

复杂系统中的因果涌现研究综述^{*}

王志鹏¹⁾ 张江^{1,2)†}

(1)北京师范大学系统科学学院, 100875, 北京; (2)北京集智研究院, 100085, 北京)

摘要 聚焦复杂系统:就涌现、因果以及因果涌现进行了定量描述;阐述了因果与涌现的联系;系统梳理了基于粗粒化和信息分解2种定量刻画因果涌现的方法,以及基于信息分解和神经信息压缩2种因果涌现辨识方法;详细介绍了各方法的基本原理、优缺点,以及相关应用等。基于粗粒化方法定义了因果涌现框架,其可应用于离散动力系统;利用信息分解方法求解时结果会依赖冗余信息;基于信息分解所提出的指标可识别数据中的因果涌现,且找到一个充分条件;基于数据驱动的神经信息压缩方法可扩展并应用于连续动力系统,并可自动提取不同层级的粗粒化函数,以及构建不同层级的动力学,还可识别不同类型动力学系统的因果涌现;改进现有方法并用于识别更为复杂的系统 with 学习自动化分组,解决通用动力学大模型等问题。

关键词 复杂系统;因果涌现;有效信息;粗粒化;信息分解;神经信息压缩

中图分类号 N949

DOI: 10.12202/j.0476-0301.2023049

0 引言

复杂系统无处不在,从电网、交通、通信系统等基础设施网络到城市、社会、经济、生命甚至整个宇宙,这些都可以看作复杂系统^[1-6]。复杂系统是由许多相互作用的元素构成的。复杂系统组成元素之间以及系统和环境之间存在依赖、竞争、关联等复杂的相互作用,通过这些复杂的相互作用系统往往表现出非线性、涌现、自组织等性质^[7-10]。

涌现一直是复杂系统的一个重要特性和研究对象^[9,11]。1972年著名物理学家Anderson发表在*Science*的文章“More is Different”^[7]中给出了涌现的一般性定义,即“涌现是指系统上定量的变化导致行为上定性的变化”。涌现可简单理解为“整体大于部分之和”,即整体上展现了构成它的个体所不具备的新特性^[12-13]。复杂系统都存在着涌现特征,这使得学者们不能借助还原论的方法将复杂系统分解为一个个元素来理解。此外,Fromm^[14]从反馈视角出发,将涌现分为简单涌现、弱涌现、多重涌现和强涌现4类。

因果也是复杂系统的一个重要概念,它既描述了动力系统中因和果在时间上的联系,也描述了一个动力过程的因果效应^[15-16]。Neuberg^[17]进一步完善了因果科学理论,用数学框架量化因果,并把因果引入对涌现的衡量。涌现和因果实际上具有天然的联系,这

是因为虽然涌现有一般分类,但却无法定量刻画。学者们往往把很多东西归结为一些微观因素,但是宏观涌现出来的模式,往往无法归因到微观层面并进行解释,所以也就无法找到对应的原因^[14]。

因果涌现概念的定性描述为:当一个系统在宏观尺度相较其在微观尺度上具有更强的因果效应,此即产生了因果涌现。

在定性描述的基础上,已有研究提出了定量刻画因果涌现的方法,主要有如下2类。

1)基于粗粒化的方法^[18-19]以及基于信息分解的方法^[20],后者包括基于特有信息和协同信息2种。

2)因果涌现的识别问题。包括:Rosas等^[20]基于互信息提出的3个指标;Zhang等^[21]基于神经信息压缩方法,从数据中识别出系统的因果涌现。

因果涌现结合了人工智能中的因果和复杂系统中的涌现这2个核心概念,较好地刻画了系统宏观和微观状态之间的联系和区别。因果涌现为回答一系列的哲学问题提供了一个量化的视角,比如借助因果涌现框架讨论生命系统或社会系统中自上而下的因果特性等^[22]。因果涌现理论框架为定量理解涌现和其他重要概念提供了新的途径^[23-24]。

量化的因果涌现方法已被广泛应用于许多复杂系统中,例如:布尔网络^[18]、弹簧振子模型^[21]、具

^{*} 国家自然科学基金资助项目(61673070)

[†] 通信作者:张江(1978—),博士,教授,博士生导师。研究方向:复杂系统建模、因果涌现、规模法则、机器学习。E-mail: zhangjiang@bnu.edu.cn
收稿日期:2023-04-04

有成百上千节点的复杂网络^[25-27], 以及神经网络^[28]; 具有明显涌现现象的生命游戏^[20]、鸟群模型^[20]、蛋白质交互^[29]、生物^[30], 以及真实的大脑网络^[31]等。

1 因果涌现量化方法概述

本研究详细梳理了 2 种因果涌现的定义与 2 种因果涌现辨识方法的基本原理, 并且比较了各方法的优缺点。

1.1 因果涌现的定义

1.1.1 Hoel 的粗粒化方法 Hoel 等^[18]最早提出并定义了因果涌现框架, 本研究将其进行抽象, 如图 1 所示。

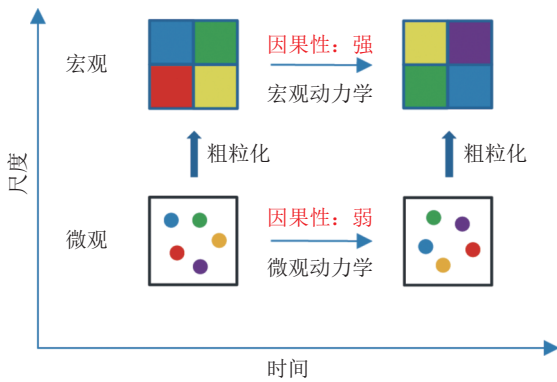


图 1 抽象的因果涌现框架

图 1 中, 横轴表示时间尺度, 纵轴表示空间尺度, 由于微观态往往具有很大的噪声, 导致微观动力学的因果性比较弱, 所以需要微观态进行粗粒化, 以便获得噪声较小的宏观态, 从而使得宏观动力学的因果性更强。

