AGI 的一些基本概念

YKY 甄景贤

Independent researcher, Hong Kong

generic.intelligence@gmail.com

May 4, 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
   基本结构

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 就是 **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，而没有很好的 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”

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:

$$\sigma\langle \boldsymbol{x}, \boldsymbol{w} \rangle \tag{1}$$

- 这非线性函数 可以有很多种，例如:

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

神经网络的结构

- 一层神经元 是 一个 矩阵 乘法:

$$\sigma(W \cdot \boldsymbol{x}) \tag{3}$$

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

$$[\sigma W]^L \, \boldsymbol{x} \tag{4}$$

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 Weirstrauss 定理，即: 任意连续函数 可以用 多项式 近似

神经网络 的威力来自「深度」

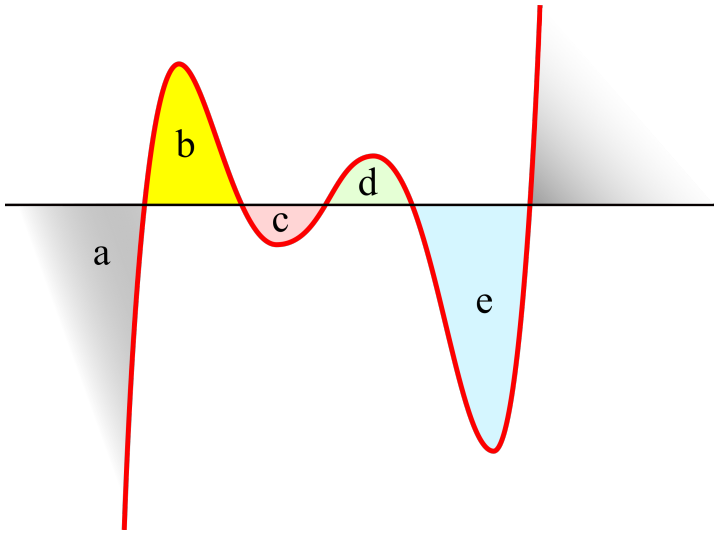
- 例如，假设  是 3 次多项式
- 每增加一层 神经网络，等如

$$(\text{多项式} \circ \text{多项式}) \tag{5}$$

- 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 一个 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，彻底改革了 机器视觉领域，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:

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

- 而 卷积 具有 **平移 不变性**，有利於 视觉:

$$T_x(f) * g = T_x(f * g) \tag{8}$$

- 这是一种 **归纳偏好**，令 学习 更快

卷积 神经网络 的 启示

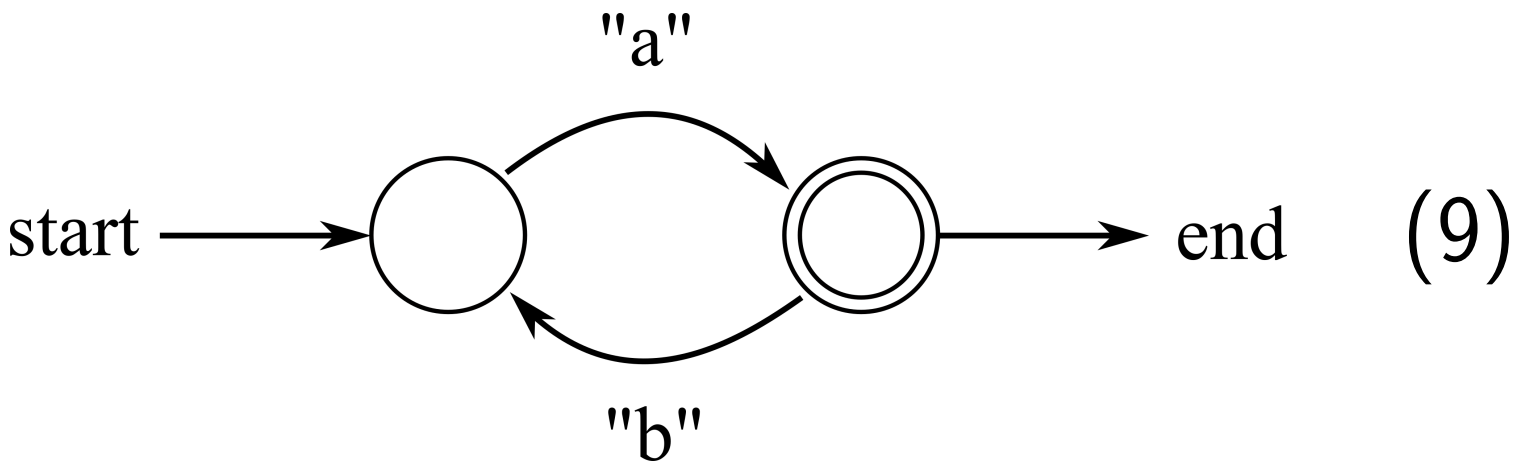
- 其实 **视觉** 需要的是 **仿射** (affine) 不变性，它包括 平移、旋转、放大缩小 等
- 但 似乎 单单是 平移不变性 所带来的 **学习加速**，已足以令 CNN 在 2012 年超越了人类水平
- 可见，**归纳偏好** 在 深度学习 里 仍然是很有用的

