

# AGI 的一些基本概念

YKY 甄景贤

Independent researcher, Hong Kong  
*generic.intelligence@gmail.com*

May 12, 2019

# Talk summary

- 1 什么是 归纳偏好？  
「没有免费午餐」
- 2 神经网络 的 力量  
来自什么？
- 3 Turing 机 与 逻辑 的 宇宙性
- 4 经典逻辑 AI 系统 的 结构

# Section 1

什么是 归纳偏好？  
「没有免费午餐」

# 机器学习 的 目的

- 机器学习的目的，是在某些「学习机器」的空间中，搜寻符合要求的某些机器
- 例如在所有给定大小的神经网络中，搜寻符合目标函数的那些神经网络的 weights

# AI Winter

- 一般来说，AI 的 樽颈问题 是**搜寻空间太大**，导致 学习 太慢
- 历史上「AI 寒冬」出现的原因，是因为**基於逻辑**的学习方法，导致 搜寻空间的 组合数量爆炸，而没有好的 heuristic (算法窍门)

# Inductive bias (归纳偏好)

- 每种学习方法都有它的 归纳偏好
  - 换言之，在 搜寻空间 里预先 划分 某些部分 是不会搜索的
  - 所以 偏好 令学习更快
  - 但如果 偏好 太强，连 答案 所在的空间也删除了
- “Throw the baby out with the water”

# 「没有免费午餐」定理

- 在搜寻空间 没有 *a priori* (先验的) 分布下, 每种 归纳偏好, **必然 对某些问题 提升效能, 但对余下的问题 降低效能**, 这就是 “no free lunch” 的意思
- 例如 人类视觉 有 3 维 Euclidean 空间的几何不变性, 所以 人脑的视觉神经 很可能有 对这种不变性的偏好
- 又或者 人类思维 有「逻辑」结构, 这种结构可以令 机器 更快学习 人类智慧

# Kolmogorov complexity

- 最后还要一提 Kolmogorov 复杂性
- 它**不可计算** (incomputable) 但**可以近似** (approximable)
- 逻辑命题之间的 **语义距离** (semantic distance metric) 和它有关，一个 逻辑推导 步骤 等如 语义距离的一单位
- 寻找一套 逻辑 法则，**解释** 这世界，但法则的数目不可太多，也不能推出错误结论——这些要求等於 **间接地** 近似 Kolmogorov 复杂性



## Section 2

神经网络 的 力量  
来自什么？

# 神经网络的结构

- 一粒神经元 就是 一个 dot product 接著一个 非线性函数:

$$\sigma \langle x, w \rangle \quad (1)$$

- 这非线性函数 可以有很多种, 例如:

$$\sigma(\xi) = \frac{1}{1 + e^{-\xi}} \quad (2)$$

# 神经网络的结构

- 一层神经元 是 一个 矩阵 乘法:

$$\mathcal{O}(W \cdot x) \quad (3)$$

- 一个神经网络 是很多层的函数 composition ( $f \circ f$ ):

$$[\mathcal{O}W]^L x \quad (4)$$

# 神经网络 的 特性

- 神经网络 是一个有很多**参数** 的 函数
- 它是 **万能的 函数 近似器** [Cybenko 1989]
- 定理 的 证明 可追溯到 Weierstrauss 近似定理 (1885), 即: 任意连续函数 可以用 多项式 近似
- 但这定理的证明**和 深度 无关**

# 神经网络 的威力来自「深度」

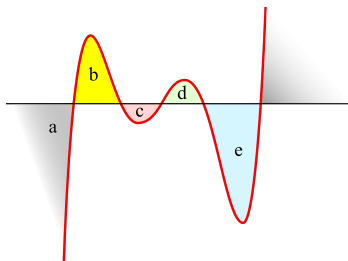
- 假设  $\odot Wx$  是 3 次多项式
- 每增加一层 神经网络，等如

(多项式  $\circ$  多项式) (5)

- 故，总体的 多项式 次数  $= 3^L$
- 换句话说 整体次数 呈**指数式增长**

# 神经网络 的威力来自「深度」

- 代数基本定理：多项式 次数 = 曲线 跨过  $x = 0$  多少次



(6)

- 高维：曲面对分类空间分割成多少块

# 神经网络 的威力来自「深度」

- 这和 VC-dimension 道理一样 [Vapnik–Chervonenkis 1971]
- VC-dimension = 函数 能将 空间 分割成多少块
- 多层 神经网络 的 VC-dim 是  $O(N \log N)$  其中  $N$  是 网络参数 的总个数，但证明用的是不连续的 阀函数
- 我估计 VC-dim 会是 指数增长的，但未有证明

