层级时间感受窗与生物学约束尺度不变深度网络中的零样本时间尺度泛化

Aakash Sarkar 心理与脑科学系 波士顿大学

Marc W. Howard 心理与脑科学系 物理系 波士顿大学

arXiv:2601.02618v1 [q-bio.NC] 6 Jan 2026

人类认知整合跨越嵌套时间尺度的信息。尽管大脑皮层呈现层级化的时间感受窗（TRW），局部回路却常常表现出异质的时间常数。为调和这一矛盾，我们基于尺度不变海马时间细胞训练了生物学约束深度网络，使其执行模拟语言层级结构的语言分类任务（例如，“字母”组成“词语”）。首先，利用前馈模型（SITHCon），我们发现，尽管网络在各层内具有完全相同的时间常数谱，跨层仍会自然形成时间感受窗层级结构。随后，我们将这些归纳先验提炼为一种生物学上合理的循环架构 SITH-RNN。对从通用循环神经网络到这一受限子集的一系列架构进行训练后发现，尺度不变的 SITH-RNN 以少数几个数量级的参数更快地完成学习，并能零样本泛化到分布外时间尺度。这些结果表明，大脑采用尺度不变的序列先验——编码“发生了什么”以及“何时发生”——因此，具有此类先验的循环网络尤其适合描述人类认知。

人类认知需要在远长于单个神经元内在时间尺度——膜时间常数、通道动力学——的时间尺度上进行记忆。例如，要理解这个句子末尾的含义，读者必须整合来自句子开头的信息。随着叙事结构通过许多句子和段落逐步展开，必须整合和提取信息的时间尺度也随之延长。

类似于空间感受野描述视觉系统中的神经元倾向于响应视网膜空间特定区域内信息的方式，时间感受窗（TRWs）界定了神经元群体处理信息的持续时间。大量采用认知神经影像学方法的研究已在人类皮层中发现了时间感受窗的层级结构（图1a-c；Lerner、Honey、Silbert和Hasson，2011年；Honey等，2012年；Hasson、Chen和Honey，2015年）。例如，参与者聆听以正常顺序呈现或按不同时间尺度（词语、句子、段落）打乱的故事。如果某个脑区在特定尺度上具有时间感受窗，那么在较短时间尺度上打乱输入时，该脑区的反应将保持不变。早期感觉区（如听觉皮层）即使在叙事中的单词顺序被打乱时，也会对单个单词作出稳定响应；而“高阶”脑区，包括顶叶和额叶中的一些区域，只有在较长时段（例如短语或句子）保持完整时，才会表现出稳定响应。

循环神经网络可能是大脑皮层时间感受窗层级结构的一种机制

时间感受窗意味着神经群体动力学必须在至少数秒的时间尺度上表现出记忆。循环神经网络（RNNs）中快速神经元之间的活动反复传递，可能是缓慢网络动力学的成因（Hebb，1949年）。采用具有循环连接的真实脑模型进行的计算研究能够展现缓慢的网络动力学（Compte、Brunel、Goldman-Rakic和Wang，2000年；Ding、Froudist-Walsh、Jaramillo、Jiang和Wang，2024年）。[1]循环神经网络在数学上易于处理且形式优雅（Maass、Natschl¨ager和Markram，2002年；White、Lee和Sompolin-sky，2004年；Rajan和Abbott，2006年；Dahmen、Gr¨un、Dies-mann和Helias，2019年），因此非常适合进行理论分析。此外，循环神经网络与持久的神经活动模式之间存在天然联系（Wong、Huk、Shadlen和Wang，2007年；Sussillo、Churchland、Kaufman和Shenoy，2015年；Rajan、Harvey和Tank，2016年）。

与计算神经科学领域的大量研究并行，人工智能领域的研究表明，循环神经网络可以通过训练完成复杂任务（Elman，

[1]缓慢的动力学也可能由神经元或突触的内在特性引起（Egorov、Hamam、Frans´en、Hasselmo 和 Alonso，2002年；Frank、Rudy、Levy 和 O’Reilly，2005年；Hasselmo、Giocomo、Brandon 和 Yoshida，2010年；Sheffield、Best、Mensh、Kath 和 Spruston，2011年；Tiganj、Hasselmo 和 Howard，2015年；Guo、Huson、Macosko 和 Regehr，2021年）。然而，在许多情况下，这些网络都可以用数学形式写成循环神经网络。本文不深入探讨这些函数方程的生物学成因。

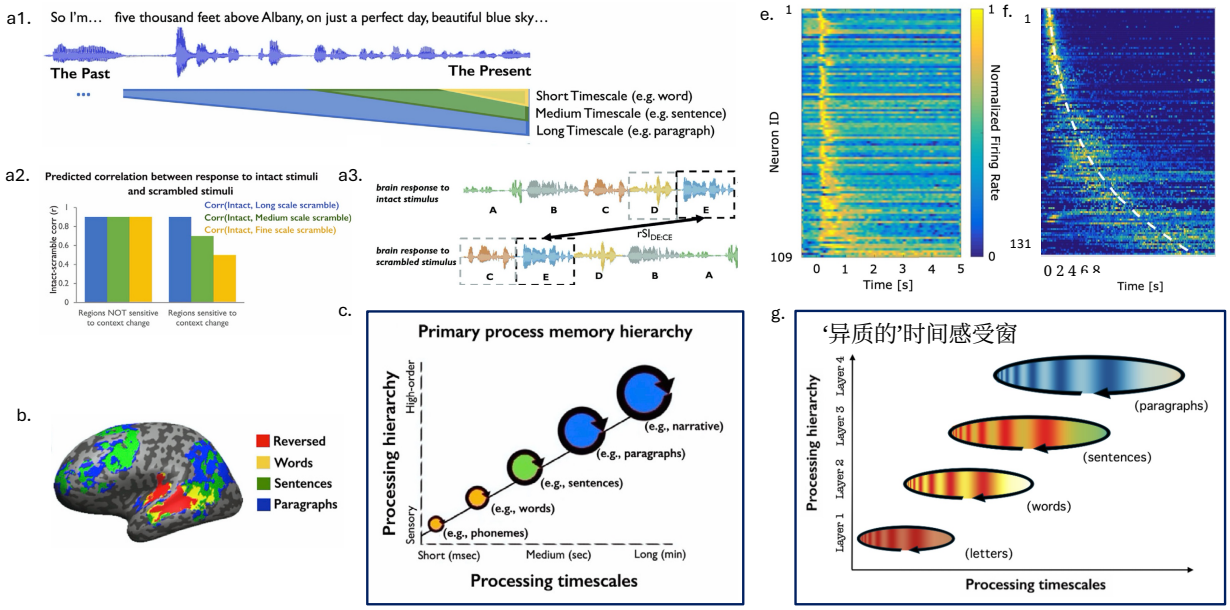

图 *1*。 有证据表明，尽管各脑区的时间尺度存在异质性，它们仍可能呈现出时间感受窗逐渐增大的层级结构。*左*： 口语包含多种时间尺度。为了理解当前单词，人们必须同时追踪词语、句子和段落（*a1*）。不同脑区需要不同时间尺度的上下文来处理当前输入，这种上下文范围被定义为时间感受窗（TRW）。为了测量这一处理时间尺度，已有大量研究通过测量脑区对自然情境刺激（包括完整刺激和打乱后的刺激）的反应活动相关性来进行探究（*a2*）。在不同时间尺度上打乱序列，相当于替换输入之前不同长度的上下文（*a3*），并借此检验脑区对这些不同尺度上下文的敏感程度（*b*）。这些研究广泛支持了大脑皮层存在时间感受窗层级结构这一假设（*c*）：感觉区处理输入前的即时上下文，而高级脑区则收集并整合句子和段落等更长时间尺度的上下文。*右*： 记忆神经科学的近期研究发现了一些细胞群，称为时间上下文细胞（*e.*）和时间细胞（*f.*），它们主要见于海马体及其他一些脑区；这两类细胞都以不同的时间感受野编码近期过去的压缩记录。这些细胞群呈现出单神经元时间常数的连续分布，其中编码近期过去的细胞更多。结合其他关于脑区内单神经元时间尺度存在异质性的相互印证证据，这些发现引出了一个问题：大脑是否实际上在每个脑区内都具有时间常数分布，以支持灵活性和重标度？然而，在处理具有层级时间结构的序列时，这些脑区会呈现出不同的涌现处理时间尺度。图 1a–c 改编自 Chien & Honey（2020）。

1990年；Hochreiter& Schmidhuber，1997年；Graves、Mohamed 和 Hinton，2013年；Feng、Tung、Ahmed、Bengio 和 Hajimir- sadeghi，2024年）。

设想一系列彼此提供输入的神经群体，每个群体都具有循环连接。由于每个区域内部存在循环动力学，每个区域的输出理应比其输入具有更强的自相关性。这意味着，网络第二层的输入理应比第一层的输入具有更强的自相关性。在深层循环神经网络的各层之间，人们预期动力学会逐渐变慢，从而为分层时间感受窗提供一种可能的机制。要理解由循环神经网络构成的深层网络，必须了解区域内部的网络动力学。本文探讨一种关于循环连接的特定假设所带来的影响；这一假设受到人类记忆计算模型以及特定脑区内部神经动力学的启发。这些脑区包括海马体、内嗅皮层和前额叶区域，并被认为与工作记忆和情景记忆有关记忆。

时间细胞、尺度不变性与深层网络

与不同脑区之间的时间尺度层级结构并行，有大量证据表明，脑区内部的功能时间常数存在平滑的异质性*within* 大脑脑区。这一基本发现已在不同物种和任务中得到观察（Bernacchia、Seo、Lee 和 Wang，2011年；Wasmuht、Spaak、Buschman、Miller 和 Stokes，2018年；Fascianelli、Tsujimoto、Marcos 和 Genovesio，2019年；Cavanagh、Hunt 和 Kennerley，2020年；Spitmaan、Seo、Lee 和 Soltani，2020年；Rossi-Pool 等，2019年；Danskin 等，2023年），这表明单个神经元时间尺度在脑区内部的异质性是生物神经网络的一项基本特征。对被认为与记忆密切相关的脑区中的神经元放电模式进行研究发现，这种异质性反映了一种更具体的时间编码形式（图 1e、f）。