本研究借鉴了整合信息的量化方法^[32], 提出了一种因果效应度量指标 EI (effective information), 即有效信息, 用以量化一个马尔可夫动力学的因果性强弱。有效信息的计算借助因果理论框架^[17], 利用因果阶梯中的第 2 层干预方法对上一时刻的状态做干预, 再计算干预分布及干预情况下, 经过马尔可夫动力学的下一时刻分布之间的互信息, 作为因果效应的度量指标。EI 的计算式为

$$I_E(S) = I_M(I_D; E_D) = \sum_{i \in I_0} p(\text{do}(s_{t-1} = i)) \cdot \sum_{s_t \in E_0} p(s_t | \text{do}(s_{t-1} = i)) \log_2 \frac{p(s_t | \text{do}(s_{t-1} = i))}{p(s_t)}, \quad (1)$$

式中: I_M 为互信息; s_{t-1} 和 s_t 分别表示 $t-1$ 和 t 时刻的状态; $I_D = \text{do}(S_{t-1} \in U(I))$; $E_D = S_t | \text{do}(S_{t-1} \in U(I))$; do 表示对状态进行干预操作, 即设定上一时刻的状态 S_{t-1} 为均匀分布, 强行使 $S_{t-1} \in U(I)$, I 表示系统的状态空间, $U(I)$ 表示空间上的均匀分布。进行 do 操作可使得有

效信息能客观衡量动力学的因果特性而不受状态分布的影响。由于有效信息依赖状态空间的大小, 为了便于在不同尺度进行比较, 本文定义了一个归一化指标有效系数衡量动力学的因果性强弱。有效系数和有效信息的关系为

$$E_F(S) = \frac{I_E(S)}{\log_2 n}, \quad (2)$$

式中: n 表示系统的状态个数; $E_F(S) \in [0, 1]$ 。此外, 有效系数可以进一步分解, $E_F(S) = \langle s_0(\text{确定性}) \rangle - \langle s_0(\text{简并性}) \rangle$ 。确定性和简并性的计算式分别为

$$s_0(\text{确定性}) = \frac{1}{\log_2 n} \sum_{s_t \in E_0} p(s_t | \text{do}(s_{t-1} = s_0)) \cdot \log_2(n \cdot p(s_t | \text{do}(s_{t-1} = s_0))), \quad (3)$$

$$s_0(\text{简并性}) = \frac{1}{\log_2 n} \sum_{s_t \in E_0} p(s_t | \text{do}(s_{t-1} = s_0)) \cdot \log_2(n \cdot p(s_t)). \quad (4)$$

故此, 可通过提高系统的确定性和降低简并性增加系统的有效信息。

根据动力学因果效应的度量指标, 通过比较有效信息的大小判断是否发生因果涌现。通过选择合适的粗粒化策略, 可将微观状态转换成宏观状态, 如果宏观动力学的有效信息大于微观动力学的有效信息 ($I_E(S_M) > I_E(S_m)$), 则认为在该粗粒化尺度上的宏观动力学具有因果涌现特性。

给定一个包括 4 种状态马尔可夫链的状态转移矩阵, 其中前 3 种状态之间等概率转移, 最后 1 种状态是独立的, 通过将前 3 种状态粗粒化成一个状态, 可得到确定的宏观系统, 即

$$S_m = \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{粗粒化操作}} S_M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

此时 $I_E(S_m) = 0.81 \text{ bit} < I_E(S_M) = 1.00 \text{ bit}$, 系统发生了因果涌现。

在已有研究中学者们给了一些实例, 通过对微观变量进行分组, 定义一个从微观态到宏观态的映射函数, 即可得到因果涌现。而他们并没有给出一般的粗粒化策略, 因此, 他们的工作只是定义了动力学的因果效应度量指标, 无法实现判断系统能否发生因果涌现。

要找到使得因果涌现能够发生的粗粒化函数, 需

要搜索所有可能的粗粒化策略,且搜索空间巨大,这对于较大的系统特别是复杂的真实系统是不可行的。故此,如何设计有效的粗粒化策略是一个值得研究的问题。

2017年,Hoel^[19]从香农信息论的角度进一步解释了因果涌现:将其看作噪声信道编码的一种特例;寻找宏观动力学最大化有效信息的过程等价于搜索所有可能的输入分布,以便找到以一定速率沿信道传输的最大信息量,即信道容量。该方法并未解决有效粗粒化的策略选择。

1.1.2 信息分解方法 Rosas等^[20,34]从信息理论角度出发,提出了一种基于信息分解方法定义系统中的因果涌现。他们认为发生因果涌现有因果解耦(causal decoupling, CD)和向下因果(downward causation, DC)2种可能性,其中CD表示宏观态对其他宏观态的因果效应,DC表示宏观态对于微观元素的因果效应。具体地,用 $X_t(X_t^1, X_t^2, \dots, X_t^n)$ 表示微观状态输入, V_t 表示宏观状态是 X_t 的随附特征, X_{t+1} 和 V_{t+1} 分别表示下一时刻的微观和宏观状态。该方法基于Williams等^[33]提出的多元信息非负分解方法,他们将微观态 $X(X^1, X^2)$ 与宏观态 V 之间的互信息分解为4部分,计算式为

$$I(X^1, X^2; V) = \text{Red}(X^1, X^2; V) + \text{Un}(X^1; V|X^2) + \text{Un}(X^2; V|X^1) + \text{Syn}(X^1, X^2; V), \quad (5)$$