# 神经网络 的威力来自「深度」

- VC-dim **指数增长** 的意义，表示 神经网络 能 代表 一些 非常复杂的 函数家族
- 而 神经网络 的参数个数 相对地 很少，可以在 电脑上实现



# 卷积神经网络的启示

- Yann LeCun 在 1989 发明了 ConvNet, 触发了 机器视觉 的革命, 最近得了 Turing 奖



# 卷积 神经网络 的 启示

- CNN 将普通 NN 的 点积 用 卷积 代替:

$$\boxed{\text{点积}} \quad \mathcal{O}(\langle x, w \rangle) \rightsquigarrow \mathcal{O}(f * g) \quad \boxed{\text{卷积}} \quad (7)$$

- 而 卷积 具有**平移 不变性**, 有利於 视觉:

$$T_x(f) * g = T_x(f * g) \quad (8)$$

- 这是一种 归纳偏好, 令 学习 更快

# 卷积神经网络的启示

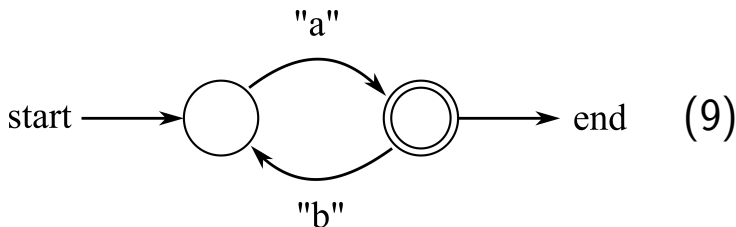
- 其实视觉需要的是仿射 (affine) 不变性，它包括平移、旋转、放大缩小等
- 但似乎单单是平移不变性所带来的学习加速，已足以令 CNN 在 2012 年超越了人类水平
- 可见，归纳偏好 在深度学习里 仍然是很有用的

## Section 3

# Turing 机 与 逻辑 的 宇宙性

# 有限自动机

- 有限自动机 通常用一些 tuple 定义 (从略), 例如:



- 这个 自动机**接受** “a”, “aba”, “ababa...” 等 字符串

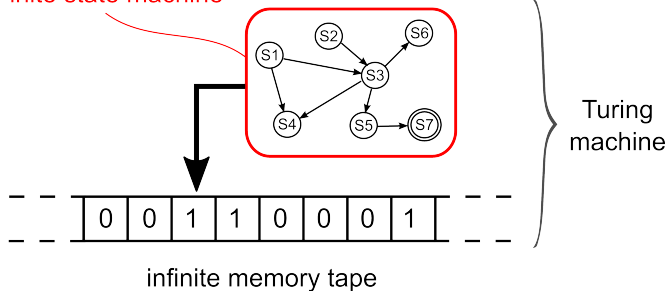
# 有限自动机

- 有限自动机 可以接受  $a^m b^n$  这种字串,  $m$  和  $n$  不同
- 但它不能接受  $a^k b^k$  这种字串, 因为它里面没有办法「记住」 $k$  是多少次
- 有限自动机 能辨认的 语言, 称作 regular languages
- Noam Chomsky 在 1950s 定义, 他是计算机科学家 + 语言学家, 现在主要谈政治, 从左派角度批评美国资本主义

# Turing 机

Turing 机 = 有限自动机 + 无限长的**记忆磁带** (每个 state 可以 读 / 写 一个字符)

Finite state machine



(10)

# Turing 机

- 亦即是说：有限自动机 + 无限读写带  
= 可以计算任何函数，此即  
Church-Turing 假设
- Turing 机 等价於  $\lambda$ -calculus、  
combinatory logic、cellular automata、  
game of life、recurrent 神经网络、等

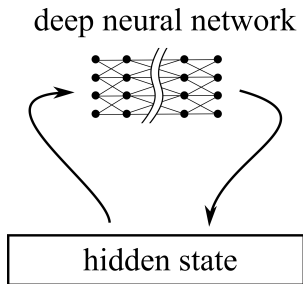


# Alan Turing (1912-1954)

- Turing 是一个 远超越於时代 的人
- 他考虑过 神经网络 作为学习机器
- 也考虑过 进化算法 (evolutionary algorithms)
- 而当时 1940s 还未有电脑 — 电脑是他发明的!
- 他求出所有可计算函数的形式, 从而将 AI 的问题 限制 在一框框内

# 回路 神经网络

其实一个 RNN 可以看成类似 (10) 的结构：



(11)