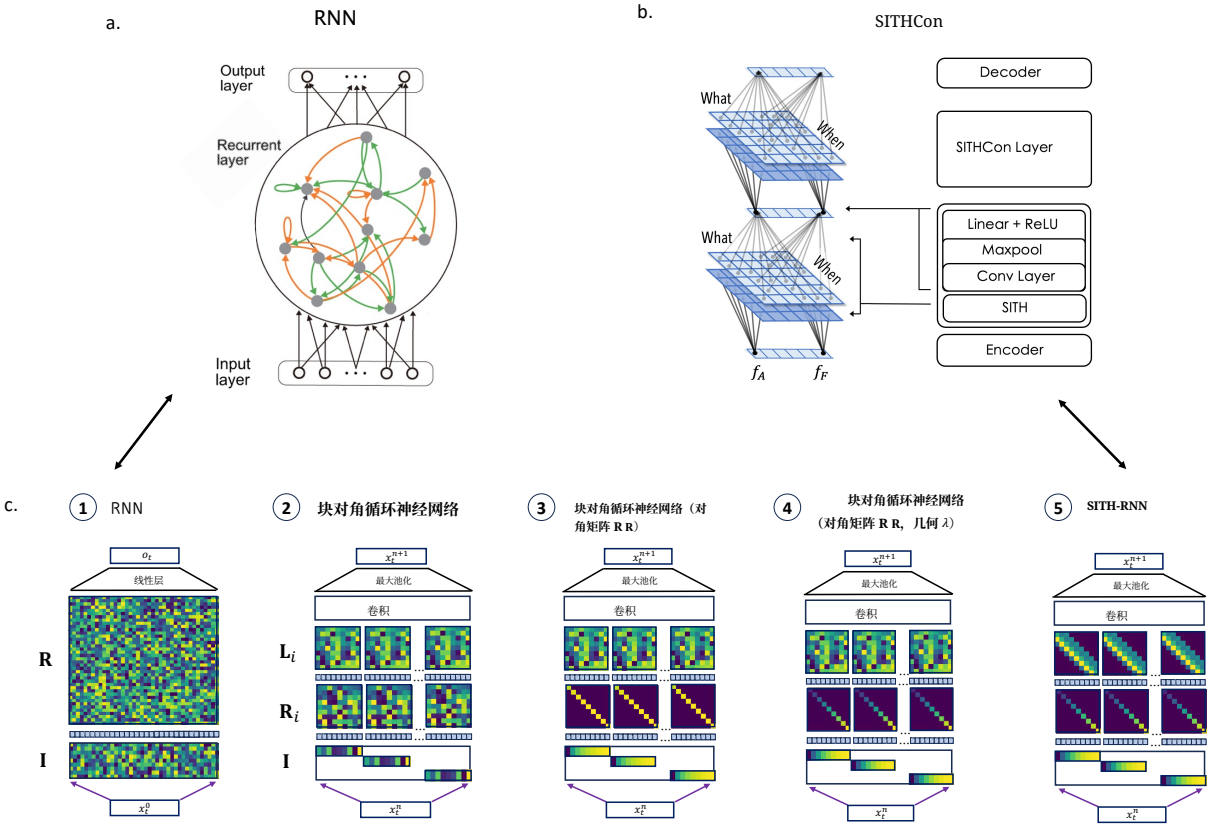

图 *2*。一系列带有归纳先验的循环神经网络。*a.* 神经科学中的通用循环神经网络具有将输入x_t投射到隐藏层的前馈权重（**I**）、作用于隐藏层的循环权重（**R**），以及将最后一个隐藏层投射到输出的前馈权重（**L**）（如*c1.*所示）。*SITHCon* 是一种前馈卷积神经网络，其归纳先验受到大脑中时间细胞群体的启发；这些细胞会维持一份对过去进行对数压缩的记录。这些先验赋予网络时间尺度不变性等特征，使其无需重新训练即可泛化到较慢或较快的输入。*SITHCon* 的每一层都有一个矩阵记忆，其中具有不同时间常数的时间细胞（When）会跟踪每个特征（What）。每层末端的卷积会捕捉这份对数压缩记忆中的模式，并生成可由下一层跟踪的特征。*c.* 我们提出一系列归纳先验，逐步将其加入神经科学中使用的通用循环神经网络（*c1*），从而构建出一个与*c5*中的*SITHCon* 非常相似的循环网络。为了模拟类似*SITHCon*的 What x When 记忆结构，我们约束投射矩阵（**I**）和递归矩阵（**R**）具有块对角结构（*c2*）。因此，What x When 矩阵记忆在这里表现为一个较长的隐藏状态，其中每个特征都在各自的时间子空间中独立地被跟踪和演化。与*SITHCon*类似，一组输出权重（**L**）生成输出状态，而卷积 + 最大池化层则将所有特征中的模式组合并重新映射为下一层的新特征。我们发现，要获得尺度不变的序列激活，还需要额外的先验，例如对角递归矩阵（在*c3*中引入）、（**R**）中的几何特征值（在*c4*中引入），以及（**L**）中的平移基元（在*SITH-RNN*、*c5*中引入）。

时间细胞。在海马体及其他脑区，所谓的时间细胞会在触发刺激后依次放电，将对事件发生时间的记忆延伸至长达数分钟的时间跨度（Pastalkova、Itskov、Amarasingham 和 Buzsaki，2008年；Jin、Fujii 和 Graybiel，2009年；MacDonald、Lepage、Eden 和 Eichenbaum，2011年；Tiganj、Cromer、Roy、Miller 和 Howard，2018年；Cruzado、Tiganj、Brincat、Miller 和 Howard，2020年；Taxidis 等，2020年；Shikano、Ikegaya 和 Sasaki，2021年；Liu 等，2022年；Schonhaut、Aghajan、Kahana 和 Fried，2022年）。

时间上下文细胞。此外，所谓的时间上下文细胞，其感受野会随时间呈指数衰减不同神经元的时间常数构成连续谱（Tsao 等，2018；Bright 等，2020；Ning、Bladon 和 Hasselmo，2022；Atanas 等，2023；Zuo 等，2024）。关键的是，至少对于海马时间细胞而言，时间常数的分布似乎在对数尺度上均匀间隔（Cao、Bladon、Charczynski、Hasselmo 和 Howard，2022）。

这种结果模式——由时间常数呈对数分布的时间上下文细胞和时间细胞构成的双重群体——可以用一个简单的数学假设来解释（Shankar 和 Howard，2013年）。根据这一假设，时间上下文细胞编码

过去随时间变化的实拉普拉斯变换，而时间细胞编码其近似逆拉普拉斯变换。这种映射为理解神经群体所表征的信息提供了高层次视角，因此可以用于构建一系列任务的认知模型（霍华德与艾肯鲍姆，2013年；霍华德等人，2014年；霍华德、卢萨多与蒂甘，2018年；蒂甘、克鲁萨多与霍华德，2019年；霍华德与哈塞尔莫，2020年；Salet、Kruijne、van Rijn、Los 与 Meeter，2022年）。在某些情况下，这种映射还可以用连续吸引子神经网络来解释（Daniels 与 Howard，2025年；萨卡尔、王、左与霍华德，2024年）。正如我们将要看到的，这一数学假设可以理解为循环神经网络的一种特殊情形。很容易写出一个线性循环神经网络，使其输出时间上下文细胞和时间细胞作为一种解（Liu & Howard，2020年）。

带有时间细胞的深层网络。Jacques、Tiganj、Sarkar、Howard 与 Sederberg（2022年）研究了由时间细胞组成的深层网络的性质。在该网络（称为 SITHCon）中，时间细胞通过矩阵记忆进行建模；矩阵记忆使用具有一系列时间尺度的时间基函数来存储特征（“内容”）和时间信息（“时机”）。时间常数按几何级数排列，从而使记忆呈现对数尺度。在每一层的末端，卷积层会在这种对数压缩的记忆中搜索模式，随后最大池化和线性投影会返回重映射特征，供下一层跟踪。SITHCon 对时间重缩放具有鲁棒性，这是因为对输入时间序列进行重缩放，会在其对数压缩的时间记忆中产生平移。卷积层天生对平移具有等变性，而后续的最大池化操作则通过舍弃最大激活值的确切时间索引，使网络对这些平移具有不变性。

论文概述

为了研究局部尺度不变动力学与全局层级处理之间的关系，我们采用了一种构造性计算方法。首先，我们定义了一项行为合成语言任务，其具有跨越多个时间尺度的内在层级结构，映射了自然语言的嵌套组合。接着，我们使用前馈网络模型（SITHCon）进行训练，该模型明确编码了假设的生物学约束——尺度不变性、几何时间常数和时间平移等变性——并在这一玩具语言上进行训练，以观察各层中涌现的处理时间尺度。最后，我们系统地推导出一系列循环神经网络，以确定在生物学上合理的循环神经回路中实现这些功能属性所需的最小结构先验集合。

处理时间尺度层级结构的涌现

在研究循环电路的动力学之前，我们首先检验一个基本假设：每一层都具有相同的对数压缩时间常数的系统，能否自发组织成处理层级？为回答这个问题，我们采用 SITHCon——一种充当典型尺度不变参考架构的前馈架构。通过明确硬编码对数压缩记忆基（近似于完美的时间细胞），SITHCon 让我们能够隔离这种归纳偏置所带来的功能后果，而不必引入训练循环动力学所带来的额外复杂性。

任务：一种层级玩具语言

我们构建了一种层级玩具语言，其中的符号分别对应字母、词语和句子，并具有确定性转换以及为每个层级设定的不同时间尺度（图3）。

这些序列由符号 1-9 构成，包含固定的组合，并由这些组合构成下一个层级单元——3个“字母”组合成“单词”，3个“单词”组合成“句子”，依此类推。只有某些组合是允许的，要预测第三个符号，了解前两个符号至关重要。这些规则旨在构造异或式组合——组合中的第一个、第二个和第三个位置总是分别从前三个符号（1 – 3）、中间三个符号（4 – 6）和最后三个符号（7 – 9）中选取。这也确保了在最终序列中，每个“字母”的边缘统计都均匀分布，因此它们出现的频率相同。知道第一个和第二个位置上的符号，就能唯一确定第三个符号，从而确定该组合的身份。由于这些组合发生在不同的尺度上，要理解序列的性质，就需要在多个层级的时间尺度上记住该序列，这类似于人们聆听并理解语音的过程。

最终构建的序列包含 4 个指定的层级，生成由 81 个符号组成的序列。网络的任务是根据这些层级中不同的符号组合对序列进行分类。我们以固定尺度训练网络对这些序列进行分类。

深层自发地组织出不同的时间事件边界

为了直观呈现网络如何分割连续输入流，我们分析了各层群体活动的时间自相似性（图 4a）。我们根据不同时刻的群体输出向量 $\mathbf{o}^l$ 之间的皮尔逊相关，计算了递归图（见方法）。从几何角度看，该指标量化了时刻t_i 和t_j经 Z 分数标准化的群体状态之间夹角的余弦，从而让我们能够直观地观察神经表征随时间的稳定性。

在较浅层中（图 4a，左），递归图呈现出细粒度的棋盘格模式。这表明神经元群体状态变化迅速，在各个符号的边界处有效重置，以表征即时的感觉输入。然而，随着深入到更深层（图 4a，右），出现了显著变化：对角线以外涌现出大块且稳定的高自相似性区域。

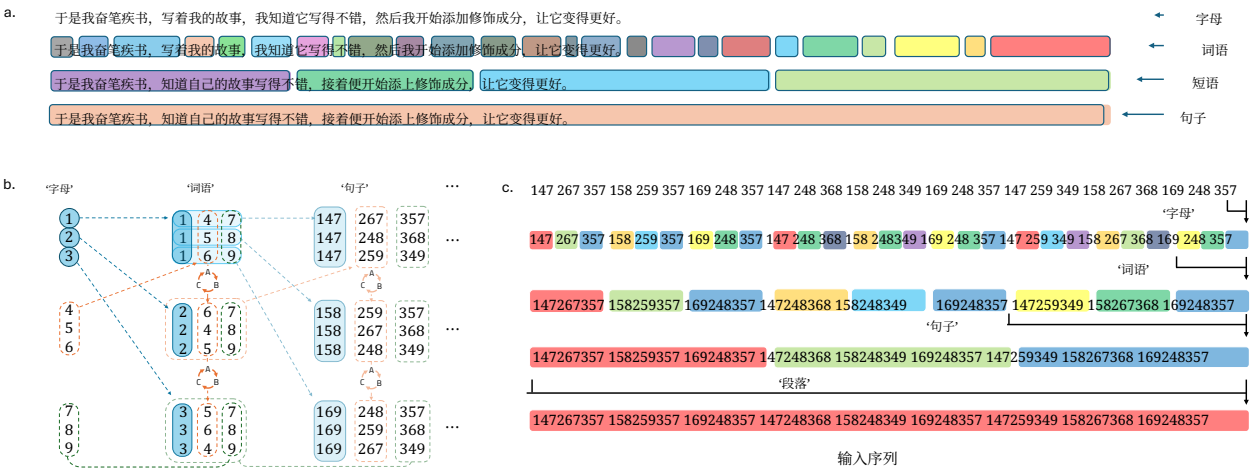

图*3*。 具有跨尺度结构的合成层次语言。*a.* 自然语言具有符号层级，每个层级组合起来形成下一个层级。为了理解当下，我们需要同时追踪这些时间尺度上的信息。*b.* 这种玩具语言在每个层级中包含9个符号，由3个'字母'组合成一个'单词'，依此类推。符号组合只采用特定形式，且三个位置各有不同规则——因此，仅凭前两个符号不足以预测第三个符号。例如，每个'单词'中的第一个'字母'只能是前三个字母（1、2和3）之一，第二个位置由中间的字母（4、5和6）排列而成，第三个位置则只能是最后三个'字母'（7、8和9）之一；不同规则共同生成所有可能的单词。从一个层级到下一个层级的符号由相同的模式生成——在每个层级上，符号转移都是确定性的。*c.* 由此产生的玩具语言也具有类似自然语言的层级结构。到达高层符号的边界时，低层的符号转移仍会突然改变，因此，试图对这种层次化序列进行分类的网络需要追踪多个时间尺度上的符号，正如理解语音需要回忆单词、段落以及更大范围的上下文。

这些区域表明，神经元群体在较长时间内维持着稳定表征，有效忽略了单个输入词元的高频抖动。这些稳定区域之间的急剧转变似乎与任务结构的层级边界（词语和句子）相吻合。这表明，尽管训练期间没有明确的分割线索，网络仍会自发地将输入分块为越来越抽象的叙事单元，而更深层的网络则基于经过重归一化的事件时间线运作，而非原始时间步。

层级时间感受窗的涌现：镜像语言结构

为量化每一层的有效处理范围，我们应用了镜像功能性神经影像学研究（Lerner等人，2011年）的时间扰乱扰动分析。我们以不同的置乱尺度s对输入序列x进行置换，置乱尺度范围从单个符号（3^0）到最长的9个训练序列（并排放置，$9 * 3^4 = 3^6$），并将层（l）对置乱输入的响应与其对完整序列的响应进行比较，$\mathbf{o}^l(x)$。

关键的是，由于打乱会扰乱时间线，我们无法直接比较输出序列。相反，我们首先应用逆置换，对层针对置乱输入的响应进行"反打乱"，$P_s^{-1}[\mathbf{o}^l(P_s[x])]$, 从而在时间上使其与原始输入序列重新对齐[2]。通过观察原始活动模式与反打乱后的活动模式之间的差异随s 的变化（图 4b），我们可以描述该层对先前上下文时间尺度的敏感性。

结果显示出明显的时间敏感性分化。只有当输入在非常细的尺度上被打乱时（例如单个符号），早期层的相关性才会下降；由于其整合窗口过短，无法检测到长程上下文的破坏，因此它们基本不受粗粒度置乱的影响。相比之下，即使输入在较大尺度上被打乱，更深层也会表现出性能的急剧下降。