式中: $\text{Red}(X^1, X^2; V)$ 表示冗余信息,是2个微观态 X^1 和 X^2 同时向宏观态 V 提供的信息; $\text{Un}(X^1; V|X^2)$ 和 $\text{Un}(X^2; V|X^1)$ 均表示特有信息,是每个微观态单独向宏观态提供的信息,其中 X^1 的特有信息是指在另一个微观态 X^2 给定的情况下,微观态 X^1 向宏观态 V 提供的信息; $\text{Syn}(X^1, X^2; V)$ 表示协同信息,是将所有微观态 X 联合向宏观态 V 提供信息。

Williams等^[33]基于分解后的信息提出了如下2种因果涌现的定义方法:

1) 基于特定宏观状态 V_t , 当特有信息 $\text{Un}(V_t; X_{t+1}|X_t^1, \dots, X_t^n) > 0$, 表示当前时刻的宏观态 V_t 能超过当前时刻的微观态 X_t , 向下一时刻的整体系统 X_{t+1} 再多提供一些信息, 则存在因果涌现;

2) 绕开选择特定的宏观态 V_t , 仅基于系统当前时刻的微观态 X_t 和下一时刻的微观态 X_{t+1} 之间的协同信息定义因果涌现, 当协同信息 $\text{Syn}(X_t^1, \dots, X_t^n; X_{t+1}^1, \dots, X_{t+1}^n) > 0$, 且 $\text{Un}(V_t; X_{t+1}|X_t) \leq \text{Syn}(X_t; X_{t+1})$, 则系统发生了因果涌现。

值得注意的是, 对于方法1), 通过计算特有信息判断因果涌现是否发生依赖于宏观态 V_t 的选择, 然而实际又不知如何选择 V_t , 因此该方法一般不可行。如

果借助协同信息判断因果涌现是否发生, 但协同信息的计算依赖冗余信息, 冗余信息又存在计算问题^[33], 因此, 方法2)一般也不可行。相较而言, Hoel提出的基于粗粒化的方法定义因果涌现是一个更好的方案。

1.2 因果涌现识别 因果涌现识别是指从可观测数据中辨别系统能否发生因果涌现。如果能找到一个宏观尺度下动力学有效信息大于微观动力学的, 则认为系统发生了因果涌现; 反之, 如果遍历所有宏观尺度都无法找到“宏观打败微观”的效果, 系统则不会发生因果涌现。此外, 对于因果涌现识别往往需要给出发生因果涌现的最优尺度。下面介绍信息分解和神经信息压缩这2种因果涌现的识别方法。

1.2.1 信息分解方法 Rosas等^[20]忽略了对特有信息和协同信息的求解, 而提出了一个判定因果涌现的充分条件: 基于互信息给出了 Ψ 、 Δ 和 Γ 这3个新指标, 用于识别系统中的因果涌现。3个指标的关系分别为:

$$\Psi_{t,t+1}(V) := I(V_t; V_{t+1}) - \sum_{j=1}^n I(X_t^j; V_{t+1}), \quad (6)$$

当 $\Psi > 0$ 时, 宏观状态 V 会发生涌现; 当 $\Psi < 0$ 时, V 不确定是否会发生涌现。

$$\Delta_{t,t+1}(V) := \max_j \left(I(V_t; X_{t+1}^j) - \sum_{i=1}^n I(X_t^i; X_{t+1}^j) \right), \quad (7)$$

当 $\Delta > 0$ 时, 宏观状态 V 发生DC。

$$\Gamma_{t,t+1}(V) := \max_j I(V_t; X_{t+1}^j), \quad (8)$$

当 $\Psi > 0$ 且 $\Gamma = 0$ 时, 宏观状态 V 发生因果涌现且发生CD。

Rosas等^[20]将该方法应用于生命游戏和鸟群实验, 发现生命游戏实验存在因果涌现, 且存在CD; 在鸟群实验中, 固定聚集及对齐参数不变, 调节回避参数至中间值时能产生因果涌现的效果, 同时取0.05时, 因果涌现效果最强。

3个指标中都含有宏观状态 V , 但在具体的实验中, 如鸟群, 先固定宏观变量, 然后遍历其变量值, 获得不同的宏观态用以判断因果涌现是否发生, 所以该方法避开粗粒化策略。3个指标均是基于互信息计算, 而没有基于因果方法定义涌现, 且该方法获得的仅仅是发生因果涌现的充分条件, 是一种近似求解方法。此外, 该方法无法得到显式宏观动力学以及粗粒化策略, 然而这2项对于下游任务十分重要。与Hoel提出的基于粗粒化方法一样, 也需要为宏观和微观动力学定义明确的马尔可夫转移矩阵。对于微观输入为时序数据时, 需要从数据中估计状态转移矩阵; 当

输入是连续数据时,这种估计又会存在很大的偏差,因此,该方法也不是一种最优方法。

1.2.2 神经信息压缩方法 近年来,随着以神经网络为代表的机器学习方法的快速发展,其在药物的靶向预测、蛋白质结构预测等许多跨学科领域都取得了较好的效果^[35-39],这使得基于以数据驱动方式自动发现复杂系统中的因果关系甚至动力学成为可能^[40-44]。

同时通过机器学习,也可找到合适的粗粒度策略^[45-46]。Zhang 等^[21]借助神经网络提出了一种基于数据驱动的神信息压缩(NIS)方法,能够从时间序列数据中识别系统中的因果涌现,还可自动提取有效的粗粒度策略和宏观动力学,克服了信息分解方法的种种不足。NIS 方法的模型框架如图 2 所示。

