\bigcirc

AGI 大统一理论

0 综述

- ·大统一理论是在强化学习的框架下进行的,这是以 Richard Sutton 为代表人物提出的理论框架。
- •在强化学习里最辣手的一个问题,就是如何储存和计算所有状态之上的概率分布。对AGI来说,状态 = 思维空间。我们需要的是所有可能的思维之上的概率分布,而这是AGI的一个硬性需求,无法避免。由于思维空间是高维的向量空间,它上面的概率分布是一个庞大的mathematical object,很难在计算机上表示。如果用神经网络表示,则问题是如何对这个概率分布进行采样(sampling),在神经网络里,这是很困难的。
- Hopfield 网络的权重 定义了一个 能量地势 (energy landscape), 它可以看成是一个 implicit 的 概率分布。透过 Hopfield 网络的 learning,可以改变这个概率分布。但这需要修改 Hopfield 网络的算法,将能量诠释成概率,而这正是 Boltzmann machine,也称作 EBM (Energy-Based Models).
- 根据 "Hopfield Network is All You Need" 论文¹, 现代 Hopfield 网络的 state update rule 跟 Transformer 重合²。换句话说,每执行一次Transformer,就会趋向 Hopfield 的能量最低点。
- Transformer 的 softmax 可以看成是 **大脑**中某种 "winner-takes-all" 机制。从这个角度,可以类比大脑思考的机制,互相参考以获取更多灵感。
- 我最新的论文 提出,Transformer 具有 逻辑结构,可以在逻辑基础上建立 AGI.

¹感谢 Eric Zeng 给我推荐这篇论文。

²注意这是 state update rule 而不是 learning update rule. 前者 更改 Hopfield 网络的 激活 状态;后者 更改 Hopfield 网络的权重 / 记忆。

2 Hopfield 网络 1

1.1 经典 Hopfield 网络 我们的符号跟随 Hopfield Network is All You Need.

总能量 (Hamiltonian):

 \mathbf{x}^i = 需要记忆的 patterns (有 N 个), x_s^i 是它的 s-th bit.

 $\mathbf{X} = (\mathbf{x}^1, ..., \mathbf{x}^N)$ 是所有 patterns 的矩阵。

连接权重 between s-th and t-th neurons:

 ξ = 网络的 **状态** $, \xi_s$ = s-th 神经元 的 激活状态。

(1)

(2)

(3)

(6)

(9)

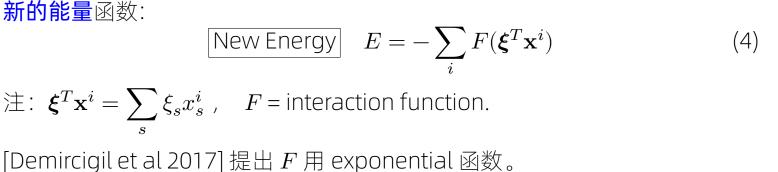
(10)

Hamiltonian 能量函数,令干扰减弱,可以储存数量更多的 patterns: pattern pattern

> energy landscape

在经典 Hopfield 网络里,当 A和 B两个 patterns 太靠近的时候,它们会

互相干扰,导致可以储存的 patterns 数量不大。现代 Hopfield 网络 改变



State update rule:

 $\boldsymbol{\xi}^{\mathsf{NEW}} = \mathbf{X} \; \mathsf{softmax}(\beta \mathbf{X}^T \boldsymbol{\xi})$ (5)

 $\mathbf{Z} = \operatorname{softmax} \left(\frac{1}{\sqrt{d_k}} \mathbf{Q} \, \mathbf{K}^T \right) \mathbf{V}$

1.3 Hopfield-Transformer 对应

传统 Transformer's state update rule:

Modern Hopfield network's state update rule:
$$\mathbf{Z} = \operatorname{softmax} \left(\beta \hat{\mathbf{R}} \hat{\mathbf{Y}}^T \right) \hat{\mathbf{Y}}^T \tag{7}$$

这里 $\hat{\mathbf{R}},\hat{\mathbf{Y}},\hat{\mathbf{X}}$ 是 $\mathbf{R},\mathbf{Y},\mathbf{X}$ 分别乘上了适当的 \mathbf{W}' s 矩阵,但为了数式的简洁,

使用了代号。
$$\text{Query patterns} = \mathbf{R} = (\mathbf{r}_1,...,\mathbf{r}_M) \text{ 是网络的状态。之所以有 M 个状态,是$$

因为他们将 Transformer 的 M 个输入 摊开来,才构成一个大的 Hopfield 网。不这样做根本无法将 Hopfield 网和 Transformer 等同起来。

Y 就是 Hopfield 记忆里的 patterns,它们担任 Transformer 里 keys 的角色。

在 Self-Attention 里, $\mathbf{R} = \mathbf{Y}$. 1.4 Boltzmann 机

Let $O = (O_1, ..., O_n)$ be the state vector.

 $W = \{W_{s,t}\}$ are connection weights. **State update rule**: *i*-th unit is set to 1 with probability

注意:以下是 Boltzmann machine 跟 经典 Hopfield network 的对应。

$$\frac{1}{1 + e^{-S_i/T}}\tag{8}$$

 $P(O) = P(O|W) = \frac{e^{-\mathcal{E}(O)/T}}{Z}$ Boltzmann distribution

悉的 Q-Learning 的 temporal difference update:

如果用以上的 update rule,则 Hopfield 能量变成概率分布:

where partition function
$$Z = \sum_{U} e^{-\mathcal{E}(U)/T}$$

强化学习的 Bellman update 是根据某个 output的 reward, 例如我比较熟

TO-DO: 求出 现代版的 Hopfield network 的 Boltzmann state update rule.

s = state,

a = action,

1.5 与强化学习 结合

where T is a temperature.

 η = learning rate, γ = discount factor,

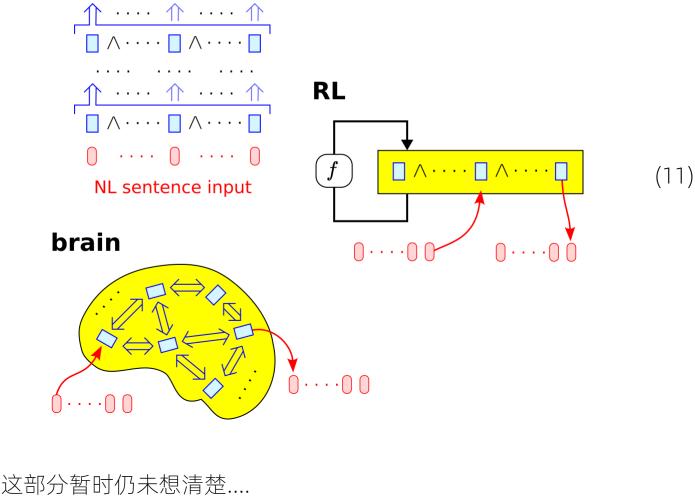
 $Q(s, a) += \eta [R + \gamma \Delta Q]$

而 这个
$$Q$$
 值可以看成是某种 **能量**。

LLM

RL

下图中, 我们比较 LLM, RL 和 大脑。它们的状态 有什么对应关系?





TO-DO: 求出 Hopfield-Boltzmann machine 的learning update rule. 但它似 乎是根据 **记忆** 而 update 的??