为量化这种尺度变化，我们将每一层的有效时间感受窗（TRW）定义为相关性降至最大值一半时的扰乱尺度s_{50}（使用线性插值计算，见图4c）。这一指标提供了对整合时间尺度的稳健、无参数估计，使我们能够比较各层的处理带宽。我们观察到，这些时间尺度随网络深度呈指数级扩展，这一点可从半对数图上的线性变化看出。

[2]尽管未打乱的输出在时间上与原始输入对齐，但每个时间步之前的上下文 在尺度s上仍然是混乱的。为了严格量化这种差异，我们计算原始活动的递归矩阵（$\mathbf{o}^l(x)$，记为 ρ^l）与未打乱活动$P_s^{-1}[\mathbf{o}^l(P_s[x])]$, 的递归矩阵（记为$\rho^{l}{}_{\mathrm{shuff}}(s)$）之间的相关性。该指标$C(\rho^l,\rho^{l}{}_{\mathrm{shuff}}(s))$衡量该层在多大程度上依赖于超出$s$的时间上下文。

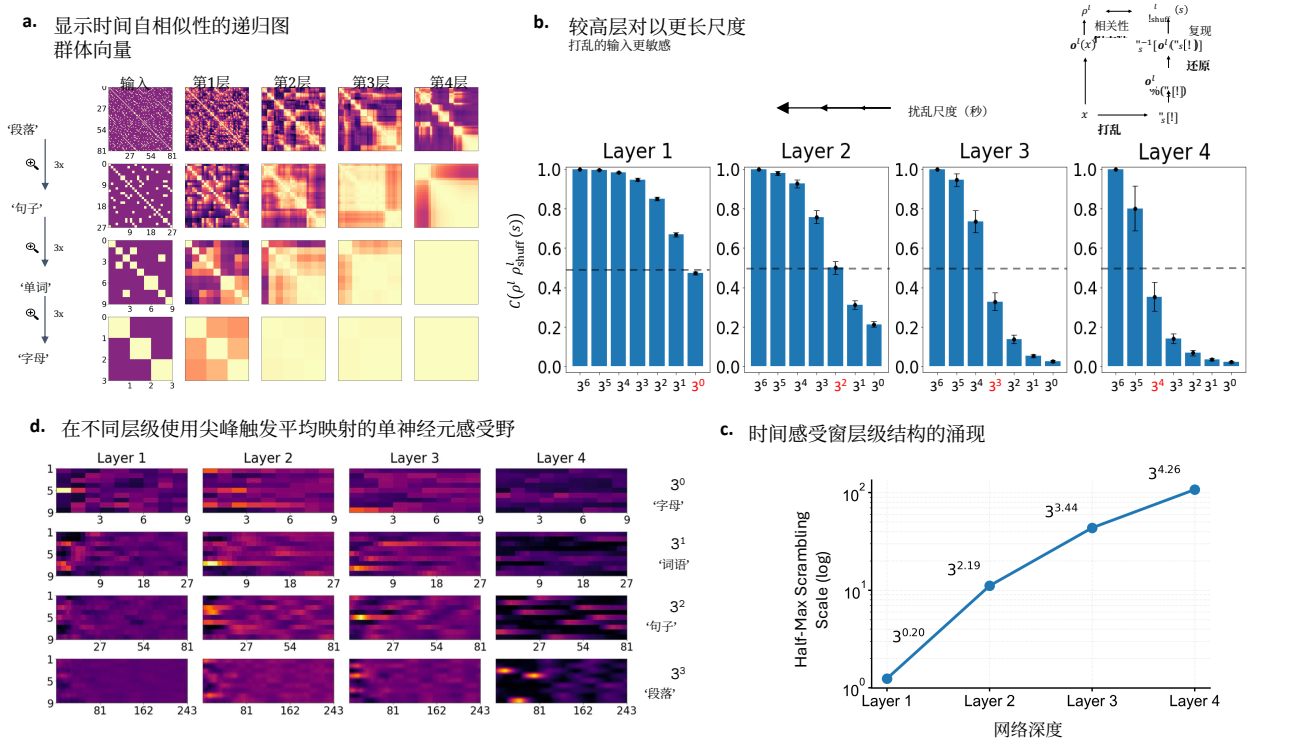

图*4*。 *SITHCon* 中层级时间感受窗的涌现。 *a.* 递归图揭示了各层逐渐增强的时间结构。图中显示了从输入（左）到第4层（右）的群体向量的时间自相似性。各行依次展示了 3× 时间放大（从上到下）。浅层呈现细粒度的时间动力学，而深层则能在更长时间内维持稳定状态。 *b.* 对层级打乱的敏感性。与时间感受窗文献中的方法类似，我们测量了模型对原始序列和按不同时间尺度打乱的序列（示意图见右上）的响应相关性。浅层仅对细尺度打乱（较小的 s）敏感，对大尺度打乱则保持稳健（较大的 s 下相关性较高）。相较之下，深层（例如第4层）依赖长程结构，因此对细尺度和粗尺度打乱都敏感（直至 3^4）；只有当打乱尺度超过训练序列的长度时，才恢复响应（$> 3^4$）。灰色虚线表示用于计算时间感受窗的50%相关性阈值。 *c.* 时间整合窗的涌现层级结构。各层的有效时间感受野（TRW）根据 *b* 中的数据，通过线性插值计算为半最大打乱尺度（s_{50}）。半对数图上的趋势近似线性，表明整合时间尺度呈指数扩展；第4层略微饱和，反映出网络与训练序列有界的全局时间尺度相匹配（$> 3^4$）。 *d.* 各层对层级的调谐。脉冲触发平均（STA）将代表性神经元（列）的线性感受野映射到不同层级的刺激（行）。第1层神经元（左）追踪基本符号，但忽略高阶结构。另一方面，第4层神经元（右）呈现嵌套的组合式结构：在'段落'层级上的单个宽激活带，在'句子'层级上分解为三个不同的激活带，依此类推。这表明深层神经元已学会玩具语言的层级映射，并通过语法中特定的组合句法定义其感受野（各层全部9个神经元的感受野见图S2）。

整合窗口从第1层中的基本符号$s_{50} \approx 3^{0.2}$扩展到第4层中近乎整个训练序列的长度$s_{50} \approx 3^{4.3}$，这证实网络已自主形成一个处理时间尺度呈指数级递增的层级结构。

更深层的神经元发展出"句子"和"段落"级感受野

最后，我们从群体层面的分析转向单个神经元的具体表征内容。我们利用脉冲触发平均（尖峰触发平均，STA）计算神经元的线性感受野（感受野，RF），以识别最能预测神经元放电的刺激特征。标准的尖峰触发平均通常假定单一时间尺度，而我们的层级任务则允许同时在多个层级计算感受野：我们可以探究什么样的'字母'、'单词'或'句子'最能驱动特定神经元。这种方法在此尤其适用，因为简化后的 SIT HCon 网络（去除了最终非线性）能够对神经元响应进行直接的线性解释。

图4d展示了四层网络中（从左到右）具有代表性的神经元在四个层级（从上到下）构建的线性感受野。第1层的输出充当简单的符号检测器，仅在第1级（'字母'）表现出可辨别的激活，而对更高阶结构

视而不见。然而，随着层级不断升高，神经元开始表现出特征抽象。第4层最右侧的一个代表性神经元不会对单个符号作出响应；相反，它会对特定的高层序列（一个‘句子’或‘段落’）产生选择性响应。

值得注意的是，这种高层次的选择性具有组合性。将这一分析扩展到整个群体后，可发现感受野呈现出稳健的类分形细分（图S2）：第4层神经元在“段落”层级上的宽广激活带，可以明确分解为对三个“句子”特定组合的偏好，而每个“句子”又可进一步分解为九个“词语”，依此类推。这表明，网络执行了某种逐层的“时间重整化”：每一层都将其输入历史压缩为离散特征，作为下一层的有效原子时间步。这一机制使系统能够构建对叙事结构的理解，而无需任何显式的分割线索。

尺度不变动力学的循环实现

上一节表明，对数压缩时间基分布（SITHCon）在功能上足以生成分层时间感受窗。然而，要在生物回路中实现这一机制，它必须能够在循环动态框架中实现。循环神经网络是神经科学中用于建模神经动力学的标准计算工具，但通用循环神经网络缺乏可解释性和结构。在此，我们将 SITHCon 的关键数学特性——尺度不变性、模块性和平移等变性——提炼到一种具有生物学合理性的循环神经网络（SITH-RNN）中，从而弥合这一差距。这使我们能够提出，SITH-RNN 不仅是一种机器学习架构，也可作为皮层处理的规范性架构框架候选。

推导 *SITH-RNN* 架构

循环动力学产生的尺度不变活动。在循环矩阵受到某些约束的情况下，线性循环网络已被证明能够产生连续的尺度不变活动（Liu & Howard，2020）。神经活动的尺度不变性意味着，序列中任意两个神经元的响应在时间上互为缩放版本。从数学上说，这一要求可以写为

$$x_i(t) = x_j(\alpha_{ij} t),\ \forall i, j \in 1, 2, \cdots, N \tag{1}$$

这意味着，对于每一对神经元i和j，都存在一个因子 α_{ij}，使它们的神经活动$x_i(t)$ 和$x_j(t)$分别按该因子在时间上缩放。

近期研究表明，具有隐藏状态$\mathbf{h(t)}$连续时间动力学和循环矩阵$\mathbf{R}$的单层循环网络

$$\frac{d\,\mathbf{h}(t)}{dt} = \mathbf{R}\mathbf{h}(t)\,, \tag{2}$$

能够维持尺度不变活动，但须满足某些约束——$\mathbf{R}$的特征值分解必须同时产生几何特征值和具有平移不变性的特征向量（Liu & Howard，2020）。具有这些限制的线性网络能够在$\mathbf{h}_t$ 中产生单调衰减的时间动力学；当网络扩展出一组下游线性权重$\mathbf{L}$[3]时，甚至能够表现出尺度不变序贯活动（如时间细胞）。本节将研究并扩展这些约束，以适用于具有多个输入特征的深层循环神经网络。

块对角权重矩阵产生“是什么” ×“何时”表征。通用循环神经网络具有隐藏状态$\mathbf{h}_t$，该状态会随新输入x_t 更新，并在每个时间点生成输出$\mathbf{o}_t$ 。通用循环神经网络单层的演变过程可以用方程 2 的离散时间版本建模，其中前馈权重$\mathbf{I}$和$\mathbf{L}$ 分别负责向隐藏空间投射输入和输出

$$\mathbf{h}_t = \mathbf{R}\mathbf{h}_{t-1} + \mathbf{I}\mathbf{x}_t\,, \tag{3}$$

$$\mathbf{o}_t = \mathbf{L}\mathbf{h}_t\,. \tag{4}$$

在标准循环神经网络的表述中，输入（$\mathbf{I}$）和循环（$\mathbf{R}$）权重是稠密且完全可训练的，因此会产生无结构的表征，将刺激身份（“是什么”）与时间历史（“何时”）混杂在一起。

为避免这种纠缠，我们施加一种结构先验（Sarkar 等人，2024），通过将隐藏状态定义为特征子空间与时间子空间的张量积，为每个特征保留独立的时间线。

$$\mathbf{h_t} = \bigotimes_{k=1}^{n_f} \mathbf{h_t^k} \tag{5}$$

其中，$\bigotimes$ 表示张量积，n_f 表示特征数量，而$\mathbf{h_t^k}$ 表示每个特征k对应的时间子空间。每个时间子空间$\mathbf{h_t^k}$ 由n_τ 个具有不同时间常数的神经元组成，从而形成对该特征的记忆[4]。

循环权重（$\mathbf{R}$）和前馈权重（$\mathbf{I}$）也必须具有额外的结构，以保留$\mathbf{h_t}$中“是什么”与“何时”的张量积结构，并确保每个特征的时间子空间独立演化——具体而言，每个子空间都必须由独立算子$\mathbf{R}_k$ 作用于其上，这意味着$\mathbf{R} = \bigotimes_{k=1}^{N} \mathbf{R}_k$具有分块对角结构。此外，每个特征必须以相同方式演化（以维持可分离表征），因此各矩阵应具有相同的子块；等价地说，它们应是对应的特征空间与时间空间的张量积。

[3]这类具有两组权重的网络，其活动也可以等效地用一组有效循环权重表示，定义为$\mathbf{R}_{\text{eff}} = \mathbf{L}\mathbf{R}\mathbf{L}^{-1}$——由于这种形式与$\mathbf{R}_{\text{eff}}$的特征分解相同，上述约束可简化为：$\mathbf{R}$ 为具有几何特征值的对角矩阵，而$\mathbf{L}$ 的各列包含重复的平移基元，从而提供平移后的特征向量。

[4] 这形成了 $\mathbf{h_t} \in \mathbb{R}^{n_f \times n_\tau}$，令人联想到 SITHCon 中用于“什么”和“何时”的矩阵记忆，不过不同之处在于，此处的时间子空间 $\mathbf{h_t^k}$可以串联起来并表示为一个向量。

$$\mathbf{R} = \mathbf{R}_{what} \otimes \mathbf{R}_{when} \tag{6}$$

其中，就我们的目的而言，$\mathbf{R}_{what}$ 是一个对角矩阵，因为假设特征空间是正交的，且$\mathbf{L}$ 和$\mathbf{I}$ [5]具有类似的先验。