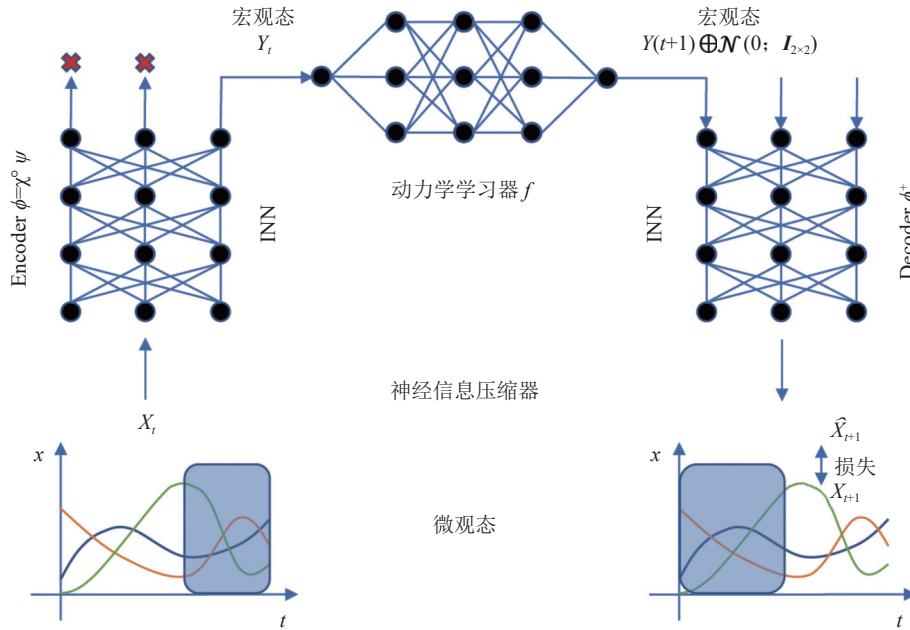


图 2 神经信息压缩方法的模型框架示意

他们借鉴了深度学习任务中广泛使用的 Encoder-Decoder 框架:利用具有强大拟合能力的神经网络构建动力学学习器(f);利用可逆神经网络(INN)构建编码器(Encoder)和解码器(Decoder)。该模型框架可以看成是一个神经信息压缩器,将包含噪声的微观状态压缩成宏观态,丢弃无用信息,从而使得宏观动力学的因果性更强。该方法总的目标函数是希望在微观状态预测误差很小的约束条件下最大化有效信息。该模型主要包括编码器、动力学学习器,以及解码器 3 个部分。

设模型输入为微观状态 $X_t(X_t^1, X_t^2, \dots, X_t^p)$, 输出为下一个时刻的微观状态预测 $\hat{X}_{t+1}(\hat{X}_{t+1}^1, \hat{X}_{t+1}^2, \dots, \hat{X}_{t+1}^p)$, 则总目标为

$$\begin{aligned} & \max_{\Phi_q, \hat{f}_{\Phi_q}, \Phi_q^+, q} I_E(\hat{f}_{\Phi_q}), \\ & \text{s.t.: } \|\Phi_q^+(Y(t+1)) - X_{t+1}\| < \varepsilon. \end{aligned} \quad (9)$$

最终希望得到有效的粗粒化维度 q 、粗粒化策略 Φ_q 和宏观动力学 $\hat{f}_{\Phi_q}$ 。

为了优化目标函数,他们将优化过程分为 2 个阶段:

- 1) $\min_{\Phi_q, \hat{f}_{\Phi_q}, \Phi_q^+, q} \|\Phi_q^+(Y(t+1)) - X_{t+1}\|$;
- 2) 遍历不同的 q , 找到有效信息最大的宏观尺度 $\max_q I_E(\hat{f}_{\Phi_q})$ 。

神经信息压缩器用可逆神经网络实现编码器 Φ_q 和解码器 Φ_q^+ , 其中编码器由双射 Ψ_q 和投影操作 χ_q 这 2 部分构成, 分别对应于信息转换和丢弃 2 个过程。双射 Ψ_q 使用 RealNVP^[47] 可逆神经网络实现, 可逆神经网络能降低模型的复杂度, 调换编码器中可逆神经网络的方向即可构建解码器, 因此, 编码器和解码器可以共享神经网络参数, 同时, 可逆神经网络能保证不丢失信息, 且还可保持该模型框架具有很好的数学性质。给定压缩后维度为 q , 编码器负责将微观状态 X_t 编码到隐空间, 同时希望将输入的有用信息尽可能压缩至 q 维隐空间状态, 剔除 $p-q$ 维噪声信息, 得到宏观态 Y_t ; 在隐空间利用全连接神经网络学习宏观动力学 f , 得到下一时刻的宏观态 Y_{t+1} 。为了保证输入输出维度一致, 需要添加 $p-q$ 维高斯噪声 ($z_{p-q} \in N(0, I_{p-q})$), 然后输入解码器, 解码器负责将宏观预测的结果解码到微观状态空间, 最终得到下一时刻的微观预测 $\hat{X}_{t+1}$ 。

在保证预测误差足够小的情况下, NIS 方法可以

避免 trivial 解的出现. 例如, 当宏观动力学是一个恒等映射时, 通过计算所获得的有效信息十分有用, 但它却不是有意义的动力学, 而 NIS 方法可有效地避免这种情况发生.

除了基于时序数据自动识别因果涌现, 鉴于该框架采用了可逆神经网络, 因此, 有 2 个比较重要的定理^[21].

定理 1 信息挤压器的信息瓶颈, 对于任意的双射 Ψ_α 、投影 χ_q 、宏观动力学 f 以及高斯噪声 $z_{p-q} \in N(0, I_{p-q})$, $I(Y_i; Y_{i+1}) = I(X_i; \hat{X}_{i+1})$ 恒成立, 其中宏观动力学是整个通道的信息瓶颈.

定理 2 对于一个训练较好的模型, $I(X_i; \hat{X}_{i+1}) \approx I(X_i; X_{i+1})$, 模型的互信息接近于数据的互信息.

