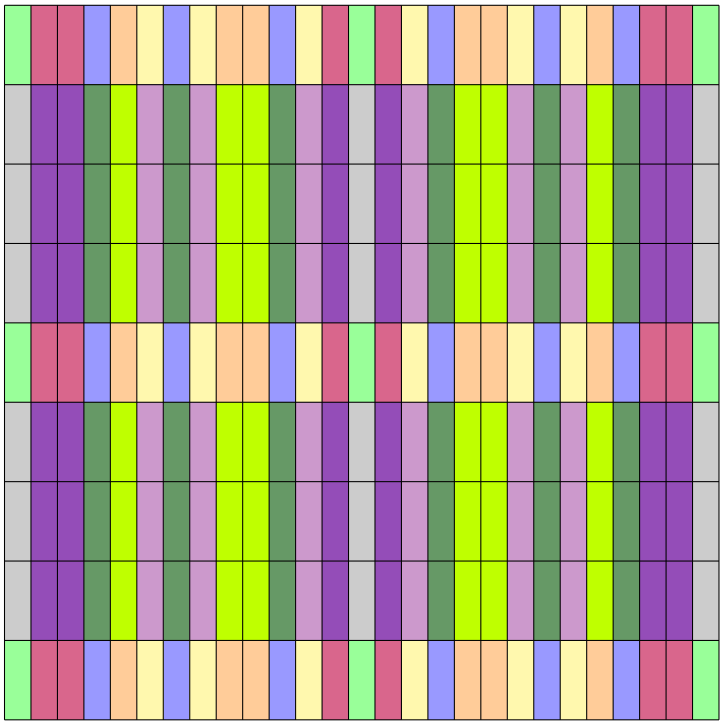

FIGURE 6. 每个小正方形都对应于由长度为 (2, 3) 的带点序列所给出的划分中的一个正方形。颜色相同的正方形具有相同的相等模式，因此，在字母表的不同 G¨odel 编码下，它们可以相互映射。

假设我们正在考虑长度为 (l,r) 的带点序列的不变划分，这意味着左侧长度为 l ，右侧长度为 r。那么我们知道，正方形 Y 的划分由 $E^{(i,j)}=\left[\frac{i}{m^l},\frac{i+1}{m^l}\right)\times\left[\frac{j}{m^r},\frac{j+1}{m^r}\right)$ 给出。矩形的每个左角都对应于大小为 (l,r) 的带点序列的哥德尔化位置。每个点 $(y_1,y_2)=(\frac{i}{m^l},\frac{j}{m^l})$ 的坐标都具有以 m 为底的唯一展开。也就是说，存在一些

c_1，...， c_l，其中 $0 \le c_i \le m_1$，使得

$$y_1 = \frac{i}{m^l} = \frac{c_1}{m} + \frac{c_2}{m^2} + \cdots + \frac{c_l}{m^l}. \tag{19}$$

这些 $\{c_1, \ldots, c_l\}$ 也以定义2.3中给出的相同方式定义了 $\{1, \ldots, l\}$ 的一个划分。因此 $\{d_1, \ldots, d_k\} \in \mathcal{P}_x \iff c_{j_1} = \cdots = c_{j_k}$。此过程同样适用于 y_2 坐标。因此，定义该划分中一个不变片段的角，将是那些共享 $\{1, \ldots, l\} \times \{1, \ldots, r\}$相同划分的角。换句话说，我们可以通过在基底 m 上展开 y_1 和y_2，并改变0，...， $m-1$ 在展开式中的出现顺序，来得到与给定y相关的角。

例如，如果 $m = 3$ 和 （l， r） = （2， 3），我们就有 $3^2 \cdot 3^3 = 3^5$ 个矩形。现在以矩形$\left[\frac{6}{3^2}, \frac{7}{3^2}\right) \times \left[\frac{10}{3^3}, \frac{11}{3^3}\right)$ 为例，求出它的不变划分。首先，我们将其分解为

$$y_1 = \frac{6}{3^2} = \frac{2}{3} + \frac{0}{9} \text{ and } y_2 = \frac{10}{3^3} = \frac{1}{3} + \frac{0}{9} + \frac{1}{27}.$$

因此，同一不变划分中的矩形必须具有如下形式： $E^{(i,j)} = \left[\frac{i}{m^l}, \frac{i+1}{m^l}\right) \times \left[\frac{j}{m^r}, \frac{j+1}{m^r}\right)$ 与 $\frac{i}{3^2} = \frac{a}{3} + \frac{b}{9}$，以及 $\frac{j}{3^3} = \frac{c}{3} + \frac{b}{9} + \frac{c}{27}$与 a、 b、 $c \in \{0, 1, 2\}$，且[4]不同。这给出了以下矩形

$$\left[\frac{1}{9}, \frac{2}{9}\right) \times \left[\frac{23}{27}, \frac{24}{27}\right) \qquad \left[\frac{3}{9}, \frac{4}{9}\right) \times \left[\frac{20}{27}, \frac{21}{27}\right)$$

$$\left[\frac{2}{9}, \frac{3}{9}\right) \times \left[\frac{16}{27}, \frac{17}{27}\right) \qquad \left[\frac{6}{9}, \frac{7}{9}\right) \times \left[\frac{10}{27}, \frac{11}{27}\right)$$

$$\left[\frac{7}{9}, \frac{8}{9}\right) \times \left[\frac{3}{27}, \frac{4}{27}\right) \qquad \left[\frac{5}{9}, \frac{6}{9}\right) \times \left[\frac{6}{27}, \frac{7}{27}\right)$$

通过这种方式，我们可以根据相等模式构造单位正方形的划分。

2.3.1. 不变可观测量。现在，我们的目标是为神经自动机定义一个可观测量 $f \in B(X)$，其含义与第 2 节中的相同。也就是说， $f\colon\ X \to \mathbb{R}$ 应满足公式（3），其中映射 α_π^* 对应于由字母表的 G¨odel 重新编码所诱导的对称性。这里， π 表示从一种 G¨odel 编码转换到另一种编码所需的字母表置换，稍后将对此作出说明。