当前任务通常可能需要在不同层之间重新映射特征（在当代机器学习架构，如状态空间模型中，称为通道混合）——为此，我们在下游权重$\mathbf{L}$之后引入一个卷积层（类似于 SITHCon 中的卷积层，可使其对时间重缩放具有鲁棒性）。因此，对于我们的情况，式 4 更新为

$$\mathbf{o}_t = \max_{n_\tau}(g \star \mathbf{L}\mathbf{h}_t) \tag{7}$$

其中，二维卷积滤波器g 在特征 × 时间矩阵重塑后的$\mathbf{L}\mathbf{h}_t$ 版本上搜索模式（这使卷积核能够同时观察所有特征的历史，而不是常规的拼接向量形式）；最大池化操作（$\max_n$）则沿时间维度运行，为每种模式返回一条轨迹，并将检测到的特征组合起来传递给下一层（详细示意图见图 9）。

约束连续体：从通用循环神经网络到 *SITH-RNN*。为单独考察上述特定约束——分块对角结构、固定几何时间尺度和卷积读出——所产生的作用，我们构建了由五个网络组成的架构连续体（图 2；图 5 还展示了这些网络的时间动态和泛化性能）。虽然网络*1* 是一个标准线性系统，但网络 *2* 至 *5*共享式 5–7 中描述的张量积结构，并使用共享的卷积特征混合器（g）和循环块（$\mathbf{R}_{when}$），它们在时间子空间上独立运行。

网络*1*（通用循环神经网络）作为基准线性动力系统，其中$\mathbf{I}$ 和$\mathbf{R}$ 均为稠密且完全可训练的。

网络*2*（稠密分块对角）强制采用张量积结构，但保留稠密且可训练的循环子块（$\mathbf{R}_{when}$）和投影权重（$\mathbf{I}_{when}$以及$\mathbf{L}_{when}$），从而允许在时间子空间内进行任意混合。

网络 *3*和*4*（对角循环神经网络）将循环矩阵$\mathbf{R}_{when}$ 限制为对角矩阵，并将输入投影固定为$\mathbf{I}_{when} = \left[1 \cdots 1\right]$，使其在每个特征的时间空间中向所有神经元进行相同投影。这迫使隐藏状态充当具有不同时间常数的泄漏积分器组。

- 网络 *3*采用均匀分布的时间常数。
- 网络 *4* 采用几何分布的时间常数。[6]

网络 *5*（*SITH-RNN*）完成了这一转变，实现了时间记忆与特征绑定的完全解耦。与之前$\mathbf{L}_{when}$ 灵活变化的网络不同，SITH-RNN 冻结了循环连接（使用网络 4 中的几何特征值），并将$\mathbf{L}_{when}$ 中的学习限制在固定宽度的平移时间模式上，该模式以带状（托普利兹）矩阵的形式实现。[7]

固定的几何动力学与平移不变读出核的结合确保了SITH-RNN在构造上具有尺度不变性，能够同时在所有时间尺度上搜索相同的相对模式（有关SITH-RNN内部张量运算和信号流的详细示意图，请参见《方法》中的图9）。

尺度不变时间细胞源于几何递归特征值和平移读出特征向量

为理解 SITH-RNN 的底层动力学，我们分析了网络在接收增量脉冲输入后的脉冲响应。在图 5e、f 中，我们追踪了一系列循环神经网络架构中隐藏状态（$\mathbf{h}_t$）和投影输出（$\mathbf{L}\mathbf{h}_t$）的时间动力学演变。我们观察到，将循环权重$\mathbf{R}$ 限制为对角形式，无论采用均一特征值（见网络 *3*）还是几何特征值（见网络 *4*），都会使隐藏层变成一组积分器，其中的神经元以不同速率平滑衰减。

然而，只有在 SITH-RNN（网络 *5*）中才能观察到前馈动力学中的序贯活动；该网络还施加了第二项约束：读出$\mathbf{L}$采用带状矩阵结构——这会生成平移特征向量，将衰减的隐藏状态转化为读出$\mathbf{L}\mathbf{h}_t$ 中平滑、依序激活的峰值（图 5f，右）。与 SITHCon 中固定的解析权重（用于近似逆拉普拉斯变换）不同，这里的$\mathbf{L}$ 中的模式完全可训练；唯一的约束是零均值结构，而这足以形成感受野逐渐增宽的序贯场。

图 6 显示，这些激活确实具有尺度不变性。当将$\mathbf{h}_t$ 和$\mathbf{L}\mathbf{h}_t$ 的激活曲线重新绘制为归一化时间（t/τ_i）的函数时（其中 τ_i 是该神经元的时间常数），这些曲线会重合为一个单一的通用函数（插图），满足方程 1 中尺度不变性的条件。我们通过消融分析确认了这一特性是严格的：当且仅当

[5] 虽然整个循环矩阵$\mathbf{R}$ 有（$n_f \times n_\tau$）[2] 个条目，但实际的有效参数仅涉及块$\mathbf{R}_{when} \in \mathbb{R}^{n_\tau \times n_\tau}$，并且该块会针对每个特征重复。同样，对于投影权重，我们有$\mathbf{L}_{when} \in \mathbb{R}^{n_\tau \times n_\tau}$，以及$\mathbf{I}_{when} \in \mathbb{R}^{n_f \times n_\tau}$。[6]这引出了尺度不变性的第一个约束：如果将循环动力学和前馈动力学视为一个单一的有效算子$\mathbf{R}_{\mathrm{eff}}$，那么这一步确保$\mathbf{R}_{\mathrm{eff}}$ 具有几何特征值。[7]这引出了尺度不变性的第二个约束：带状结构意味着有效循环矩阵$\mathbf{R}_{\mathrm{eff}}$ 的特征向量具有平移不变性（Liu & Howard，2020）。

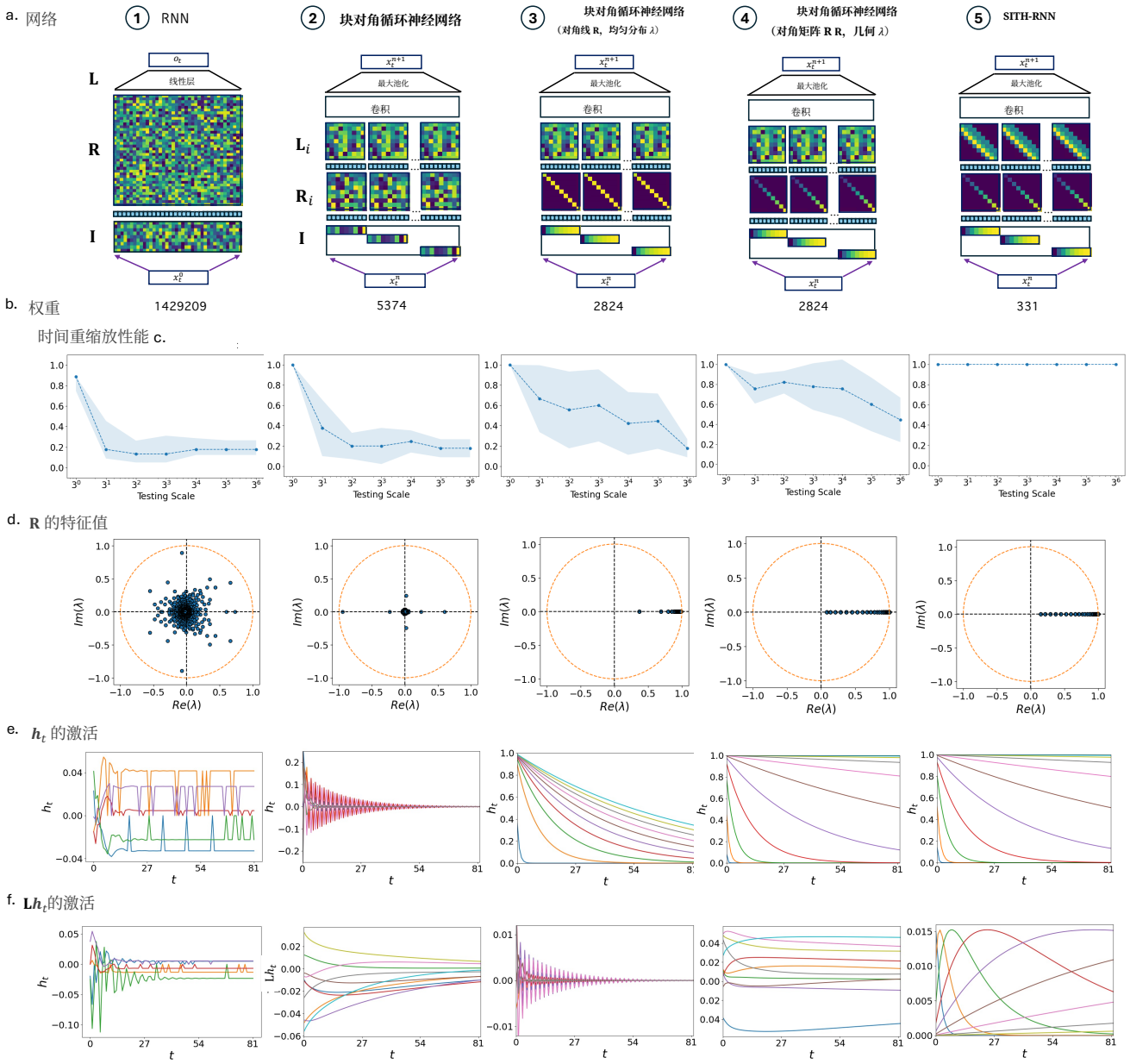

图*5*。 生物学约束的 *SITH-RNN* 通过施加尺度不变动力学，以更少的参数实现更强的泛化能力。 *a.* 架构约束的演变。我们构建了由五种循环网络组成的连续谱，从通用线性 RNN（左）出发，逐步加入受 SITHCon架构启发的归纳先验（例如，块对角连接、几何时间常数），最终得到 SITH-RNN（右）。 *b.* 参数效率。与通用 RNN 相比，加入这些结构约束后，可训练权重的数量大幅减少（减少了多个数量级），而隐藏状态维度保持相同。*c.* 时间重缩放的零样本泛化。在固定时间尺度（3^0）下使用层级序列训练，并在按最高达 3^6，的因子重缩放的序列上测试时，随着先验逐步增加，模型表现越来越好（从左到右）。SITH-RNN 在时间尺度跨越全部六个数量级的情况下均达到完美分类准确率，展现出稳健的零样本泛化能力。*d.* 循环的谱结构。循环矩阵的特征值**R** 从复单位圆上的均匀分布（通用 RNN，左），转变为分布在坐标轴上的实数值（块对角），最终呈现出沿实轴铺排的几何间距（SITH-RNN，右）；这是实现对数压缩（韦伯–费希纳定律）和尺度不变性的必要条件。 *e.* 隐藏状态的时间动力学（$\mathbf{h}_t$）。在t = 0处输入增量脉冲后，第一层隐藏状态的响应。对于循环矩阵**R** 为对角矩阵的网络（右），神经元表现为泄漏积分器，以不同速率平滑衰减，类似于生物学中的时间上下文细胞。 *f.* 读出的时间动力学（$\mathbf{L}\mathbf{h}_t$）。投影输出神经元的响应。**L** 中的带状矩阵结构（SITH-RNN 特有，右）产生平移特征向量，进而形成平滑、连续的序列激活，其时间常数谱与生物学时间细胞相似。

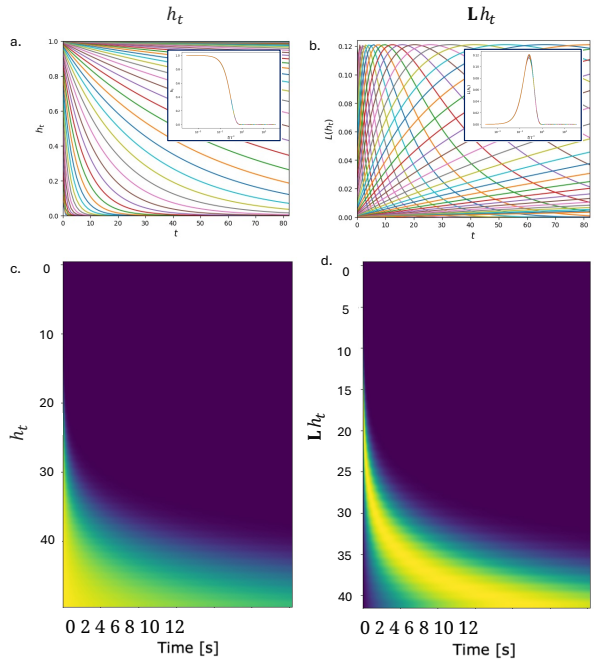

图 *6*。 *SITH-RNN* 在“对偶”的隐藏层和投影输出神经元群体中呈现尺度不变的序列活动。这里展示了隐藏层$\mathbf{h_t}$（左）以及前馈输出$\mathbf{Lh_t}$（右）的时间动力学；在$t = 0$时向单个特征输入增量脉冲，并分别显示单个神经元的激活轨迹（上）以及群体热图（纵轴为神经元单元，横轴为经过的时间）（下）。上方两幅图中的插图显示了将横轴按每个神经元的时间常数重新缩放后（t/τ_i）的激活情况，此时各曲线彼此重合。两个不同的群体呈现出明显不同的特征：隐藏层神经元同时开启，但以不同速率衰减（类似于时间情境细胞）；前馈神经元则在脉冲输入后依次开启，且时间场逐渐增大（类似于时间细胞）。