由定理 1 和 2 可知, 对于一个训练好的模型, 可以得到 $I(Y_i; Y_{i+1}) \approx I(X_i; X_{i+1})$.

NIS 方法的另外一个贡献是其拓展了 Hoel 所提出的、基于粗粒化方法只能计算离线动力学有效信息的局限性. NIS 方法将神经网络看作给定输入 X 条件下的高斯分布 $P(Y|X)$, 定义了新的有效信息计算式

$$I_{\text{EL}}(f) = I(\text{do}(X \in U([-L, L]^n)); Y) \approx \frac{n + n \ln(2\pi) + \sum_{i=1}^n \sigma_i^2}{2} + n \ln(2L) + E(\ln |\det(\partial_X f(X))|), \quad X \in U([-L, L]^n), \quad (10)$$

式中: $U([-L, L]^n)$ 表示 $[-L, L]$ 上的 n 维均匀分布; σ_i 是输出 Y_i 的标准差, 可以通过 Y_i 的均方误差来估计; $\det$ 表示函数 f 的雅可比行列式.

式(10)既能计算离散动力学系统的有效信息, 也适合连续动力系统. 但其会受到输入维度的影响, 当输入维度差异很大时, 有效信息差异也会很大, 这会干扰因果涌现. 为了消除维度影响, Zhang 等^[21] 定义了新的有效信息计算式

$$I_{\text{ELD}}(f) \approx -\frac{1 + \ln(2\pi) + \sum_{i=1}^n \sigma_i^2/n}{2} + \ln(2L) + \frac{1}{n} E(\ln |\det(\partial_X f(X))|), \quad X \in U([-L, L]^n). \quad (11)$$

NIS 方法能克服信息分解方法的不足, 但是也存在一些缺点. 该方法并没有真正最大化 I_E , 而是通过比较不同尺度下的 I_E , 选择其中最大的作为最优尺度; NIS 方法目前只能适用于规模较小的简单系统.

2 相关应用

2.1 应用于简单系统 在 Hoel^[18] 的研究中, 给出了 1 个只包含 4 个节点的布尔网络例子, 每个节点有

0 和 1 这 2 种状态, 节点之间通过边进行链接. 他通过特定的粗粒化使得宏观尺度下的因果性更强. 另外, 他还给出了 3 个例子, 分别通过空间粗粒化、时间粗粒化和时空粗粒化, 均能产生因果涌现的效果. 然而并没有给出一般系统应该如何进行节点分组, 如何定义微观到宏观之间的映射函数. 在后续的工作中, 他尝试解决了这个问题, 同时希望将他的方法拓展到其他类型的系统中.

Zhang 等^[21] 在弹簧振子模型中对比了普通神经网络, 结果发现 NIS 方法不仅学到了真实的隐藏宏观状态, 同时也学到了真实的宏观动力学. 而普通神经网络的预测效果很差, 真实的隐藏宏观状态和动力学 2 个方面都没有学好. 在多步预测方面, NIS 方法能有效地预测弹簧的移动轨迹. 当宏观尺度 $q=2$ 时, 因果涌现效果最强, 对应于真实的宏观动力学中的速度和位置 2 个宏观状态. 他们还在简单马尔可夫链与布尔网络上进行了试验: 对于马尔可夫链, 有 8 个微观状态, 其中 7 个状态可以等概率互相转移, 最后 1 个状态是独立的; 对于布尔网络, 只含有 4 个节点, 每个节点与其中 2 个节点相链接, 遵循相同的微观动力学机制, 因此, 一共含有 16 个微观状态. 2 个试验结果均发现, $q=1$ 时, 因果涌现效果最强, 同时将微观状态输入学习好的编码器中, 得到隐空间的宏观状态与期望的结果保持一致.

2.2 应用于复杂网络 2020 年, Klein 等^[25] 将基于粗粒化的方法应用于复杂网络, 借助随机游走子定义网络中的马尔可夫链, 将随机游走子放在节点上, 等价于对节点做干预 $\text{do}(\cdot)$; 基于随机游走概率定义节点的转移概率矩阵; 建立了有效信息与网络连通性的联系. 网络中的连通性可通过节点出边与入边的权重的不确定性表征 2 项衡量: 1) 节点输出的不确定性可通过节点出权的香农熵定义, 即 $H(W_i^{\text{out}})$, 因此整个网络的不确定性可通过 $\langle H(W_i^{\text{out}}) \rangle$ 得到; 2) 基于网络的出边权重分布计算, $H(\langle W_i^{\text{out}} \rangle)$ 反映了确定性如何在网络中分布. 通过这 2 项就可得到复杂网络中的有效信息定义, $I_E = H(\langle W_i^{\text{out}} \rangle) - \langle H(W_i^{\text{out}}) \rangle$. 为了识别复杂网络中的因果涌现, 需要对网络进行粗粒化, 然后比较宏观网络与微观网络的有效信息, 判断能否发生因果涌现. 他们采用贪婪算法对节点进行分组, 并构建宏观网络. 将微观节点合并成宏观节点时, 对应宏观网络的转移概率矩阵也要进行相应的处理. 通过使用高阶节点显式地对高阶依赖项建模 (HOMs), 保证分组后的宏观网络和原始网络具有相同的随机游走动力学; 通过动力学的一致性检验进一步验证了 HOMs 方

法的有效性.