请注意，在前面的讨论中，我们假设已知所编码字符串的长度。然而，实际情况并非如此，而且这可能

[4]这里我们假设字母表 A 和 G¨odel 编码在点的两侧相同，否则我们会有更大的自由度，并得到更多区间，但这一过程无论如何都行得通。

会导致问题，因为字符串的长度在每次迭代时都可能变化。例如，对于字母表 $\{a, b\}$，如果在某种 G¨odel 编码γ 下，符号 a 映射为 0，而符号 b 映射为 1，那么一旦我们假设字符串长度为 r，对于 $r \in \mathbb{N} \cup \{0\}$，数值 $x = 1/2 \in [0, 1]$ 就会对应于词 $w_r = ba^{r-1}a$。然而，如果我们不知道所编码字符串的长度，那么每个 w_r 在将 b 映射为 0、将 a 映射为 1 的 G¨odel 编码下，都会有不同的 G¨odel 数，即 $\sum_{k=2}^{r-1} 1/2^k$。因此，在 G¨odel 重新编码中，用数字 0 编码符号会使某些字符串无法区分，因为没有符号会被解释为含有任意多个以 0 编码的符号。只需向字母表中添加一个符号 $\sqcup$，并将其解释为空白符号，就能轻松避免这个问题；任何 G¨odel 编码都会强制将其编码为 0。

假设我们有一个由多功能移位定义的 NDA，并且满足以下条件：字母表 A 中添加了表示空白符号的空白符号 $\sqcup$，且在任何 G¨odel 编码[5]下，该符号都映射为 0。我们假设添加空白符号后，A 有 m 个符号（也就是说，添加前有 $m-1$ 个符号）。那么对于任意一对 $(r, l) \in \mathbb{N} \times \mathbb{N}$，我们可以将单位正方形 Y划分为以下矩形划分：

$$\mathcal{R} = \{E^{(i,j)} \subset Y |\ 1 \le i \le m^r,\ 1 \le j \le m^l\}. \tag{20}$$

接下来，我们将 NDA 的相空间划分扩展到整个神经自动机的相空间。NDA 的相空间等于较大 NA 的机器配置层子空间。只需定义另一个划分，即可轻松实现这一点：

$$\mathcal{Q} = \{E^{(i,j)} \times [0, 1]^{n-2} \subset X |\ 1 \le i \le m^r,\ 1 \le j \le m^l\}. \tag{21}$$

现在，对于每个左下角 $(x_1^{(i,j)}, x_2^{(i,j)}) \in E^{(i,j)}$，我们确定它们的相等模式$\mathcal{P}_{ij}$，并假设置换只作用于符号 $\{2, \ldots, m\}$（因为在任何编码下，第一个符号都必须映射为 0）。

设 $\{\mathcal{P}_{ij_1}, \ldots, \mathcal{P}_{ij_s}\}$ 是所有不同的、出现过的相等模式，并定义示性函数 χ_k：$X \to \{0, 1\}$ 为

$$\chi_k(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in E^{(i,j)} \times [0, 1]^{n-2} \text{ and } \mathcal{P}_{ij} = \mathcal{P}_{ij_k} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

对于 $\mathbf{x} \in X$。然后，我们可以选择 c_1、.....、$c_s \in \mathbb{R}$ 作为 s 个不同的实数，并将宏观可观测量 f: $X \to \mathbb{R}$ 定义为阶跃函数

$$f(\mathbf{x}) = \sum_{k=1}^{s} c_k \chi_k(\mathbf{x})\,. \tag{23}$$

显然，我们有 $f \in B(X)$。

我们的目标是证明，该可观测量在由式（18）中的神经自动机给出的动力系统（X， Φ）的对称群$S_{m-1} \times S_{m-1}$ 作用下是不变量

[5]为了简化问题，我们假设栈上的字母表与输入符号使用相同的字母表。如果有需要，总可以取这两个字母表的并集来作此假设。

其中， S_{m-1} 表示由 $m-1$ 个元素构成的对称群。首先，我们必须证明 $S_{m-1} \times S_{m-1}$ 是该神经自动机的对称性。

在进行此操作之前，我们先定义一个辅助映射。令 $\pi = (\pi_1, \pi_2) \in S_m \times S_m$ 为直积中固定 1 的任意元素（在集合 $\{1, 2, \ldots, m\}$ 上，其中 S_m 作用）。注意， S_m 中固定第一个元素的元素构成了 S_m 的一个子群，该子群同构于 S_{m-1}。现在令 x 为 X 中的任意一点。考虑 $\mathbf{y} = (y_1, y_2)$，即 x 的前两个坐标，它们由非确定自动机（NA）的机器配置层的激活给出。随后，我们可以确定它位于划分 $\mathcal{R}$ 的哪个区间中；设该区间为 $(y_1, y_2) \in E^{(i,j)} = \left[\frac{i}{m^l}, \frac{i+1}{m^l}\right) \times \left[\frac{j}{m^r}, \frac{j+1}{m^r}\right)$。因此，我们可以计算每个角点在基 m 上的展开，并将所得系数视为字母表 $\mathbf{M} = \{0, 1, \ldots, m-1\}$ 上的词，记为 $c_1 \ldots c_l \in \mathbf{M}^l$ 和 d_1, ..., $d_r \in \mathbf{M}^r$。然后，我们计算 $g_{\pi_1}(c_1 \ldots c_l)$ 和 $g_{\pi_2}(d_1 \ldots d_r)$，并通过规范的 G¨odel 编码（即由 M 上的恒等映射给出的编码）对这些词进行编码。这样，我们便得到划分后的相空间中某个矩形的一个新角点，记为 $E^{(i',j')} = \left[\frac{i'}{m^l}, \frac{i'+1}{m^l}\right) \times \left[\frac{j'}{m^r}, \frac{j'+1}{m^r}\right)$。现在定义映射 $\rho_\pi \colon Y \to Y$ by $\rho_\pi(y_1, y_2) = \left(y_1 + \frac{i'-i}{m^l}, y_2 + \frac{j'-j}{m^r}\right)$。显然，该映射可以扩展为从 X 到 X 的映射，在其余坐标上取恒等映射。略去记号上的区别，我们也将 ρ_π 称为这个映射。非正式地说，映射 ρ_π 按照 g_{π_1} 和 g_{π_2} 对表示角点的词的作用，刚性地置换划分 $\mathcal{R}$ 上的正方形。