网络同时具备几何级数递归特征值以及读出中的平移模式（图 S1）。

最后，图6中的群体热图揭示了模型与生物记忆回路之间的直接同构。模型中的两个不同群体分别对应已知的神经表型：隐藏神经元（$\mathbf{h}_t$）同时开启并以不同速率衰减，重现了时间上下文细胞的动态；与此同时，读出神经元（$\mathbf{Lh}_t$）依次激活且感受野逐渐扩大，重现了时间细胞的动态。这证实了 SITH-RNN 不仅能完成任务，还能通过实例化一种生物学上合理、尺度不变的时间基来完成任务。

尺度不变先验使时间重缩放的零样本泛化成为可能

为了量化这些归纳先验带来的益处，我们在第 2 节介绍的层次分类任务上训练了网络

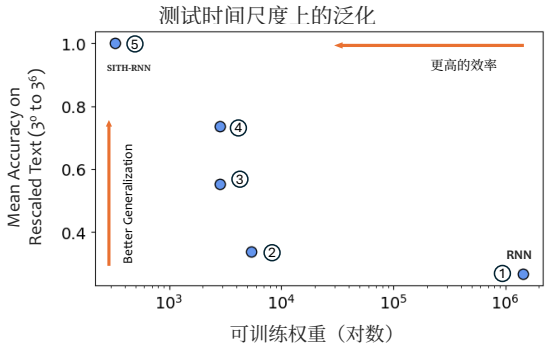

图7。尺度不变归纳先验以最小模型复杂度实现稳健泛化。我们绘制了各架构在连续时间尺度范围内的零样本泛化表现（跨越六个数量级时间尺度的平均分类准确率，取自图5c）与可训练参数数量之间的关系。通用循环神经网络（网络*1*）模型复杂度高，却无法泛化到分布外时间尺度。相比之下，施加基于SITH的归纳先验（网络 *2–5*）既能减少参数数量，又能提升稳健性。最终的 SITH-RNN 架构（网络 *5*）优化了效率与泛化能力之间的关系，以显著更少的可训练权重（通用循环神经网络的 < 0.05%）实现了完美的零样本泛化。

该任务要求模型隐式学习字母、词语和句子之间的边界，以预测下一个词元。我们在标准时间尺度（3^0）的序列上训练了所有模型。随后，我们测试它们能否零样本泛化到经过重新缩放的序列，缩放因子最高为 3^6（从*1* 倍到*729* 倍减速）。

如图5c所示，通用线性 RNN 成功学会了训练尺度下的任务（3^0），表明简单的线性递归具有足够的记忆容量来解决该任务。然而，随着时间尺度发生偏移，其性能立即下降。我们观察到架构连续体呈现出清晰的趋势：随着各项归纳先验依次引入，模型从左到右表现越来越好。在右侧，SITH-RNN 在测试尺度的全部六个数量级范围内都达到了完美的分类准确率。这种分化表明，尽管通用线性动力学能够记忆特定的时间模式，但只有几何归纳偏置才能使网络泛化相对的层级结构，而不受绝对持续时间的影响。

讨论

此处呈现的结果，为局部回路异质性与全局皮层层级之间看似矛盾的现象提供了解释。通过验证一组特定的归纳先验——它们为每个特征创建独立的对数压缩时间线，并在前馈和循环架构中实现尺度不变动力学——我们找到了一个候选机制，可解释大脑如何从感觉流中构建叙事结构。在本文

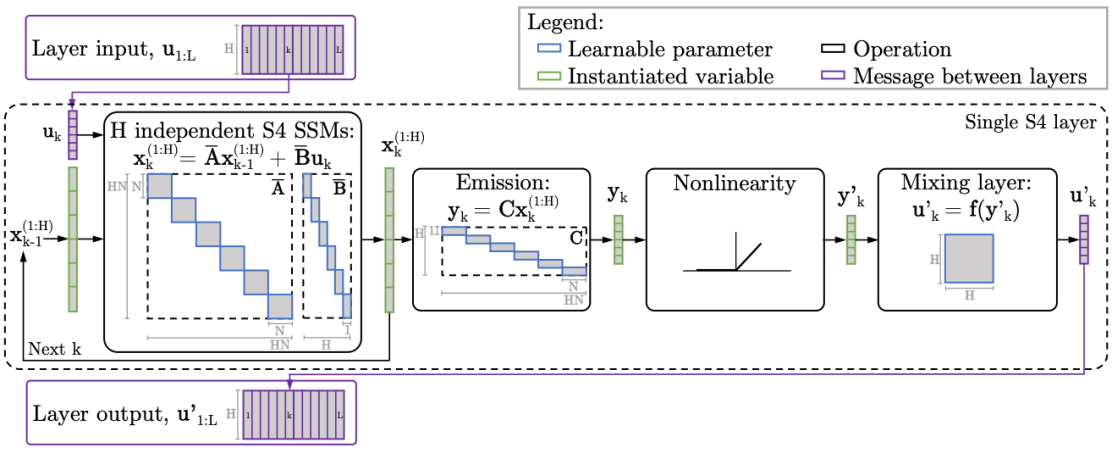

图 *8*。状态空间模型（*SSMs*）可以表示为一个块对角系统，与 *SITH-RNN* 类似，但每个特征都有独立的权重。该示意图展示了 S4（Gu等人，2021年）处理具有 L 个时间步的输入 $\mathbf{u}$ 时的动态过程，其中输入包含 H 个特征。在单个时间步中，每个输入特征都会传递给 H 个独立的 SSM，由它们对该特征进行扩展，并创建一个跨越 N 个维度的状态空间——生成维度为 HN 的记忆 $\mathbf{x}$。这也可以表示为一个组合而成的单一循环网络，其递归矩阵 $\mathbf{A}$ 为块对角矩阵，形状为 $HN \times HN$；块对角投影权重接收输入向量 $\mathbf{u}_k$，将每个特征扩展到其状态空间（$\bar{\mathbf{B}}$），再将 HN 维状态向量 $\mathbf{x}_k$ 投影回 H 维输出 $\mathbf{y}_k$。递归权重和投影权重均为块对角矩阵，因此每个特征的状态空间不受其他特征影响（末尾的混合层除外），这与 SITH-RNN 类似，不过在此类 SSM 中，每个特征对应的子块彼此独立。图引自 Smith等人（2022年）。

的讨论中，我们将考察这一机制的理论意义，提出可证伪的电生理学预测，并将研究发现置于机器学习时间建模这一更广阔的背景下。

局部时间异质性如何产生离散处理层级

人类皮层组织成一个时间感受窗（TRW）逐渐增大的功能层级结构，这表明大脑通过离散、级联的时间尺度处理信息。然而，支撑这一全局架构的回路机制仍未明了，尤其是考虑到局部回路中观察到的时间常数具有异质性。在此，我们提出，局部时间常数的这种异质性不仅能够共存，还能自发形成一个逐渐扩展的处理时间尺度层级结构。

我们使用 SITHCon 对此进行了验证。SITHCon 是一种深度前馈网络，其中每一层都具有相同的内在时间常数分布，但受到约束，必须表征对“何事”在“何时”发生的压缩记忆。尽管细胞时间尺度缺乏内在梯度，训练过程中仍自发形成了时间感受窗的功能层级结构。当输入序列以不同粒度被打乱时，浅层仅对细粒度扰动（例如交换字母）敏感，而深层不仅对局部抖动敏感，也会对广泛叙事结构（例如“段落”）遭到破坏作出反应，这与人类神经影像学研究中观察到的功能组织相呼应（Hasson, Yang, Vallines, Heeger, & Rubin, 2008年；Lerner 等，2011年）。这一发现与慢特征分析（Wiskott & Sejnowski, 2002年）的原理一致，指出局部时间异质性是将快速感觉流“重整化”为稳定、抽象叙事结构所必需的数学基底。

循环神经网络中尺度不变性的归纳先验

尽管前馈架构能够捕捉这些原理，大脑却必须通过循环动力学来实现它们。我们将受记忆神经科学启发的归纳先验系统地引入循环神经网络框架。首先，我们采用乘积隐藏状态，为每个特征设置独立的时间空间，从而强制形成可分离的表征。这要求权重矩阵呈块对角结构，使循环项能够分解为$\mathbf{R}_{\texttt{what}} \otimes \mathbf{R}_{\texttt{when}}$，从而形成对特征历史的模块化记忆。

然而，我们发现，仅有模块化不足以实现稳健的时间处理。通过将网络分析为一个由循环权重$\mathbf{R}$ 和下游读出权重$\mathbf{L}$构成的动力系统，我们证明，只有满足严格的谱约束，网络才能生成尺度不变序列活动：$\mathbf{R}$必须具有间隔呈几何级数的简并特征值（实现对数压缩），而$\mathbf{L}$ 必须采用平移的模式结构（实现平移等变性）（Liu & Howard, 2020）。

由此形成的架构 SITH-RNN 呈现出双群体动力学，类似于生物学中的时间上下文细胞（位于隐藏状态中）和时间细胞（位于读出中）。

这些约束使网络能够对分布外时间尺度进行零样本泛化——这是通用循环神经网络无法实现的——同时所需的可训练权重数量大幅减少。这证实了生物时间细胞特有的几何时间尺度分布，是实现尺度不变计算的一种规范性解决方案。

将内在动力学与功能处理时间尺度区分开来

大脑如何弥合毫秒级细胞过程与叙事结构缓慢展开之间的差距？本文认为，这种层级结构源于网络架构，而不仅仅是单神经元特性；功能层级的形成，关键取决于读出连接的拓扑组织，而非内在时钟的减速。

对内在时间尺度与涌现时间尺度进行实验性分离。在自然优化的生物回路（或标准部署的 SITH-RNN）中，通常会预期内在时间常数的分布随层级结构而变化，较深层能够访问逐渐变长的时间尺度 τ_{max} (Jacques、Tiganj、Howard 和 Sederberg，2021年；Jacques等，2022)。然而，在本研究中，我们有意限制模型，使所有层的内在时间常数分布都完全相同。这一实验控制至关重要：它确保观察到的时间感受窗层级结构并非仅仅由细胞特性预先设定，而必须作为网络架构的一种涌现功能属性而产生。事实上，尽管存在这一限制，深层仍涌现为“句子”级和“段落”级处理器。

慢动态的生物物理不可知论。这些发现对这样一种观点提出了挑战：时间感受窗（TRWs）的皮层层级结构完全由细胞内在特性的梯度所决定，例如受体动力学或膜时间常数（Murray 等人，2014年；Chaudhuri、Knoblauch、Gariel、Kennedy 与 Wang，2015年）。尽管这类解剖梯度已有充分记录（Siegle 等人，2021年），我们的结果表明，它们可能不足以解释大脑中所观察到的功能分离的全部幅度。事实上，近期来自雪貂听觉皮层的电生理证据证实，整合窗在不同区域之间呈层级组织——其驱动因素是回路拓扑，而非皮层层，表明仅凭局部层状梯度无法解释叙事时间尺度的出现（Sabat 等人，2025年）。

此外，本文推导出的数学原理并不依赖于这些时间尺度的特定生物物理来源。在我们的模型中，慢动力学源于循环矩阵**R**的特征值。然而，在生物回路中，这种动力学不一定仅依靠循环反馈环路；它们也可以由内在的细胞机制产生，例如钙激活非特异性阳离子电流（Frans´en、Alonso 和 Hasselmo，2002年；Liu、Tiganj、Hasselmo 和 Howard，2019年）或慢突触动力学（Guo 等，2021年）。事实上，未耦合的对角线结构SITH-RNN在功能上等同于一组由细胞内过程驱动的并行神经元。

拓扑组织驱动时间重整化。这一层级结构的出现表明，从感觉到叙事的转变是由时间线的结构性重整化驱动的，而不仅仅是单神经元时钟的减速。通过对时间模式进行最大池化，深层明确舍弃了事件的精确时间戳，转而保留其语义身份。关键在于，这意味着仅生成缓慢时间尺度是不够的；真正的关键约束在于读出的组织，它实际上实现了第4.2节所述的平移等变模体。事实上，共享或冗余的功能连接对于稳健的时间处理至关重要（Machens、Romo 与 Brody，2010 年）。因此，大脑中的解剖连接不可能是无结构的；它们必须遵循时间记忆的拓扑结构，以确保跨时间尺度的连续性。无论这种连续性是通过连续吸引子网络中的主动放电（Daniels 与 Howard，2025年；Sarkar 等人，2024）维持，还是通过活动静默的突触痕迹（Stokes，2015年；Mongillo、Barak 与 Tsodyks，2008年）维持，系统都必须遵守这些几何约束，才能支持将感觉信息重整化为叙事结构。