他们在随机(ER)、偏好依赖(PA)等人工网络,以及 4 类真实网络中做了试验.对于 ER 网络,有效信息的大小只依赖连接概率 P ,并且随着网络规模的增大收敛到 $-\log_2 P$.在某点之后,随机网络结构不会随着其规模的增加而包含更多的信息,这个相变点近似在网络的平均度 $\langle k \rangle = \log_2 N$ 的位置.随着 ER 网络连接概率的增大,巨连通集团相变点位置对应出现.对于 PA 网络: $\alpha < 1.0$ 时,有效信息的大小随着网络规模的增加而增大; $\alpha > 1.0$ 时,结论相反; $\alpha = 1.0$ 对应的无标度网络则是 I_e 增长的临界边界.对于真实网络,他们发现,生物网络因为具有很大的噪声,所以有效信息最低,通过有效的粗粒化能去除这些噪声.相较于其他类型,因果涌现最显著;技术类型网络是更稀疏、非退化的网络,因此,平均效率更高,节点关系更加具体,所有有效信息也最多.

他们还利用贪婪算法构建了宏观尺度的网络,对于大规模网络其效率很低.

Griebenow 等^[26]提出了一种基于谱聚类的方法识别偏好依附网络中的因果涌现.相较于贪婪算法以及梯度下降算法,谱聚类算法的计算时间最少,同时找到宏观网络的因果涌现也更加显著.

2.3 应用于生物系统 鉴于生物网络中充满噪声,其中有些是系统固有噪声,有些是测量或观察引入的,因而学者们很难理解其内部运作机制. Klein 等^[29]通过探索生物网络中的噪声、简并性和确定性以及它们三者之间的关系发现,基因表达网络中的高度确定性关系可理解为一个基因肯定带来另一个基因的表达,生物系统在进化过程中普遍存在简并现象.这 2 个原因使得目前尚不清楚在何种尺度分析生物系统,才能更好地理解它们的功能. Brennan 等^[30]分析了超过 1 800 个物种蛋白质的相互作用网络,发现宏观尺度的网络具有更小的噪声和简并性,组成宏观尺度交互群中的节点更具有弹性.为了适应进化要求,生物网络需要演化成宏观尺度以提高确定性、增强网络的弹性,以及提高信息传输的有效性.

Hoel 等^[48]借助有效信息理论进一步研究了生物系统中的因果涌现.他们将有效信息应用于基因调控网络中,用以识别最有可能提供信息的心脏发育模型,从而控制哺乳动物的心脏发育;通过量化酿酒酵母基因网络的最大连通集团的因果涌现,揭示了富有信息的宏观尺度普遍存在于生物学,以及生命机制的运行尺度也是宏观尺度.他们的方法为生物学家提供了一个可计算的工具,用以识别最具有信息的宏观尺度,并且在此尺度上可建模、预测、控制和理解复

杂的生物系统.

为了验证猕猴运动有关的信息是其皮层活动的一个涌现特征, Rosas 等^[20]做了如下实验:基于猕猴的皮质脑电图(EEG)和动作捕捉(MoCap)数据,其中 EEG 和 MoCap 分别由 64 个通道和 3 个通道的数据构成微观和宏观数据;由于最原始的 MoCap 数据不满足随附特征的条件独立假设,因此,他们利用偏最小二乘和支持向量机算法,推断出与预测猕猴行为有关的编码在 EEG 信号中的那部分神经活动,并推测该信息是潜在神经活动的涌现特征;基于计算所得宏观特征与微观状态,验证了因果涌现的存在.

Swain 等^[49]探索了蚁群的交互历史对任务分配和任务切换的影响,利用有效信息研究了噪声信息如何在蚂蚁之间传播.结果表明:蚁群之间历史交互程度影响任务的分配,并且具体交互中,蚂蚁群体的类型决定了交互中的噪声;即使蚂蚁切换功能群时,蚁群涌现出来的凝聚力也能保证群体的稳定,不同功能的蚁群在维持蚁群凝聚力方面发挥的作用不同.

2.4 应用于人工神经网络 对于深度神经网络(DNN)的研究,学者们一般关注其输入和输出之间的互信息,很少关注其因果结构.检验模型预测能力的一个重要指标就是其泛化能力,它是 DNN 因果结构的函数. Marrow 等^[28]尝试引入了一个基于信息论的指标,即有效信息,用以量化和跟踪训练过程中 DNN 因果结构的变化,其中有效信息用于评估节点和边对每层下游目标的因果的影响程度.有效信息可分解为灵敏性和简并性,通过观察模型训练过程中的有效信息,包括灵敏性和简并性的变化,就可以确定模型的泛化能力,从而更好地理解 and 解释 DNN 的工作原理.

3 结束语

本文讨论了复杂系统中的因果涌现,给出了因果涌现的定量描述,阐述了将因果引入涌现衡量的原因以及引入因果涌现的意义,详细梳理了 2 种因果涌现的定义和 2 种因果涌现的辨识方法及其各自优缺点,同时也给出了各自方法的相关应用. Hoel 提出的粗粒化方法首次定义了因果涌现的计算框架,但是粗粒化策略、宏观尺度以及马尔可夫链需要预先给出;信息分解方法避免了粗粒化操作,但是仅找到了一个因果涌现发生的充分条件,并且计算的时间复杂性高,很难应用到大的复杂系统中;他们所提出的计算指标只是基于互信息计算,没有考虑因果,严格来说只是涌现的识别;神经信息压缩方法基于时间序列数据建模,大大拓宽了因果涌现的应用范围,适合离散和连续动力学系统,并且能自动学习粗粒化策略和宏观动

力学. 故此, 数据驱动因果涌现的识别方法将是未来一个主要的研究方向, 局限于目前的研究对象仅停留于简单的系统中, 如何将现有方法拓展至更大的复杂系统是一个具有挑战性和值得研究的问题. 此外, 该方法针对不同的动力学系统需要训练不同的模型, 因此, 训练一个适用任意一个动力学系统的通用大模型, 是一个很有意义且亟待解决的问题.