现在，我们可以如下定义 $\alpha_\pi \colon B(X) \to B(X)$。对于任意 $f \in B(X)$，我们定义

$$\alpha_\pi(f)(\mathbf{x}) = f(\rho_\pi(\mathbf{x}))\,. \tag{24}$$

不难验证，若 π、$\gamma \in S_{m-1} \times S_{m-1}$ 是两个群元素，则 $\alpha_{\gamma\pi}(f) = (\alpha_\gamma \circ \alpha_\pi)(f)$，因此 $S_{m-1} \times S_{m-1}$ 是该系统的对称性。

由此，我们最终得到主要结果。

定理 2.6。设 $f \in B(X)$ 是神经自动机状态空间上的一个宏观可观测量，如式（23）所定义。那么， f 在自动机符号字母表的 G¨odel 重新编码所构成的对称群 $S_{m-1} \times S_{m-1}$ 下是不变的。

值得一提的是，该过程会产生无穷多个不同的不变观测量。事实上，任意选择 $(r, l) \in \mathbb{N} \times \mathbb{N}$ 都会得到一个更细的不变划分，并相应地得到一个更精确的可观测量。

3. Neurolinguistic application

作为一个有启发性的例子，我们考察一种句法语言处理的玩具模型；这种模型常用于计算心理语言学和计算神经语言学（Arbib 和 Caplan 1979年，Crocker 1996年，beim Graben 和 Drenhaus 2012年，Hale 等人 2022年，Lewis 2003年）。

为了处理 beim Graben 和 Potthast（2014年）在例 3.1 中给出的句子，语言学家通常会从短语结构树中推导出上下文无关文法（CFG）（Hopcroft 和 Ullman 1979年）。

例 3.1. 狗追赶猫

在我们的例子中，CFG 由重写规则组成

$$\texttt{S} \to \texttt{NP}\ \ \texttt{VP} \tag{25}$$
$$\texttt{VP} \to \texttt{V}\ \ \texttt{NP} \tag{26}$$
$$\texttt{NP} \to \texttt{the}\ \ \texttt{dog} \tag{27}$$
$$\texttt{V} \to \texttt{chased} \tag{28}$$
$$\texttt{NP} \to \texttt{the}\ \ \texttt{cat} \tag{29}$$

其中，左侧始终给出一个非终结符号，该符号将在右侧扩展为由非终结符号和终结符号组成的字符串。省略词汇规则（27–29）后，我们将分别表示“名词短语”和“动词”的符号 NP、V 视为终结符号，并将符号 S（“句子”）和 VP（“动词短语”）视为非终结符号。

然后，通过一个简单的自顶向下识别器（Hopcroft 和 Ullman 1979年）处理该文法的多功能移位定义如下：

$$\begin{aligned} \texttt{S}.a &\mapsto \texttt{VP}\ \texttt{NP}.a \\ \texttt{VP}.a &\mapsto \texttt{NP}\ \texttt{V}.a \\ a.a &\mapsto \epsilon.\epsilon \end{aligned} \tag{30}$$

此时，纸带左侧称为“栈”，右侧称为“输入”。在（30）中，a 表示任意输入符号。请注意，点左侧的栈中符号顺序是反向的。（30）中的前两个操作是根据 CFG 规则进行预测，而最后一个操作则是将后续输入与已经预测的材料进行附接。

然后，该机器按照 beim Graben 和 Potthast（2014年）的表 1 所示，对格式良好的句子 NP V NP 进行解析。我们在此重现该表，作为表 1。

TABLE 1. 多功能移位处理例 3.1 中格式良好的字符串 NP V NP 时的状态转移序列。操作说明如下：“预测 (X)”表示根据上下文无关文法的规则 (X) 进行预测；attach 表示从栈和输入中同时消去已成功预测的终结符；“接受”表示接受该字符串，认为其格式良好。

time	状态	操作
0	S . NP V NP	预测 (25)
1	VP NP . NP V NP	附加
2	动词短语 . 动词 名词短语	预测 (26)
3	名词短语 动词 . 动词 名词短语	附加
4	名词短语。名词短语	附加
5	ϵ。 ϵ	接受

一旦得到多功能移位，就可以生成一个模拟它的非确定自动机。为此，我们选择了符号的一种特定哥德尔编码。假设我们选择了以下两种哥德尔化方式 $\gamma = (\gamma_1, \gamma_2)$ 和 $\delta = (\delta_1, \delta_2)$，它们分别定义在

by

$$\begin{aligned}\gamma_1 : \{\sqcup, \texttt{NP}, \texttt{V}\} &\to \{0,1,2\}\\ \sqcup &\mapsto 0\\ \texttt{NP} &\mapsto 1\\ \texttt{V} &\mapsto 2\\ \gamma_2 : \{\sqcup, \texttt{NP}, \texttt{V}, \texttt{VP}, \texttt{S}\} &\to \{0,1,2,3,4\}\\ \sqcup &\mapsto 0\\ \texttt{NP} &\mapsto 1\\ \texttt{V} &\mapsto 2\\ \texttt{VP} &\mapsto 3\\ \texttt{S} &\mapsto 4\end{aligned}$$

一方面，以及由

$$\begin{aligned}\delta_1 : \{\sqcup, \texttt{NP}, \texttt{V}\} &\to \{0,1,2\}\\ \sqcup &\mapsto 0\\ \texttt{NP} &\mapsto 2\\ \texttt{V} &\mapsto 1\\ \delta_2 : \{\sqcup, \texttt{NP}, \texttt{V}, \texttt{VP}, \texttt{S}\} &\to \{0,1,2,3,4\}\\ \sqcup &\mapsto 0\\ \texttt{NP} &\mapsto 4\\ \texttt{V} &\mapsto 3\\ \texttt{VP} &\mapsto 1\\ \texttt{S} &\mapsto 2\end{aligned}$$

另一方面。在选择 $(l, r) = (2, 3)$ 和随机选取 c_i-s 后，按照式（23）定义阶跃函数 $f\colon X \to \mathbb{R}$ 。该神经自动机由 $n = 72$ 个神经元组成，即其相空间由超立方体 $X = [0,1]^{72}$ 给出。以这两种编码运行神经网络，并在每次迭代 $i = 1$、.....、6 时计算阶跃函数 f ，我们从图7中看到， f 确实在哥德尔重新编码下保持不变量。