生物记忆与机器记忆中的趋同机制和分异缩放定律

大脑用于存储时间历史的机制与人工智能中的现代序列建模技术之间正日益趋同。在本节中，我们追溯机器学习中时间基函数的演变——从早期循环架构到现代状态空间模型。我们强调它们在结构上与生物记忆回路趋同，但在谱特性上存在分异。

机器学习中时间基函数的演变。循环神经网络长期以来一直难以建模长程时间依赖，主要原因是梯度消失问题。尽管长短期记忆（LSTM）网络等架构缓解了这一问题（Hochreiter & Schmidhuber，1997年），但它们缺乏生物学合理性，在捕捉跨越极长时间尺度的依赖关系时往往仍然表现不佳。

为解决这一问题，机器学习研究日益转向这样一类模型：将历史投影到一个固定的高维时间基上——这种方法在结构上让人联想到储备池计算（Maass 等，2002年；Jaeger，2001年）。值得注意的是，勒让德记忆单元（LMUs）（Voelker、Kaji´c 和 Eliasmith，2019年）以及 HiPPo框架（Fu 等，2022年）都试图利用正交多项式基来近似函数历史。

这种多项式投影方法逐渐发展成广义的状态空间模型（SSM；参见 Patro & Ag-

neeswaran，2024年），其遵循以下一般方程：

$$\mathbf{h}_t^{SSM} = \mathbf{A}\mathbf{h}_{t-1}^{SSM} + \mathbf{B}x_t \,, \tag{8}$$

$$y_t = \mathbf{C}\mathbf{h}_t^{SSM} \tag{9}$$

其中，**A** 演化隐藏状态，而**B**、**C** 是投影算子。关键在于，这些模型通常会为每个特征分别实例化独立的状态空间动力学，从而形成与我们提出的块对角先验相呼应的结构（见图 8）。

共享的对角架构，不同的时间基。受生物约束的 SITH-RNN 与 S4D（Gu、Goel、Gupta 和 R´e，2022年）或 Mamba（Gu 和 Dao，2023年）等现代状态空间模型之间，在形式上有着显著的一致性。这两种架构都通过对角化递归矩阵来解决时间整合问题，实质上是将复杂的历史信息分解为彼此独立、相互正交的记忆通道。[8] 这表明，时间模式的解耦是生物进化与机器学习工程所共有的一项基本计算原理。

然而，尽管机制已经趋同，近期的 S4D（Gu等人，2022年）和 Mamba（Gu 和 Dao，2023年）等模型也开始采用 SITH-RNN 固有的对角初始化，谱却仍存在根本差异。标准状态空间模型为了优化记忆重建而优先采用多项式基（例如 HiPPo 中的勒让德多项式），但这些基在数学上受限于固定区间尺度（Gu等人，2021年）。相比之下，生物回路采用尺度不变的几何特征谱。由此形成的是对数比率尺度，而非线性尺度；这与韦伯-费希纳定律相符。该定律描述了物理刺激强度与内部感觉之间的对数关系（Fechner，1860/1912年）。尽管 Mamba 的近期创新提升了表达能力，但它缺少这种归纳先验，因此无法天然支持零样本时间尺度泛化。

迈向稳健、尺度不变的机器智能

我们认为，韦伯-费希纳定律并非单纯的生物学特征，而是实现稳健序列建模所必需的一项基本计算约束。具体而言，采用对数分布的分格方式，记忆时域便可随着神经元数量的线性增长而呈指数级扩展，从而提供一种几何机制，绕过当前大语言模型的上下文长度瓶颈。

这些上下文长度瓶颈很大程度上源于底层自注意力架构固有的二次复杂度——无论是在训练、推理还是记忆使用方面（Vaswani et al., 2017）。近期的研究开始将生物先验与机器学习相结合，以应对这一局限。相关方法要么将输入历史预处理为尺度不变的时间线（Dickson & Tiganj, 2025），要么将记忆时间线直接嵌入注意力模块，构建“时间局部”的 Transformer（Dickson, Mochizuki-Freeman, Kabir, & Tiganj, 2025）。这些方法表明，经过对数压缩的 SITH记忆能够扩展现代人工智能的有效上下文容量，同时避免完整自注意力带来的高昂成本。

此外，这一先验为自动语音识别提供了有理有据的解决方案：与需要大量训练数据才能处理不同语速的标准架构（例如 wav2vec；Baevski、Zhou、Mohamed 和 Auli，2020年）不同，尺度不变动力学无需重新训练，便能自然适应更快或更慢的输入。近期研究证实，这种计算能力与雪貂听觉皮层的生物学实际相吻合：即使语音的信息速率发生变化，神经元整合窗口仍保持稳定，不受语境影响（Sabat 等，2025年）。这表明，大脑与 SITH-RNN 类似，依赖的是固定时间滤波器上的群体编码，而非动态调整单个细胞时钟的时间尺度。因此，将这一生物学约束与现代状态空间模型的高效性结合起来，有望兼具机器学习的计算可扩展性与生物智能的时间稳健性。

方法

层级语言语料库

我们使用确定性语法构建了一个层级数据集，以模拟自然语言的嵌套结构。词汇表由符号 {1, ... 9} 组成，代表最低层级（‘字母’）。这些符号递归组合成更高阶的结构：三个‘字母’组成‘单词’，三个‘单词’组成‘句子’，依此类推。这个组合过程依赖步进位置；例如，将第一个、第四个和第七个位置上的符号组合起来，形成下一个层级（‘单词’）的第一个条目（147）。同样，将第一个、第四个和第七个单词（分别为147、267和357）组合起来，形成下一层级（‘句子’）中的第一个条目（见图3b）。

三元组内的符号转移遵循一种确定性的、依赖上下文的语法，旨在模拟异或逻辑：第三个符号由前两个符号唯一确定，但无法仅凭其中任意一个符号预测。为避免位置聚类（例如，符号1总是出现在层级单元的开头）并确保结构多样性，各层级组合之前都会随机打乱符号列表。

由于这些规则具有递归性，因此可以生成无限多个层级；当到达更高层级的边界时，较低层级的符号转移会突然改变。在本研究中，最终序列由四 个层级构成，总共

[8] SITH-RNN 的演化方程（方程 3–4）在形式上类似于离散化状态空间模型形式。具体而言，状态空间模型递归矩阵**A** 对应于我们的对角衰减矩阵 exp(–sΔ)；输入投影**B** 对应于我们的权重**I**（按衰减进行缩放）；输出**C** 对应于我们的读出**L**。

表1循环网络架构与约束概述

模型	递归矩阵 **R**	特征值 λ	读出 **L**
1. 通用循环神经网络	稠密	随机	稠密
2. 块对角	块对角	随机	稠密
3. 对角（均匀）	块对角	均匀间隔	稠密
4. 对角（几何）	块对角	几何间隔	稠密
5. SITH-RNN	块对角	几何间隔	带状（托普利兹）

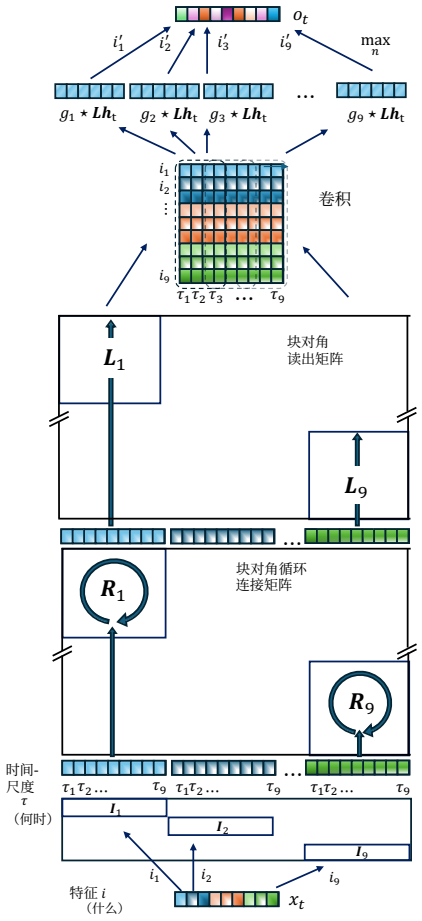

图*9*。 *SITH-RNN*示意图。我们考察了一系列权重呈块对角结构的循环神经网络，每个特征都保留一份压缩记忆。由相同块组成的投影矩阵$\mathbf{I}_{when}$生成隐藏状态，为每个特征分别保留记忆（When）。每个特征的记忆独立演化，并使用各自的循环块$\mathbf{R}_{when}$和读出块$\mathbf{L}_{when}$。将时间子空间重新堆叠成“What”×“When”矩阵后，卷积会追踪模式，而最大池化操作沿时间维度进行压缩，为下一层生成重映射特征。

包含 $3^4 = 81$ 个符号。网络经过训练，可以将九种可能的输入序列分类到相应的标签。这项任务要求网络跨多个时间尺度追踪符号，类似于保留对词语、段落以及理解语音所需的更大范围上下文的记忆

*SITHCon*架构

SITHCon架构（Jacques等，2022年）用作前馈基线。时间记忆通过一组滤波器$f_i(t)$生成，这些滤波器近似于对数时间线。这些滤波器通过对函数$F(s) = s^{-k}$（Post近似）进行逆拉普拉斯变换构造，其中k 控制时间分辨率。所有实验中均采用$k = 15$。记忆滤波器组由$n_\tau = 50$ 个滤波器组成，其时间常数 τ_i 在 $\tau_{min} = 1$和 $\tau_{max} = 81$之间按几何间隔排列。

这种设置生成了尺度不变时间历史（SITH）表征，其中，对输入$x(t) \rightarrow x(\alpha t)$进行时间重缩放，会使记忆模式沿神经元索引$i$发生平移。随后，稠密卷积层对这一记忆表征进行处理。由于卷积操作具有平移等变性，网络能够学习出对输入模式绝对持续时间不变的特征。为便于与循环神经网络进行比较，我们移除了最终的ReLU非线性和线性投影层，将学习限制在卷积核上。

网络架构与训练

我们训练了一系列循环神经网络（从通用模型到SITH-RNN）以及前馈 SITHCon 模型。循环和前馈架构均由 4 层构成，以匹配玩具语言中的层级数量。每层包含 9 个输入通道（特征），并为每个特征设置$n_\tau = 50$的时间空间，因此隐藏状态维度总计为$N_{\mathtt{hid}} = 9 \times 50 = 450$ 维。对于网络 2 至 5 和 SITHCon，读出之后连接一个卷积层，其卷积核宽度为$g_{\mathtt{width}} = 1$（覆盖单个时间常数），并通过最大池化操作为下一层创建重映射特征。

所有模型均使用 PyTorch 实现，并在层级语言分类任务上进行训练。我们采用 *AdamW* 优化器，并将权重衰减参数设置为$wd = 0.001$，以防止过拟合。训练进行了$n_{\mathtt{epoch}} = 200$个训练轮次。模型在序列上进行训练

在标准时间尺度（1×）下进行，该时间尺度对应于基本符号持续时间。

时间递归分析

为量化层表征的时间结构，我们计算了输出活动$\mathbf{o}_t^l$的自相似矩阵。任意两个时间点t_i和tj之间的相似性定义为它们群体向量之间的皮尔逊相关系数：

$$\rho_{\mathbf{o}(t_i),\mathbf{o}(t_j)} = \frac{(\mathbf{o}(t_i) - \langle \mathbf{o}(t_i) \rangle) \cdot (\mathbf{o}(t_j) - \langle \mathbf{o}(t_j) \rangle)}{\| \mathbf{o}(t_i) - \langle \mathbf{o}(t_i) \rangle \|_2 \ \left\| \mathbf{o}(t_j) - \langle \mathbf{o}(t_j) \rangle \right\|_2} \tag{10}$$

其中，$\langle \mathbf{o}(t) \rangle$表示时间$t$时神经元的平均活动。该指标等同于经 z 分数标准化的群体向量的余弦相似度。

通过置换序列测量层对不同时间尺度的敏感性

为量化每一层的有效时间递归窗口，我们采用了时间扰乱扰动分析。我们以不同粒度s对输入序列x进行置换，粒度范围从单个符号（3^0）到整个序列（3^6）。在此分析中，我们将 9 条分类序列拼接起来，构造长度为$N = 729$的较长连续流。令$P_s[x]$表示置换后的序列。我们将该置换序列输入网络，并记录层激活$\mathbf{o}^l(P_s[x])$。

为了将这些激活与原始响应$\mathbf{o}^l(x)$进行比较，我们通过对激活的时间维度应用逆置换P_s^{-1}，实质上“撤销”了输出的打乱。随后，我们计算了原始响应与撤销打乱后的响应的递归矩阵（ρ）之间的相关性：$C(\rho^l, \rho^{l\,\mathtt{shuff}(s)})$。较高的相关性表明，该层的表征对尺度$s$上的扰乱具有鲁棒性，这意味着其整合窗口小于$s$。相反，相关性下降则表明该层依赖于该尺度上的时间结构。为确保统计稳健性，我们针对每个尺度重复此分析，并使用$n_{\mathtt{shuff}} = 200$个随机置换。

循环网络序列

我们系统地约束通用循环神经网络（RNN）公式（方程3），以检验特定的生物先验。各架构变体汇总于表1。各层之间共享权重，这与 SITHCon架构一致，从而分离出递归的影响。