因果涌现除了可以帮助学者们更加深入地理解复杂系统在不同尺度下的因果规律, 其与人工智能的一些子领域, 如因果表示学习^[50]、因果模型的抽象^[51]和强化学习中的世界模型^[52]之间也有着紧密的联系, 因果涌现的研究也可为这些领域提供新的思路. 由于原因和结果之间存在多重对应关系, 识别因果涌现中的信息分解方法, 可为学者们理解复杂系统中因果关系的复杂性提供一个数学框架. 此外, 学者们常常利用网络和动力学研究复杂系统, 从因果角度而言, 网络可看作因果结构, 动力学可看作因果机制. 除此以外, 因果涌现可广泛应用于神经以及社会系统等其他领域. 在神经科学领域, 因果涌现可用来解释大脑功能和神经元之间的相互作用. 通过分析其因果关系, 可更好地了解大脑的结构和功能, 并为神经疾病的诊断及治疗提供新的思路.

因果涌现作为一种新兴的科学方法, 具有广泛的应用前景. 通过研究不同层次之间的因果关系, 可更加深入地理解自然界和社会现象的运行机制, 并为各领域的发展提供有力的支持.

4 参考文献

- [1] 智韬, 司光亚, 贺筱媛. 电力关键基础设施网络仿真模型研究[J]. *系统仿真学报*, 2010, 22(11): 2732
- [2] 叶青. 基于复杂网络理论的轨道交通网络脆弱性分析[J]. *中国安全科学学报*, 2012, 22(2): 122
- [3] KWAPIEŃ J, DROŻDŻ S. Physical approach to complex systems[J]. *Physics Reports*, 2012, 515(3/4): 115
- [4] 王飞跃, 史帝夫·兰森. 从人工生命到人工社会: 复杂社会系统研究的现状和展望[J]. *复杂系统与复杂性科学*, 2004, 1(1): 33
- [5] 王飞跃. 人工社会、计算实验、平行系统: 关于复杂社会经济系统计算研究的讨论[J]. *复杂系统与复杂性科学*, 2004, 1(4): 25
- [6] 彭雁虹, 褚启勤, 李怀祖. 生命系统理论及其应用综述[J]. *系统工程理论与实践*, 1997, 17(3): 35
- [7] ANDERSON P W. More is different[J]. *Science*, 1972, 177(4047): 393
- [8] CAMAZINE S, DENEUBOURG J L, FRANKS N R, et al. Self-organization in biological systems[M]. New Jersey: Princeton University Press, 2020
- [9] 苗东升. 论系统思维(六): 重在把握系统的整体涌现性[J]. *系统科学学报*, 2006, 14(1): 1
- [10] 方锦清, 汪小帆, 刘曾荣. 略论复杂性和非线性复杂网络系统的研究[J]. *科技导报*, 2004, 22(2): 9
- [11] MEEHL P E, SELLARS W. The concept of emergence[J]. *Minnesota Studies in the Philosophy of Science*, 1956, 1: 239
- [12] BEDAU M A. Weak emergence[J]. *Philosophical Perspectives*, 1997, 11: 375
- [13] PEPPER S C. Emergence[J]. *The Journal of Philosophy*, 1926, 23(9): 241
- [14] FROMM J. Types and forms of emergence[EB/OL]. (2005-06-13)[2023-02-15]. <https://arxiv.org/abs/nlin/0506028.pdf>
- [15] PEARL J. Graphs, causality, and structural equation models[J]. *Sociological Methods & Research*, 1998, 27(2): 226
- [16] GRANGER C W J. Investigating causal relations by econometric models and cross-spectral methods[J]. *Econometrica*, 1969, 37(3): 424
- [17] NEUBERG L G. Causality: models, reasoning, and inference, by Judea Pearl, Cambridge University Press, 2000[J]. *Econometric Theory*, 2003, 19: 675
- [18] HOEL E P, ALBANTAKIS L, TONONI G. Quantifying causal emergence shows that macro can beat micro[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(49): 19790
- [19] HOEL E. When the map is better than the territory[J]. *Entropy*, 2017, 19(5): 188
- [20] ROSAS F E, MEDIANO P A M, JENSEN H J, et al. Reconciling emergences: an information-theoretic approach to identify causal emergence in multivariate data[J]. *PLoS Computational Biology*, 2020, 16(12): e1008289
- [21] ZHANG J A, LIU K W. Neural information squeezer for causal emergence[J]. *Entropy*, 2022, 25(1): 26
- [22] MARSH H W, YEUNG A S. Top-down, bottom-up, and horizontal models: the direction of causality in multidimensional, hierarchical self-concept models[J]. *Journal of Personality and Social Psychology*, 1998, 75(2): 509
- [23] CHVYKOV P, HOEL E. Causal geometry[J]. *Entropy*, 2020, 23(1): 24
- [24] VARLEY T, HOEL E. Emergence as the conversion of information: a unifying theory[EB/OL]. (2021-04-27)[2023-02-16]. <https://arxiv.org/abs/2104.13368.pdf>
- [25] KLEIN B, HOEL E. The emergence of informative higher