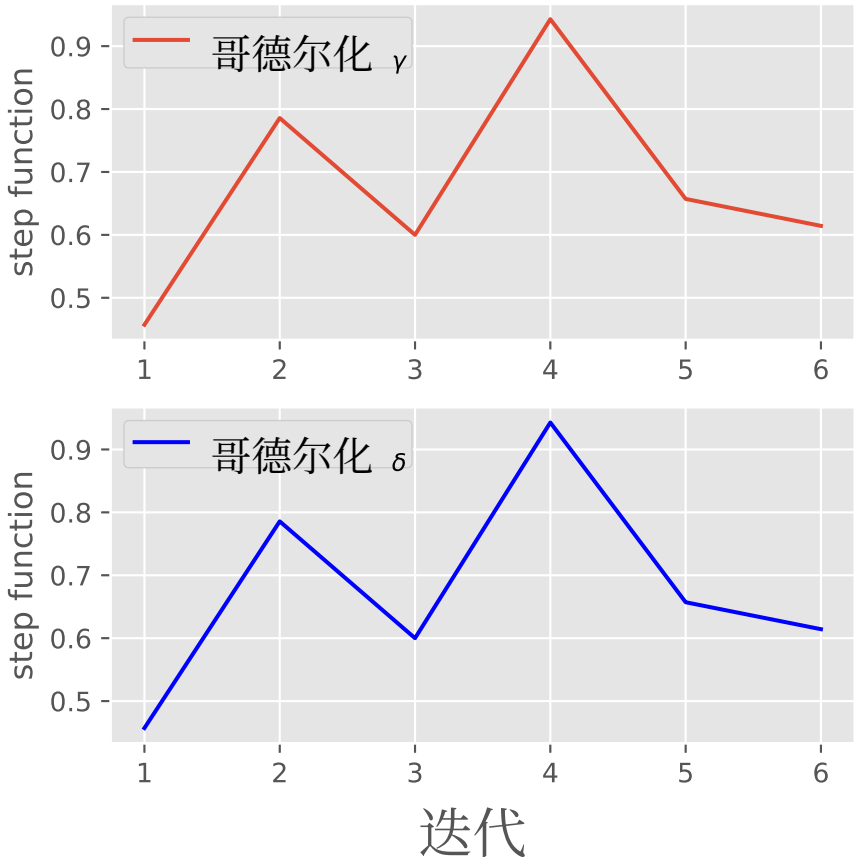

FIGURE 7. 宏观可观测量 f由阶跃函数（23）给出，在 G¨odel 重新编码下保持不变。图中展示了对顶部由 γ 编码的神经自动机以及底部由 δ 编码的同一机器进行“测量” f 的结果。

阶跃函数显然能够区分不同的状态（这里“不同”指具有不同的相等模式），但对于对应于相同相等模式的状态，也就是在 G¨odel 编码上有所差异的状态，则返回相同的值，正如所期望的那样。

相比之下，如果对同一模拟使用 Amari 观测量方程（9），我们会得到截然不同的图景，这表明如图8所示，该观测量在 G¨odel 重新编码下并非不变量。显然，该观测量很大程度上取决于我们所选择的特定 G¨odel 编码。

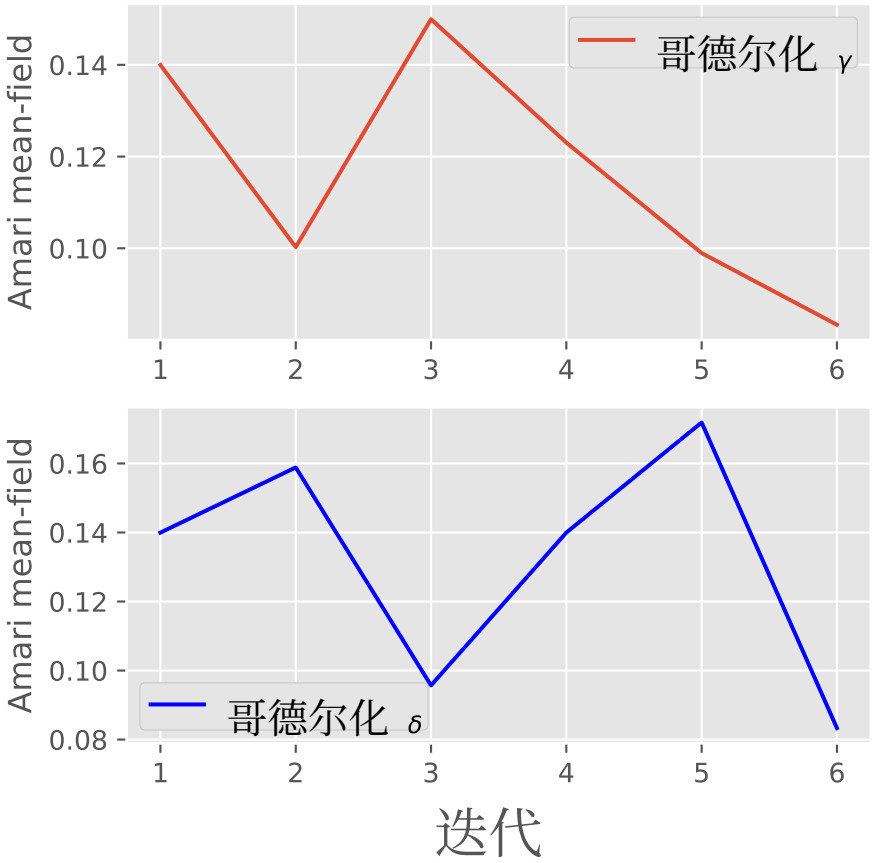

FIGURE 8. 神经自动机在两种不同哥德尔编码 γ 和 δ下的 Amari 平均场可观测量公式 (9)。

4. DISCUSSION

本研究提出了一种方法，用于寻找由符号动力系统的 G¨odel 编码生成的非线性动力系统的特定宏观观测量，例如非线性动力自动机（NDA：beim Graben 和 Potthast（2014），beim Graben 等人（2000；2004；2008））及其相应的神经网络实现，即神经自动机（NA：Carmantini 等人（2017））。具体而言，我们探讨了在什么情况下，这类观测量能够在任何特定的 G¨odel 编码选择下保持不变。