通用循环神经网络：为确保与线性SITH循环神经网络进行直接的架构比较，通用循环神经网络（模型 1）被实现为深度线性网络。它具有与块对角网络相同的隐藏维度（$N_{hid} = 450$），但采用稠密、全连接的循环矩阵。权重从均匀分布中初始化，并按层大小的平方根倒数进行缩放（He初始化的变体）；时间步之间不应用逐元素非线性。

实现细节

连续时间方程的离散化。我们将 Liu and Howard (2020) 提出的尺度不变性连续时间约束用于离散时间模拟（步长 Δt），通过对角矩阵$\mathbf{R}_{\text{when}} = \text{diag}\left(e^{-\lambda_1 \Delta_t}, \ldots, e^{-\lambda_n \Delta_t}\right)$来模拟隐藏状态演化。此外，我们将输入建模为步长起始时的瞬时脉冲，而非持续信号。因此，对输入投影进行缩放，以反映该时间区间内不同时间常数对应的异质衰减：

$$\mathbf{I}_{when} = \left[\exp(-\lambda_1 \Delta t) \cdots \exp(-\lambda_n \Delta t)\right].$$

特征值在 **R** 中的分布。对于对角线模型（3–5），每一层的时间常数分布相同。对于均匀模型（3），特征值分布为 $\lambda_i = 1/(\tau_{min} + \alpha i)$，从而实现时间常数的线性间隔。对于几何模型（4 和 5），时间常数 $\tau_i = 1/\lambda_i$ 按照以下方式分布：

$$\tau_i = \tau_{min} \left(\frac{\tau_{max}}{\tau_{min}} \right)^{\frac{i}{n_\tau - 1}}$$

这确保了n_τ 个神经元在对数时间上均匀地平铺 $\tau_{min} = 1$ 至 $\tau_{max} = 81$。

其中，τ_{min} 是最小时间常数，并且c 的取值使n_τ 个神经元在对数时间上均匀地平铺 τ_{min} 至 τ_{max}。每一层都具有相同的时间常数分布。

读出矩阵 **L**的具体约束。为实现 Liu and Howard (2020) 提出的平移特征向量约束，我们通过限制$\mathbf{L}_{\text{when}}$的形式，将*SITH-RNN* 与先前的分块对角实现区分开来。$\mathbf{L}_{\text{when}}$并非完全可训练，而是被限定为带状托普利兹型结构，从而确保有效的循环矩阵支持平移特征向量。

我们使用沿对角线滑动的重复基元来构造$\mathbf{L}_{\text{when}}$。该基元的宽度L_{width}是一个超参数，相对于时间常数n_τ的数量而言保持较小。基元中的数值可训练、在各层之间共享，并且受限于总和为零。例如，当$L_{\text{width}} = 3$ 且$n_\tau = 6$时，矩阵形式如下：

$$\mathbf{L}_{\text{when}} = \begin{bmatrix} b & c & & & & \\ a & b & c & & & \\ & a & b & c & & \\ & & a & b & c & \\ & & & a & b & c \\ & & & & a & b \end{bmatrix}.$$

在此处报告的*SITH-RNN* 这一具体情形中，我们采用宽度固定为$L_{\text{width}} = 7$的完全可训练基元。

作者注

本研究获得了 Rajen Kilachand 基金授予 Michael Hasselmo 的资助，以及授予 MWH 的 NIMH R01MH132171 资助。MWH 是 Cognitive Scientific AI 的联合创始人，该公司可能会从本研究的发表中间接受益。

参考文献

Atanas, A. A., Kim, J., Wang, Z., Bueno, E., Becker, M., Kang, D., . . . others（2023）。秀丽隐杆线虫中跨越多种时间尺度和状态的全脑行为表征。《细胞》，*186*(19), 4134–4151。Baevski, A., Zhou, Y., Mohamed, A., & Auli, M.（2020）。wav2vec 2.0：一种用于语音表征自监督学习的框架。神经信息处理系统进展，*33*，12449–12460。Bernacchia, A., Seo, H., Lee, D., & Wang, X. J.（2011）。皮层神经元中用于记忆痕迹的时间常数储备。《自然-神经科学》，*14*(3), 366-72。Bright, I. M., Meister, M. L. R., Cruzado, N. A., Tiganj, Z., Buffalo, E. A., & Howard, M. W.（2020）。猴内嗅皮层中以时间常数谱记录过去。《美国国家科学院院刊》，*117*，20274 20283-。Cao, R., Bladon, J. H., Charczynski, S. J., Hasselmo, M., & Howard, M.（2022）。啮齿动物海马体中内生生成的时间呈对数压缩。*eLife*，*https://doi.org/10.7554/eLife.75353*。Cavanagh, S. E., Hunt, L. T., & Kennerley, S. W.（2020）。内在时间尺度的多样性构成神经计算的基础。《神经回路前沿》，*14*，81。Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H., & Wang, X.-J.（2015）。灵长类皮层中层级动态处理的大规模回路机制。《神经元》，*88*(2), 419–431。Chien, H.-Y. S., & Honey, C. J.（2020）。在人类大脑皮层中构建和遗忘时间上下文。《神经元》，*106*(4), 675–686。Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X. J.（2000）。皮层网络模型中支撑空间工作记忆的突触机制和网络动力学。《大脑皮层》，*10*(9), 910-23。Cruzado, N. A., Tiganj, Z., Brincat, S. L., Miller, E. K., & Howard, M. W.（2020）。在联想记忆任务中，猴海马体和外侧前额叶皮层对内容与时间的联合表征。《海马体》，*30*，1332-1346。Dahmen, D., Gr¨un, S., Diesmann, M., & Helias, M.（2019）。大脑中的第二类临界性揭示了丰富的多神经元动力学。《美国国家科学院院刊》，*116*(26), 13051–13060。Daniels, B. C., & Howard, M. W.（2025）。用于拉普拉斯神经流形的连续吸引子网络。计算脑与 & 行为，1–18。Danskin, B. P., Hattori, R., Zhang, Y. E., Babic, Z., Aoi, M., & Komiyama, T.（2023）。压后皮层中对多种时间尺度进行指数历史整合，支持双曲行为。《科学进展》，*9*(48), eadj4897。检索自 https://www.science.org/doi/abs/10.1126/sciadv.adj4897 DOI：10.1126/sciadv.adj4897 Dickson, B., Mochizuki-Freeman, J., Kabir, M. R., & Tiganj, Z.（2025）。时间局部变换器。计算脑科学&行为，1–13。Dickson, B., & Tiganj, Z.（2025）。渐进式遗忘：通过对数压缩扩展变换器上下文窗口。*arXiv*预印本 *arXiv:2510.22109*。Ding, X., Froudist-Walsh, S., Jaramillo, J., Jiang, J., & Wang, X.- J.（2024）。细胞类型特异性连接组可预测小鼠脑中的分布式工作记忆活动。*eLife*，*13*，e85442。Egorov, A. V., Hamam, B. N., Frans´en, E., Hasselmo, M. E., & Alonso, A. A.（2002）。内嗅皮层神经元的分级持续性活动。《自然》，*420*(6912), 173-8. Elman, J. L.（1990）。发现时间中的结构。《认知科学》，*14*, 179-211. Fascianelli, V., Tsujimoto, S., Marcos, E., & Genovesio, A.（2019）。猕猴背外侧前额叶皮层（而非眶额或极部前额叶皮层）的自相关结构可预测视觉线索策略任务中的反应编码强度。《大脑皮层》，*29*(1), 230–241. Fechner, G.（1860/1912）。《心理物理学纲要》第*I*卷。霍顿·米夫林。 Feng, L., Tung, F., Ahmed, M. O., Bengio, Y., & Hajimirsadeghi, H.（2024）。循环神经网络就是我们所需的一切吗？*arXiv*预印本 *arXiv:2410.01201*。Frank, M. J., Rudy, J. W., Levy, W. B., & O'Reilly, R. C.（2005）。逻辑失效之时：人类的隐性传递性推理。记忆& 认知，*33*(4), 742-50. Frans´en, E., Alonso, A. A., & Hasselmo, M. E.（2002）。模拟毒蕈碱受体激活的钙敏感性非特异性阳离子电流 INCM 在延迟匹配任务期间内嗅皮层神经元活动中的作用。《神经科学杂志》，*22*(3), 1081-1097. Fu, D. Y., Dao, T., Saab, K. K., Thomas, A. W., Rudra, A., & R´e, C.（2022）。饥饿的河马：迈向使用状态空间模型进行语言建模。*arXiv*预印本 *arXiv:2212.14052*。Graves, A., Mohamed, A.-r., & Hinton, G.（2013）。使用深层循环神经网络进行语音识别。收录于*2013 IEEE*国际声学、语音与信号处理会议（第6645–6649页）。Gu, A., & Dao, T.（2023）。Mamba：利用选择性状态空间进行线性时间序列建模。*arXiv*预印本 *arXiv:2312.00752*。Gu, A., Goel, K., Gupta, A., & R´e, C.（2022）。对角状态空间模型的参数化与初始化。神经信息处理系统进展，*35*, 35971–35983. Gu, A., Goel, K., & R´e, C.（2021）。使用结构化状态空间高效建模长序列。*arXiv*预印本 *arXiv:2111.00396*。Guo, C., Huson, V., Macosko, E. Z., & Regehr, W. G.（2021）。代谢型信号传导的分级异质性构成了单极刷状细胞内在时间响应的连续谱基础。《自然·通讯》，*12*(1), 1–12. Hasselmo, M. E., Giocomo, L. M., Brandon, M. P., & Yoshida, M.（2010）。情景记忆中沿时空轨迹对事件时间和地点进行编码的细胞动力学机制。《行为脑研究》，*215*, 261–274. Hasson, U., Chen, J., & Honey, C. J.（2015）。层级过程记忆：记忆是信息处理不可或缺的组成部分。《认知科学趋势》，*19*(6), 304–313. Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N.（2008）。人类大脑皮层时间感受窗的层级结构。《神经科学杂志》，*28*(10), 2539-50. doi: 10.1523/JNEUROSCI.5487-07.2008 Hebb, D. O.（1949）。《行为的组织》。纽约：威利出版社。Hochreiter, S., & Schmidhuber, J.（1997）。长短期记忆。《神经计算》，*9*(8), 1735–1780. Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., . . . Hasson, U.（2012）。缓慢皮层动力学与长时间尺度上信息的累积。《神经元》，*76*(2), 423-34. doi: 10.1016/j.neuron.2012.08.011 Howard, M. W., & Eichenbaum, H.（2013）。跨尺度的海马体、时间与记忆。《实验心理学杂志：总论》，*142*(4), 1211-30. doi: 10.1037/a0033621

Howard, M. W., & Hasselmo, M. E. (2020). 在拉普拉斯域中利用时间和空间的神经表征进行认知计算。*arXiv*预印本 *arXiv:2003.11668*。Howard, M. W., Luzardo, A., & Tiganj, Z. (2018). 拉普拉斯决策空间中的证据累积。《计算脑与行为》, *1*, 237-251. Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014). 用于在海马区编码时间、空间与序列的统一数学框架。《神经科学杂志》, *34*(13), 4692-707. doi: 10.1523/JNEUROSCI.5808-12.2014 Jacques, B. G., Tiganj, Z., Howard, M. W., & Sederberg, P. B. (2021). DeepSITH：通过跨时间尺度分解“什么”与“何时”实现高效学习。收录于 M. Ranzato, A. Beygelz- imer, P. Liang, J. Vaughan, & Y. Dauphin (编), 第*35*届神经信息处理系统进展会议（卷 arXiv:2104.04646）。Jacques, B. G., Tiganj, Z., Sarkar, A., Howard, M., & Sederberg, P. (2022). 对时间重缩放具有不变性的深度卷积神经网络。收录于国际机器学习会议（第9729–9738页）。Jaeger, H. (2001). 分析和训练循环神经网络的“回声状态”方法——附勘误说明。波恩，德国：德国国家信息技术研究中心 *GMD* 技术报告，*148*(34), 13. Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). 皮质-基底 ganglia 回路中的时间神经表征。《美国国家科学院院刊》, *106*(45), 19156–19161. Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). 利用叙述故事绘制时间感受窗层级结构的拓扑图。《神经科学杂志》, *31*(8), 2906-15. doi: 10.1523/JNEUROSCI.3684-10.2011 Liu, Y., & Howard, M. W. (2020). 线性循环网络中尺度不变序贯活动的生成。《神经计算》, *32*(7), 1379–1407. Liu, Y., Levy, S., Mau, W., Geva, N., Rubin, A., Ziv, Y., . . . Howard, M. (2022). 小鼠海马中持续数分钟尺度的群体活动。海马体，*32*(5), 359–372. Liu, Y., Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2019). 一种用于可扩展尺度不变时间表征的神经微回路模型。海马体，*29*(3), 260–274. Maass, W., Natschl¨ager, T., & Markram, H. (2002). 无稳定状态的实时计算：一种基于扰动的神经计算新框架。《神经计算》, *14*(11), 2531-60. doi: 10.1162/089976602760407955 MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). 海马“时间细胞”弥合了不连续事件记忆中的空缺。《神经元》, *71*(4), 737-749. Machens, C. K., Romo, R., & Brody, C. D. (2010). 前额叶皮层中“什么”与“何时”的功能性而非解剖学分离。《神经科学杂志》, *30*(1), 350–360. Mongillo, G., Barak, O., & Tsodyks, M. (2008). 工作记忆的突触理论。《科学》, *319*(5869), 1543–1546. Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wal- lis, J. D., Cai, X., . . . others (2014). 灵长类皮层内在时间尺度的层级结构。《自然-神经科学》, *17*(12), 1661. Ning, W., Bladon, J. H., & Hasselmo, M. E. (2022). 前额叶皮层与海马体中时间的互补表征。海马体。Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). 大鼠海马中内部生成的细胞组装序列。《科学》, *321*(5894), 1322-7。