重点是 hidden state 可以储存计算的**中途结果**，令 RNN 也变成 Turing 机

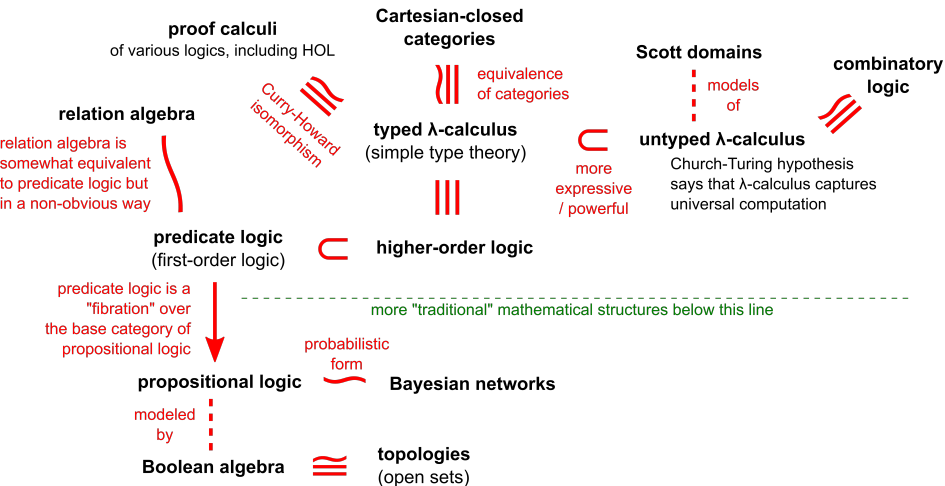
## Section 4

# 经典逻辑 AI 系统的结构

# John McCarthy (1927-2011)

- 「AI 之父」
- 1956 年 在 Dartmouth 第一次举行「人工智能」会议
- 开创了 使用**数理逻辑** 作为 AI 的 **知识表述** (knowledge representation)
- 晚年研究 **改写系统** (term rewriting systems), 是一种更 **广义** 的逻辑

# 逻辑的种类繁多

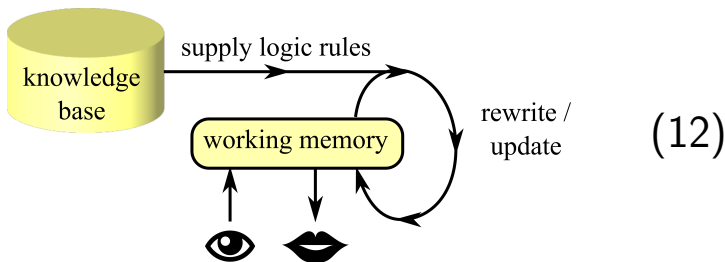


# 命题逻辑 vs 谓词逻辑

- 最重要是 搞清 命题逻辑 和 一阶 谓词逻辑 的区别
- 命题逻辑:  $P_1 = \text{「昨天下雨」}$   
 $P_2 = \text{「今天下雨」}$   
**命题 没有 内部结构**
- 谓词逻辑:  $P_3 = \text{下雨 (北京, 前天)}$
- 谓词 有 代入 (substitution) 的复杂性

# 经典 逻辑 AI 架构

这个 架构 很重要，就像 蒸汽机 时代 的  
Carnot cycle（卡诺循环）：



# 何谓 logic rule ?

- 举例：爱一个人 但他不爱你 则失恋

$$\heartsuit(x, y) \wedge \neg \heartsuit(y, x) \Rightarrow \odot(x) \quad (13)$$

- 这是一条 rule, 变量  $x, y$  需要**代入** 适当的个体, 例如  $\{x \setminus \text{John}, y \setminus \text{Mary}\}$
- 寻找 代入 的算法叫 **matching** 或 **unify**



# SOAR cognitive architecture

- SOAR 是一个著名的 **认知架构**
- 基本上 它根据 working memory 寻找 可以发动的 rules，类似图 (12)
- 它用 **Rete** 算法 快速地搜寻 可用的 rules，这是经典 AI 里的重要算法（*rete* 在拉丁文的意思是「网状」）
- 鉴於中国 AI 的后起之秀，视野不够广阔，故补充一些基础知识

# 我的理论

- 我的理论里，rules 和 matching 机制，都纳入到神经网络里
- 神经网络这件武器，优点是 可以逼近很复杂的 mappings，它是现时最强的机器学习方法
- 我将逻辑结构松弛 (relax)，务求做到足够的归纳偏好即可
- 这理论中最关键的元件，是 symmetric 神经网络

# 逻辑 交换律

- $\wedge$  的交换律 可能是 逻辑 中最重要规律：

$$\text{肚饿} \wedge \text{没钱} \Leftrightarrow \text{没钱} \wedge \text{肚饿} \quad (14)$$

- 也可以理解为：从 前提 推导出 结论，前提中 命题的**次序** 应该没有关系，甚至可以夹杂无关的命题
- 交换律 抽象了 逻辑命题 的结构，类似 抽象代数 中，交换群（又叫 阿贝尔群，纪念 Abel）的重要性

# Symmetric 神经网络

- 卷积 神经网络 具有 平移不变性
- 类似地, **对称**神经网络 具有 交换不变性 (permutation invariance)
- 它可以透过 weight-sharing 来实现, 类似 卷积层 的权重共享
- SymNet : 逻辑  $\approx$  ConvNet : 视觉

# 中国会不会有自主研发的 AGI ?

- 日本在 1980 年代研发第 5 代电脑的失败，可以作为借鉴
- 地球的资源有限，科技发展往往是国与国之间竞争的结果
- 在美国有歧视中国的人，中国内部也有拖后腿、倒向外国的人，但在美国也有帮过我的朋友
- 所以我比较支持建立全球化的 AGI 项目

多谢收看 😊