将符号动力学映射到实相空间时，符号的编号通常是任意的。因此，探究在不同可能选项下，我们在相空间中所看到的动力学信息有哪些得以保留或能够被恢复，是有意义的。沿着这一方向，我们根据相等模式，对经过特定 G¨odel 编码后可区分和不可区分的字符串进行了完整刻画。我们还为这类不变量证明了一个划分定理。

在按照 Carmantini 等人（2017）构造的非确定自动机这一具体情形中，该自动机能够模拟任意图灵机，因此我们得到了一个神经自动机的动力系统。该系统完全取决于非确定自动机字母表中符号的 G¨odel 编号选择。基于前文提到的不变划分，我们得以定义一个在任意 G¨odel 重新编码下保持不变的宏观可观测量。事实上，按照我们定义该可观测量的方式，其定义基于这样一种不变划分：根据构成非确定自动机机器纸带的带点序列左右两侧字符串的长度进行划分。这意味着，对这些字符串长度的每一种选择，都会产生更精细的不变量，从而使具有不同相等模式的字符串完全可区分。同样重要的是，宏观可观测量通常并不在 G¨odel 重新编码下保持不变。作为一个具体例子，我们计算了最初由天雨（1974）提出、后来由 Carmantini 等人（2017）在神经计算中作为一种建模的“合成事件相关电位”使用的神经网络平均激活（Barr`es 等人，2013）。

事实上，任何在 G¨odel 重新编码下保持不变的可观测量，对于相空间中对应于具有相同相等模式的字符串的 G¨odel 编码的点，都必须具有相同的定义。这或许能为寻找其他不变宏观可观测量提供重要约束。

理论上，可以在所有（或许多）可能的 G¨odel 编码下运行神经自动机，并检查哪些可观测量在动力学作用下得以保持，哪些则不会。这可以提供有关神经网络架构性能的重要信息，这种性能是动力系统固有的，而不取决于用于符号编码的编号方式。实际上，字母表所有排列的计算量会随着字母表基数的阶乘增长；对于较长的字符串，不变划分的计算量甚至会随着该数量的幂次增长。当然，这会给较大的字母表和精细的不变可观测量带来一些实际限制。

我们的研究结果对于计算认知神经动力学领域内任何相关方法都可能具有重大意义。所有依赖于以高维激活向量表示符号性心理内容，并将其作为（深度）神经网络训练模式的模型（Arbib 1995，Hertz 等人（1991），LeCun 等人（2015），Schmidhuber（2015）），尤其是向量符号架构（Gayler（2006），Mizraji（1989；2020），Schlegel 等人（2021），Smolensky（1990；2006）），都面临着符号编码任意性的问题。只要人们仅仅关注构建用于人工智能的推理机器，这其实无关紧要。然而，当必须将神经网络模拟的激活状态与人类或动物认知神经科学及心理学领域实验所得的真实世界数据相关联时，所采用的编码就可能发挥作用。因此，为了避免因特定编码选择而可能产生的混杂因素，有必要在回归分析和统计建模中考察不变可观测量。

这些结果对于数学与计算神经科学也具有启示意义。该领域旨在借助数学理论和计算建模，解释通过仪器设备在体外和体内实验中观察到的神经生理过程。我们的结果促使我们思考这样一种可能性：文献中推动模型（例如脉冲模型）发展的观测究竟在多大程度上（如果确有其事）只是副现象？最后，我们希望本研究能为更全面的研究铺平道路，

以推动计算认知神经动力学以及数学与计算神经科学领域的全面研究；在这些领域中，对宏观观测及其不变量表述的研究能够带来有趣的新见解。

4.1. 可重复性。第 3 节中 प्रस्तुत的所有数值模拟都可以使用 GitHub 代码仓库 https://github.com/TuringMachinegun/Turing_Neural_Networks 中提供的编码进行复现。该代码仓库包含用于构建神经自动机架构的编码，该架构在（ Carmantini 等人（2017））中进行了介绍，同时还包含一些具体示例。用于根据等式模式计算不变划分的编码也可以在该代码仓库中找到。用户可以利用这些编码实现各种观测量（例如阶跃函数、Amari 观测量），从而测试更多情形、利用并进一步发展我们的框架。

5. Acknowledgements

SR 感谢 Ikerbasque（巴斯克科学基金会）、巴斯克政府通过 BERC 2022—2025 计划以及西班牙科学与创新部提供的支持：BCAM Severo Ochoa 认证 CEX2021-001142-S / MICIN / AEI / 10.13039/501100011033，以及通过由（ AEI/FEDER，UE）资助、项目名称为 MathNEURO 的 RTI2018-093860-B-C21 项目提供的支持。JUA 感谢西班牙政府提供的资助，包括 PID2020-117281GB-I00 和 PID2019-107444GA-I00，这些资助部分由欧洲区域发展基金（ERDF）提供；同时感谢巴斯克政府提供的 IT1483-22 项目资助。

Appendix A. Proofs of lemmata and theorems

引理 2.1 的。排序 γ 本身诱导了两个图之间的同构。具体而言，令 $g_\gamma\colon T_\gamma \to T$ 满足：若 $s = a_{i_1}a_{i_2}a_{i_3}\cdots \in T_\gamma$，则

$$g_\gamma(s) = \gamma(a_{i_1})\gamma(a_{i_2})\gamma(a_{i_3})\dots$$

这显然属于 T。

我们必须证明它定义了顶点之间的一个双射，并且保持关联关系。

很容易证明它是一个双射。具体而言，若 $w = a_{i_1}a_{i_2}\dots a_{i_n}$ 和 $u = b_{i_1}\dots b_{i_k}$ 是树 T_γ 的任意两个顶点，则 $g_\gamma(w) = g_\gamma(u)$ 意味着两个字符串长度相同，因此 $n = k$。又因为 $\gamma(a_{i_j}) = \gamma(b_{i_j})$，且 γ 是一个双射，所以必有 $a_{i_j} = b_{i_j}$，从而 $w = u$。此外，对任意 $v = l_1\dots l_n \in T$，存在 $z = a_{\gamma^{-1}(l_1)}\dots a_{\gamma^{-1}(l_n)}$，它显然通过 g_γ 映射到 v。