- scales in complex networks[J]. *Complexity*, 2020, 2020: 1
- [26] GRIEBENOW R, KLEIN B, HOEL E. Finding the right scale of a network: efficient identification of causal emergence through spectral clustering[EB/OL]. (2022-02-16)[2023-02-17]. <https://arxiv.org/abs/1908.07565v1>
- [27] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用[M]. 北京: 清华大学出版社, 2006
- [28] MARROW S, MICHAUD E J, HOEL E. Examining the causal structures of deep neural networks using information theory[J]. *Entropy*, 2020, 22(12): 1429
- [29] KLEIN B, HOEL E, SWAIN A, et al. Evolution and emergence: higher order information structure in protein interactomes across the tree of life[J]. *Integrative Biology*, 2021, 13(12): 283
- [30] BRENNAN K, ANSHUMAN S, TRAVIS B, et al. Exploring noise, degeneracy and determinism in biological networks with the einet package[J]. *Methods in Ecology and Evolution*, 2022, 13(4): 799
- [31] MEDIANO P A M, ROSAS F E, LUPPI A I, et al. Greater than the parts: a review of the information decomposition approach to causal emergence [EB/OL]. (2022-05-23) [2023-02-17]. <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2021.0246>
- [32] TONONI G, SPORNS O. Measuring information integration[J]. *BMC Neuroscience*, 2003, 4: 31
- [33] WILLIAMS P L, BEER R D. Nonnegative decomposition of multivariate information[EB/OL]. (2010-04-14)[2023-02-18]. <https://arxiv.org/abs/1004.2515.pdf>
- [34] ROSAS F E, MEDIANO P A M, RASSOULI B, et al. An operational information decomposition via synergistic disclosure[J]. *Journal of Physics A: Mathematical and Theoretical*, 2020, 53(48): 485001
- [35] REICHSTEIN M, CAMPS-VALLS G, STEVENS B, et al. Deep learning and process understanding for data-driven Earth system science[J]. *Nature*, 2019, 566(7743): 195
- [36] SENIOR A W, EVANS R, JUMPER J, et al. Improved protein structure prediction using potentials from deep learning[J]. *Nature*, 2020, 577(7792): 706
- [37] KEISER M J, SETOLA V, IRWIN J J, et al. Predicting new molecular targets for known drugs[J]. *Nature*, 2009, 462(7270): 175
- [38] 程乐峰, 余涛, 张孝顺, 等. 机器学习在能源与电力系统领域的应用和展望[J]. *电力系统自动化*, 2019, 43(1): 15
- [39] 张伟伟, 朱林阳, 刘溢浪, 等. 机器学习在湍流模型构建中的应用进展[J]. *空气动力学学报*, 2019, 37(3): 444
- [40] GLYMOUR C, ZHANG K, SPIRITES P. Review of causal discovery methods based on graphical models[EB/OL]. (2019-06-04)[2023-02-18]. <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2021.0246>
- [41] CASADIEGO J, NITZAN M, HALLERBERG S, et al. Model-free inference of direct network interactions from nonlinear collective dynamics[J]. *Nature Communications*, 2017, 8: 2192
- [42] ZHANG Y, GUO Y, ZHANG Z, et al. Universal framework for reconstructing complex networks and node dynamics from discrete or continuous dynamics data[EB/OL]. (2022-09-16)[2023-02-18]. <https://journals.aps.org/pre/abstract/10.1103/PhysRevE.106.034315>
- [43] GLYMOUR C, ZHANG K, SPIRITES P. Review of causal discovery methods based on graphical models[J]. *Frontiers in Genetics*, 2019, 10: 524
- [44] 蔡瑞初, 陈薇, 张坤, 等. 基于非时序观察数据的因果关系发现综述[J]. *计算机学报*, 2017, 40(6): 1470
- [45] HU H Y, LI S H, WANG L, et al. Machine learning holographic mapping by neural network renormalization group[J]. *Physical Review Research*, 2020, 2(2): 023369
- [46] LI S H, WANG L. Neural network renormalization group[J]. *Physical Review Letters*, 2018, 121(26): 260601
- [47] DINH L, SOHL-DICKSTEIN J, BENGIO S. Density estimation using Real NVP[EB/OL]. (2017-02-27)[2023-02-18]. <https://arxiv.org/abs/1605.08803>
- [48] HOEL E, LEVIN M. Emergence of informative higher scales in biological systems: a computational toolkit for optimal prediction and control[J]. *Communicative & Integrative Biology*, 2020, 13(1): 108
- [49] SWAIN A, WILLIAMS S D, DI FELICE L J, et al. Interactions and information: exploring task allocation in ant colonies using network analysis[J]. *Animal Behaviour*, 2022, 189: 69
- [50] SCHÖLKOPF B, LOCATELLO F, BAUER S, et al. Toward causal representation learning[J]. *Proceedings of the IEEE*, 2021, 109(5): 612
- [51] IWASAKI Y, SIMON H A. Causality and model abstraction[J]. *Artificial Intelligence*, 1994, 67(1): 143
- [52] MATSUO Y, LECUN Y, SAHANI M, et al. Deep learning, reinforcement learning, and world models[J]. *Neural Networks*, 2022, 152: 267

Causal emergence in complex systems

WANG Zhipeng¹⁾ ZHANG Jiang^{1, 2)}

(1) School of Systems Science, Beijing Normal University, 100875, Beijing, China;

2) Swarma Research, 100085, Beijing, China)

Abstract Focusing on complex systems; a quantitative description is provided for emergence, causality, and causal emergence. The relationship between causality and emergence is elucidated. Two quantitative methods, coarse-graining and information decomposition, are reviewed for characterizing causal emergence. Two causal emergence identification methods based on information decomposition and neural information compression are presented. Basic principles, advantages, disadvantages, and relevant applications of each method are detailed. The framework of causal emergence is defined based on coarse-graining method, which can be applied to discrete dynamical systems. The use of information decomposition method in solving problems relies on redundant information. The proposed indicators based on information decomposition can identify causal emergence in data and find a sufficient condition. The data-driven neural information compression method is scalable and applicable to continuous dynamical systems. It can automatically extract coarse-grained functions at different levels and construct hierarchical dynamics, and identify causal emergence in different types of dynamical systems. Existing methods are improved and applied to identify more complex systems, automate learning groups, and address challenges in general dynamics of large-scale models and other related problems.

Keywords complex system; causal emergence; effective information; coarse-graining; information decomposition; neural information compression

【责任编辑: 陆有忠】