Patro, M. N., & Agneeswaran, V. S. (2024). Mamba-360：作为 Transformer 替代方案的状态空间模型综述：长序列建模的方法、应用与挑战。*arXiv* 预印本 *arXiv:2404.16112*。Rajan, K., & Abbott, L. F. (2006). 神经网络随机矩阵的特征值谱。《物理评论快报》, *97*(18), 188104。Rajan, K., Harvey, C. D., & Tank, D. W. (2016). 序列生成与记忆的循环网络模型。《神经元》, *90*(1), 128-42。doi: 10.1016/j.neuron.2016.02.009 Rossi-Pool, R., Zizumbo, J., Alvarez, M., Vergara, J., Zainos, A., & Ro mo, R. (2019). 背侧运动前皮层中支撑认知过程的时间信号。《美国国家科学院院刊》, *116*, 7523–7532。Sabat, M., Gouyette, H., Gaucher, Q., Espejo, M. L., David, S. V., Norman-Haignere, S., & Boubenec, Y. (2025). 听觉皮层中的神经元在受限且情境不变的时间窗口内整合信息。《当代生物学》。Salet, J. M., Kruijne, W., van Rijn, H., Los, S. A., & Meeter, M. (2022). FMTP：跨时间尺度时间准备的统一计算框架。《心理学评论》。Sarkar, A., Wang, C., Zuo, S., & Howard, M. W. (2024). 使用拉普拉斯神经流形表征“什么”与“何时”的工作记忆。*arXiv* 预印本 *arXiv:2409.20484*。Schonhaut, D. R., Aghajan, Z. M., Kahana, M. J., & Fried, I. (2022). 时空情境的神经编码。*bioRxiv*。Shankar, K. H., & Howard, M. W. (2013). 最优模糊时间记忆。《机器学习研究杂志》, *14*, 3753-3780。Sheffield, M. E., Best, T. K., Mensh, B. D., Kath, W. L., & Spruston, N. (2011). 缓慢整合使耦合中间神经元远端轴突中的动作电位持续放电。《自然-神经科学》, *14*(2), 200-7。Shikano, Y., Ikegaya, Y., & Sasaki, T. (2021). 海马-纹状体环路中的分钟编码神经元。《当代生物学》, *31*(7), 1438–1449。Siegle, J. H., Jia, X., Durand, S., Gale, S., Bennett, C., Graddis, N., . . . others (2021). 对小鼠视觉系统尖峰活动的综述揭示了功能层级结构。《自然》, *592*(7852), 86–92。Smith, J. T., Warrington, A., & Linderman, S. W. (2022). 用于序列建模的简化状态空间层。*arXiv* 预印本 *arXiv:2208.04933*。Spitmaan, M., Seo, H., Lee, D., & Soltani, A. (2020). 神经动力学的多个时间尺度及任务相关信号在大脑皮层中的整合。《美国国家科学院院刊》, *117*(36), 22522–22531。Stokes, M. G. (2015). 前额叶皮层中的“活动静默”工作记忆：一种动态编码框架。认知科学趋势, *19*(7), 394–405。Sussillo, D., Churchland, M. M., Kaufman, M. T., & Shenoy, K. V. (2015). 一种能够为肌肉活动的产生找到自然解决方案的神经网络。《自然-神经科学》, *18*(7), 1025–1033。Taxidis, J., Pnevmatikakis, E. A., Dorian, C. C., Mylavarapu, A. L., Arora, J. S., Samadian, K. D., . . . Golshani, P. (2020). 情境特异性海马序列中感觉表征与时间表征的差异化出现及稳定性。《神经元》, *108*(5), 984–998.e9。Tiganj, Z., Cromer, J. A., Roy, J. E., Miller, E. K., & Howard, M. W. (2018). 猴前额叶皮层中近期经验的压缩时间线。《认知神经科学杂志》, *30*, 935-950。Tiganj, Z., Cruzado, N. A., & Howard, M. W. (2019). 迈向神经层级认知架构：用神经元对工作记忆任务中的行为建模

。见 A. Goel, C. Seifert, & C. Freksa (编), 第*41*届认知科学学会年会论文集 (第1118-1123页)。蒙特利尔：认知科学学会。Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2015). 单个神经元中长时间常数的简单且生物物理学上合理的模型。海马体, *25*(1), 27-37。Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B., & Moser, E. I. (2018). 外侧内嗅皮层从经验中整合时间信息。《自然》, *561*, 57-62。Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jo nes, L., Gomez, A. N., . . . Polosukhin, I. (2017). 你只需要注意力机制。见《神经信息处理系统进展》(第5998–6008页)。Voelker, A., Kaji´c, I., & Eliasmith, C. (2019). 勒让德记忆单元：循环神经网络中的连续时间表示。见《神经信息处理系统进展》(第15570–15579页)。Wasmuht, D. F., Spaak, E., Buschman, T. J., Miller, E. K., & Stokes, M. G. (2018). 内在神经元动力学可预测工作记忆期间不同的功能作用。《自然·通讯》, *9*(1), 1–13。White, O. L., Lee, D. D., & Sompolinsky, H. (2004). 正交神经网络中的短期记忆。《物理评论快报》, *92*(14), 148102。Wiskott, L., & Sejnowski, T. J. (2002). 慢特征分析：不变性的无监督学习。《神经计算》, *14*(4), 715-70。doi: 10.1162/089976602317318938 Wong, K.-F., Huk, A. C., Shadlen, M. N., & Wang, X.-J. (2007). 知觉决策中随时间变化的证据累积背后的神经回路动力学。《计算神经科学前沿》, *1*, 6。Zuo, S., Wang, C., Wang, L., Jin, Z., Kusunoki, M., & Kwok, S. C. (2024). 内侧后顶叶皮层中时间顺序记忆的神经标记。*bioRxiv*, 2023年8月。

补充信息（SI）

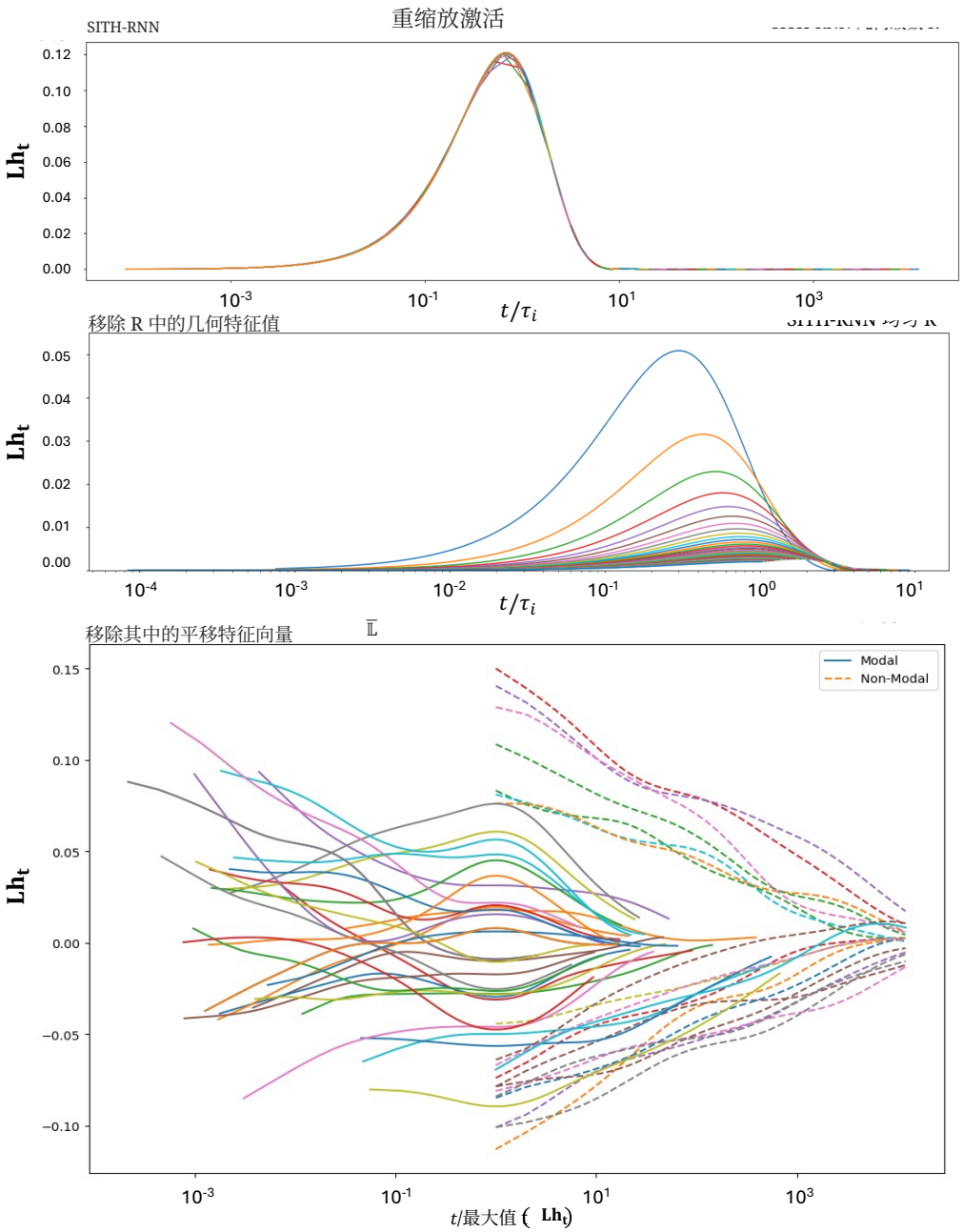

图 *S1*。 消融研究证实，尺度不变性需要几何特征值与平移等变读出的共同作用。我们分析了三种模型变体中，输出神经元群 $\mathbf{Lh}_t$ 对 $t = 0$ 时刻的狄拉克输入的脉冲响应。上图（*SITH-RNN*）：完整模型在**R** 中采用几何特征值，并在**L**中采用带状的平移等变读出。以重缩放时间（t/τ_i）为横轴作图时，不同神经元的响应曲线汇聚为同一个通用函数，表明模型具有真正的尺度不变性。中图（均匀特征值）：将**R** 中的几何分布替换为均匀分布，虽然保留了序列性，却破坏了尺度不变性。重缩放后，曲线无法对齐，这证实对数压缩是实现缩放性质所必需的。下图（稠密读出）：保留几何特征值，但将带状**L** 替换为标准的稠密可训练矩阵，会导致无序动力学。神经元群分化为“模态”（峰状）和“非模态”（单调）响应，完全丧失了可解释的“时间细胞”样序列。因此，SITH-RNN 需要$\mathbf{R}_{when}$ 中的几何特征值和$\mathbf{L}_{when}$ 中的平移基元，才能产生尺度不变的序列动力学。



图*S2*。通过脉冲触发平均对线性感受野进行全面映射，揭示了逐层重整化和习得的组合层级。该图通过展示四个隐藏层中全部九个神经元的线性感受野，对图4d中的代表性示例进行了扩展。相对于五个层级尺度上符号的起始时刻，使用脉冲触发平均（STA）恢复感受野，其范围从基本的“字母”（3^0）到完整的“序列”（3^4）。该可视化突出了时间处理深度中的一项基本转变。左： 第1层神经元在最精细的时间尺度上呈现出尖锐的对角线带状结构，表明它们充当局部符号转换的精确检测器，但无法整合全局上下文。在更高层级（例如“序列”或“段落”）上，其感受野呈现出稠密、重复的平铺模式，显示出周期性。这表明早期层会随着每次局部符号转换而有效地“重置”，追踪反复出现的构成元素（例如特定字母三元组的每一次出现），但无法辨别这些元素在更广泛的叙事结构中的独特位置。右： 相比之下，第4层神经元（右列）呈现出与语法的组合规则相一致的嵌套结构。从第4层列的底部向顶部观察，在“序列”层级上的单个宽激活带，在“段落”层级上解析为三个不同的带，并进一步在“句子”层级上细分为九个带。这种类分形细分证实，深层神经元已经成功习得了层级之间的映射，通过构成符号的适当组合学来定义其感受野（例如，将特定序列识别为三个特定段落的组合）。中： 这两个极端之间的过渡并非二元的。中间列（第2层和第3层）呈现出丰富且连续的整合窗梯度，展现了网络进行混合选择性的能力，能够处理介于单个符号的速度与完整叙事持续时间之间的中间语言结构（例如短语）。