唯一剩下需要证明的是，g_γ 保持关联关系。也就是说，给定 $w \in \mathsf{A}^*$ 和 $a \in \mathsf{A}$，那么对于某个 $l \in \{0,\dots,m-1\}$，$g_\gamma(wa) = g_\gamma(w)l$。但从 g_γ 的定义也可以清楚地看出这一点。□

引理 2.2。假设 $d(p,q) \leq \frac{1}{m^n}$。这意味着两个字符串中至少前 n 个符号相同。那么，如果 ψ 是由以下内容定义的 G¨odel 编码，

赋值 $\gamma : \mathrm{A} \to \mathrm{M}$ 满足

$$\psi(p) = \sum_{i=1}^{\infty} \gamma(a_i) \frac{1}{m^i} = \sum_{i=1}^{n} \gamma(a_i) \frac{1}{m^i} + \sum_{i=n+1}^{\infty} \gamma(a_i) \frac{1}{m^i}.$$

$_n$令我们设 $r = \sum \gamma(a_i) \frac{1}{m^i}$。现在，由于 $\gamma(a_i)$ $\leq m-1_{i=1}$

$$r = \frac{a_1 m^{n-1} + a_2 d^{n-2} + \ldots a_{n-1} m + a_n}{m^n} \leq$$
$$\leq \frac{m^n - m^{n-1} + m^{n-1} - \cdots - d + d - 1}{m^n} = \frac{m^{n-1}}{m^n}.$$

$_\infty$因此，对于某些 $k = 0$，……，$m^n - 1$。由于 $\sum \gamma(a_i) \frac{1}{m^i} < \frac{1}{m^{n+1}}$，我们得到 $_{i=n+1}\psi(p) \in \left[\frac{k}{m^n}, \frac{k+1}{m^n}\right)$。由于 q 至少在前 n 个字符串上等于 p，我们还将有 $_\infty \psi(q) = r + \sum \gamma(b_i) \frac{1}{m^i}$，出于同样的原因，它将在相同的$_{i=n+1}$区间上。

对于另一个方向，如果我们在对某些无限字符串 p 和 q进行编码后得到两个实数 $\psi(p)$ 和 $\psi(q)$，我们希望证明：如果它们位于某个$\left[\frac{k}{m^n}, \frac{k+1}{m^n}\right)$, 类型的区间内，那么它们至少具有长度为 n的相同前缀。我们总可以将这些数写成 $\psi(p)$ $= \frac{k}{m^n} + r_1$ 和 $\psi(q) = \frac{k}{m^n} + r_2$，其中$r_1$、$r_2 < \frac{1}{m^{n+1}}$。如果将数 k写成其 m进制展开式，那么它将由 l_1、……、$l_n \in \{0, \ldots, m-1\}$唯一确定，并且每个数 r_i都可以写成关于 $r_i = \sum l_{j_i} \frac{1}{m^j}$的级数，其中 $i = 1$、2。取每个 l_i的逆像，即 $_{j=n+1}$为 $\gamma^{-1}(l_i) = a_i$，我们将得到 $p = a_1 \ldots a_n a_{n+1} \ldots$和 $q = a_1 \ldots a_n b_{n+1} \ldots$。也就是说，它们之间的距离至少为 $\frac{1}{m^n}$. □

定理 2.4 的证明。首先假设，对于给定的 u，存在某个 π，使其诱导的自同构将 w 映射到 u。根据引理 2.2，条件式 (13) 显然成立。于是，我们可以写出 $u = c_1 \ldots c_n$。现在，我们知道 $g_\pi(w) = a_{\pi(1)} a_{\pi(2)} \cdots a_{\pi(n)} = c_1 \ldots c_n = u$。然后

$$\begin{aligned} \{j_1, \ldots, j_k\} \in P_w &\iff a_{i_{j_1}} = \cdots = a_{i_{j_k}} \\ &\iff a_{\pi(i_{j_1})} = \cdots = a_{\pi(i_{j_k})} \\ &\iff c_{i_{j_1}} = \cdots = c_{i_{j_k}} \\ &\iff \{j_1, \ldots, j_k\} \in P_u. \end{aligned}$$

为证明反方向，只需定义 π，使得 $g_\pi(w) = u$。也就是说，若$u = c_1 \ldots c_n$，则定义 $\pi(a_i)$ $= c_i$，并以双射方式将未出现在 w 中的 $a_j \in$M 映射到未出现在 u 中的 c_j-s。由构造可知，这样做是可行的，且根据条件式 (14)，定义良好，并在 M 上定义了一个双射。□

定理 2.6。我们需要证明方程（3）成立。也就是说，对于每个 $\pi \in S_{m-1} \times S_{m-1}$ 和 $\mathrm{x} \in X$，我们需要证明 $f(\alpha_\pi^*(\mathrm{x})) = f(\mathrm{x})$。注意，根据定义，在我们的情形下 $\alpha_\pi^* = \rho_\pi$。因此，我们必须证明 $f(\rho_\pi(\mathrm{x})) = f(\mathrm{x})$。设 $\pi = (\pi_1, \pi_2) \in S_{m-1} \times S_{m-1}$ 且 (y_1, y_2) $\in Y$。那么，存在某个 $E^{(i,j)}$，使得 $(x, y) \in E^{(i,j)}$。记 $\pi' = (\pi_1', \pi_2') \in S_m \times S_m$ 为固定 1 并将 $\pi'(i) = \pi(i) + 1$（其中 $i = 2, \ldots, m-1$）的置换（也就是说，固定第一个字母，并按 π_i 对 $m-1$ 个字母的置换方式置换其余字母，其中 $i = 1$，分别为 2）。

注意，由于我们在字母表中增加了符号 $\sqcup$，且原始编码和经 π 置换后的编码都将该符号映射为 0，若 (y_1, y_2) 被解码为 $(c_1, \ldots, c_l)$ 和 $(d_1, .., d_r)$，则每个编码末尾可能出现的 0（表示字符串长度小于 l 和/或 r）仍将是 0，因此该点不会被映射为编码更长字符串的点。

应用 $\rho_{\pi'}$ 后，我们将得到 $\rho_\pi((y_1, y_2)) \in E^{(i',j')}$ 对于某些 i', j'。然而，由于 $\rho_{\pi'}$ 是通过$g_{\pi'_1}$ 定义的，并且 $g_{\pi'_2}$ 和 $E^{(i,j)}$、$E^{(i',j')}$ 都属于同一个P_{ij}。也就是说，它们具有相同的相等模式。因此，根据 f的定义，我们得到 $f(\mathrm{x}) = f(\rho_\pi(\mathrm{x}))$，证毕。□

References

Amari SI（1974）统计神经动力学的一种方法。《控制论》14：201–215 Anderson JR、Bothell D、Byrne MD、Douglass S、Lebiere C、Qin Y（2004）心智的整合理论。《心理学评论》111（4）：1036–1060 Arbib MA（编）（1995）《脑理论与神经网络手册》。麻省理工学院出版社，剑桥（马萨诸塞州） Arbib MA 、 Caplan D （1979）神经语言学必须是计算性的。Behavioral and Brain Sciences 2（03）：449–460 Atmanspacher H、beim Graben P（2007）心理状态从神经动力学中的语境涌现。《混沌与复杂性快报》2（2/3）：151–168 Barr`es V、III AS、Arbib M（2013）合成事件相关电位：神经语言学模型与实验之间的计算桥梁。神经网络37：66–92 Boston MF、Hale JT、Patil U、Kliegl R、Vasishth S（2008）以句法分析成本预测阅读难度：使用波茨坦句子语料库进行评估。《眼动研究杂志》2（1）：1–12 Brouwer H、Crocker MW（2017）论语言理解中对P400和P600的恰当处理。《心理学前沿》8：1327 Brouwer H、Hoeks JCJ（2013）语言理解的时间与位置：将N400和P600映射到最小皮层网络。《心理学前沿》7（758） Brouwer H、Fitz H、Hoeks J（2012）正视语义错觉：重新思考P600在语言理解中的功能作用。《脑研究》1446：127–143 Brouwer H、Crocker MW、Venhuizen NJ、Hoeks JCJ（2017）语言处理中的N400和P600神经计算模型。《认知科学》41（S6）：1318–1352 Brouwer H、Delogu F、Venhuizen NJ、Crocker MW（2021）语言理解中意外度的神经行为相关性：一个神经计算模型。《心理学前沿》12：110 Carmantini GS、beim Graben P、Desroches M、Rodrigues S（2017）循环神经网络中透明计算的模块化架构。神经网络85：85–105 Cleeremans A、Servan-Schreiber D、McClelland JL（1989）有限状态自动机与简单循环网络。《神经计算》1（3）：372–381 Crocker MW（1996）《计算心理语言学》，《计算心理语言学研究》第20卷。Kluwer，多德雷赫特

Davidson DJ、Martin AE（2013）使用广义线性混合效应模型将建模准确率表示为反应时间的函数。Acta Psychologica 144（1）：83–96 Elman JL（1990）发现时间中的结构。《认知科学》14：179–211 Elman JL（1991）分布式表征、简单循环网络与语法结构。Machine Learning 7：195–225 Emch G（1964）李维尔空间中的粗粒化与主方程。Helvetica Physica Acta 37：532–544 Frank SL、Otten LJ、Galli G、Vigliocco G（2015）句子中词语所传达信息量的事件相关电位反应。Brain and Language 140：1–11 Frisch S、beim Graben P、Schlesewsky M（2004）并行化语法功能：P600和P345反映了不同的重新分析成本。International Journal of Bifurcation and Chaos 14（2）：531–549 Friston KJ、Mechelli A、Turner R、Price CJ（2000）功能性磁共振成像中的非线性响应：气球模型、沃尔泰拉核及其他血流动力学。NeuroImage 12（4）：466–477 Gayler RW（2006）向量符号架构是应对Jackendoff挑战的一种可行替代方案。Behavioral and Brain Sciences 29：78–79，DOI 10.1017/S0140525X06309028 Gazzaniga MS、Ivry RB、Mangun GR（编）（2002）认知神经科学：《心智的生物学》。W. W. Norton，纽约（纽约州） van Gelder T（1998）认知科学中的动力学假说。Behavioral and Brain Sciences 21（05）：615–628 beim Graben P（2008）神经物理学基础。载于：Graben Pb、Zhou C、Thiel M、Kurths J（编）《超级计算神经科学讲座：复杂脑网络中的动力学》，Springer复杂性系列，Springer，柏林，第3–48页 beim Graben P、Atmanspacher H（2006）经典动力系统中的互补性。Foundations of Physics 36（2）：291–306 beim Graben P、Atmanspacher H（2009）拓展互补性思想的哲学意义。载于：Atmanspacher H、Primas H（编）《重构现实：沃尔夫冈·泡利的哲学思想与当代科学》，Springer，柏林，第99–113页 beim Graben P、Drenhaus H（2012）计算神经语言学。Zeitschrift f¨ur Germanistische Linguistik 40（1）：97–125 beim Graben P、Potthast R（2009）动态认知建模中的逆问题。Chaos 19（1）：015103 beim Graben P、Potthast R（2014）通用神经场计算。载于：Coombes S、beim Graben P、Potthast R、Wright JJ（编）《神经场：理论与应用》，Springer，柏林，第299–318页 beim Graben P、Rodrigues S（2013）网络中漏积分发放神经元场电位的生物物理观测模型。《Frontiers in Computational Neuroscience》6（100） beim Graben P、Rodrigues S（2014）论神经网络的电动力学。载于：Coombes S、beim Graben P、Potthast R、Wright JJ（编）《神经场：理论与应用》，Springer，柏林，第269–296页 beim Graben P、Liebscher T、Saddy JD（2000）通过动力系统解析歧义上下文无关语言：神经网络中的消歧与相变——来自事件相关脑电位（ERP）的证据。载于：Jokinen K、Heylen D、Njiholt A（编）《学习行为》，特温特大学，恩斯赫德，TWLT 18，第II卷：《内化知识》，第119–135页

beim Graben P, Jurish B, Saddy D, Frisch S（2004）动力系统的语言处理。《International Journal of Bifurcation and Chaos》14(2):599 – 621；beim Graben P, Gerth S, Vasishth S（2008）语言相关脑电位的动力系统模型初探。《认知神经动力学》2(3):229 – 255；beim Graben P, Barrett A, Atmanspacher H（2009）神经网络中宏观态语境涌现的稳定性判据。《网络：神经系统中的计算》20(3):178 – 196；Graves A, Wayne G, Danihelka I（2014）神经图灵机。arxiv:1410.5401[cs.ne], Google DeepMind；Hale JT, Lutz DE, Luh WM, Brennan JR（2015）以不同粒度的语言结构建模fMRI时间序列。载于：《2015年认知建模与计算语言学研讨会论文集》，北美计算语言学协会，丹佛（科罗拉多州）；Hale JT, Campanelli L, Li J, Bhattasali S, Pallier C, Brennan JR（2022）语言处理的神经计算模型。《语言学年度评论》8(1):427 – 446；Hepp K（1972）量子测量理论与宏观可观测量。Helvetica Physica Acta 45(2):237 – 248；Hertz J, Krogh A, Palmer RG（1991）《神经计算理论导论》，圣菲研究所复杂性科学研究讲义，第I卷。Perseus Books，剑桥（马萨诸塞州）；Hopcroft JE, Ullman JD（1979年）《自动机理论、语言与计算导论》。Addison–Wesley，门洛帕克，加利福尼亚州；Huyck CR（2009）以模拟神经元实现的自然语言句法分析心理语言学模型。《认知神经动力学》3(4):317 – 330；Jauch JM（1964）量子力学中的测量问题。Helvetica Physica Acta 37:293 – 316；Jirsa VK, Jantzen KJ, Fuchs A, Kelso JAS（2002）利用网络建模求解脑电图和脑磁图的时空正向问题。《IEEE医学影像汇刊》21(5):493 – 504；Kelso JAS（1995）《动态模式：大脑与行为的自组织》。麻省理工学院出版社，剑桥（马萨诸塞州）；Krifka M（1995）极性项的语义学与语用学。《Linguistic Analysis》25:209 – 257；LeCun Y, Bengio Y, Hinton G（2015）深度学习。《自然》521(7553):436 – 444；Lewis RL（1998）重新分析与有限修复句法分析：走出花园路径。载于：Fodor JD, Ferreira F（编）《句子处理中的重新分析》，Kluwer，第247 – 285页；Lewis RL（2003）计算心理语言学。载于：《认知科学百科全书》，Macmillan Reference Ltd.，伦敦；Lewis RL, Vasishth S（2006）将句子处理视为熟练记忆提取的基于激活的模型。《认知科学》29:375 – 419；Lind D, Marcus B（1995）《符号动力学与编码导论》。剑桥大学出版社，剑桥（英国）；Mart´ınez-Ca˜nada P, Ness TV, Einevoll G, Fellin T, Panzeri S（2021）根据点神经元网络模型计算脑电图（EEG）。《公共科学图书馆计算生物学》17(4):1 – 41；Mazzoni A, Panzeri S, Logothetis NK, Brunel N（2008）兴奋性与抑制性神经元网络中的局部场电位频谱对自然刺激的编码。《公共科学图书馆计算生物学》4(12):e1000239；McCulloch WS, Pitts W（1943）神经活动中内在思想的逻辑演算。《数学生物物理学通报》5:115 – 133

McMillan B（1953）信息论的基本定理。《数学统计年鉴》24:196 – 219；Mizraji E（1989）线性分布式记忆中的语境依赖关联。《数学生物学通报》51(2):195 – 205；Mizraji E（2020）向量逻辑借助非的平方根实现反事实虚拟化。《IGPL逻辑期刊》；Nivre J（2008）确定性增量依存句法分析算法。《计算语言学》34(4):513 – 553；Rabinovich M, Friston K, Varona P（编）（2012）《脑动力学原理：全局状态交互》。麻省理工学院出版社，剑桥（马萨诸塞州）；Rabinovich MI, Varona P（2018）离散序列信息编码：异宿认知动力学。Frontiers in Computational Neuroscience 12:73；Rabovsky M, McRae K（2014）将N400事件相关电位成分模拟为语义网络误差：基于词义的特征联结主义吸引子模型所得启示。《认知》132(1):68 – 89；Rabovsky M, Hansen SS, McClelland JL（2018）将N400脑电位建模为意义概率表征的变化。《自然·人类行为》，第1 – 13页；Ratcliff R（1978）记忆提取理论。《心理学评论》85(2):59 – 108；Ratcliff R, McKoon G（2007）扩散决策模型：二选一决策任务的理论与数据。《神经计算》20(4):873 – 922；Schlegel K, Neubert P, Protzel P（2021）向量符号架构比较。《人工智能评论》；Schmidhuber J（2015）神经网络中的深度学习：综述。《神经网络》61:85 – 117；Sewell GL（2002）《量子力学及其涌现的宏观物理学》。普林斯顿大学出版社；Smolensky P（1986）动力系统中的信息处理：和谐理论基础。载于：Rumelhart DE, McClelland JL, PDP研究组（编）《并行分布式处理：认知微观结构探索》，第I卷，麻省理工学院出版社，剑桥（马萨诸塞州），第194 – 281页；Smolensky P（1990）张量积变量绑定与联结主义系统中符号结构的表征。《人工智能》46(1-2):159 – 216；Smolensky P（2006）语言认知中的和谐。《认知科学》30:779 – 801；Stabler EP（2011）多重上下文无关文法和最简主义文法的自顶向下识别器。载于：《第二届认知建模与计算语言学研讨会论文集》，计算语言学协会，美国俄勒冈州波特兰，第39 – 48页；Stephan KE, Harrison LM, Penny WD, Friston KJ（2004）功能性磁共振成像反应的生物物理模型。《神经生物学当前观点》14:629 – 635；Wilson HR, Cowan JD（1972）局部模型神经元群体中的兴奋性与抑制性交互作用。《生物物理学杂志》12(1):1 – 24

University of the Basque Country, Department of Mathematics, Leioa, Spain
Email address: jone.uria@ehu.eus

foldAI, Munich, Germany*Email address*: giovanni@carmantini.com
Bernstein Center for Computational Neuroscience, Berlin, Germany
Email address: Peter.beimGraben@b-tu.de
Basque Center for Applied Mathematics, Bilbao, Spain*Email address*:
srodrigues@bcamath.org