Section 3

Turing 机 与 逻辑 的 宇宙性

有限自动机

- 通常用一个 tuple 定义（从略），例如：



(9)

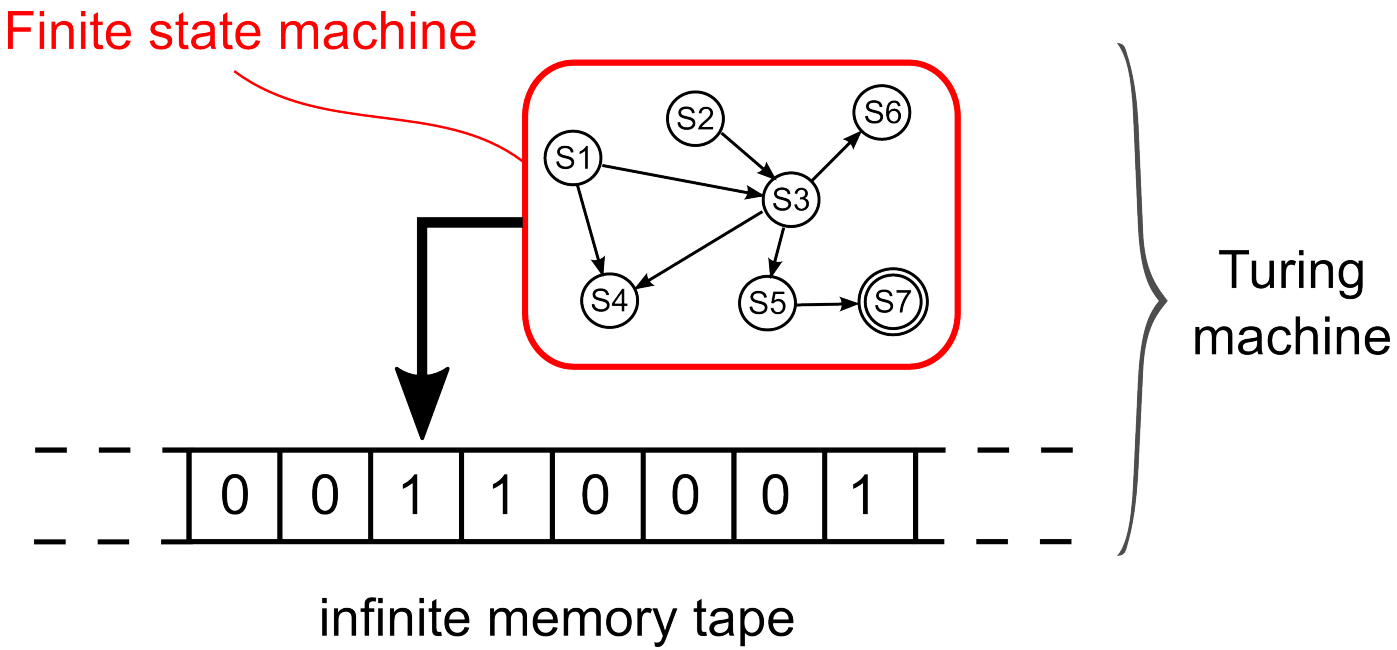
- 这个 FSA **接受**”ab”, ”abab”, ”abababab...” **等**

有限自动机

- 有限自动机 可以接受 $\mathrm{a}^m\mathrm{b}^n$ 这种字串，m 和 n 不同
- 但它不能接受 $\mathrm{a}^k\mathrm{b}^k$ 这种字串，因为它里面没有办法「记住」k 是多少次
- FSA 能够接受的 **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）



(10)

Turing 机

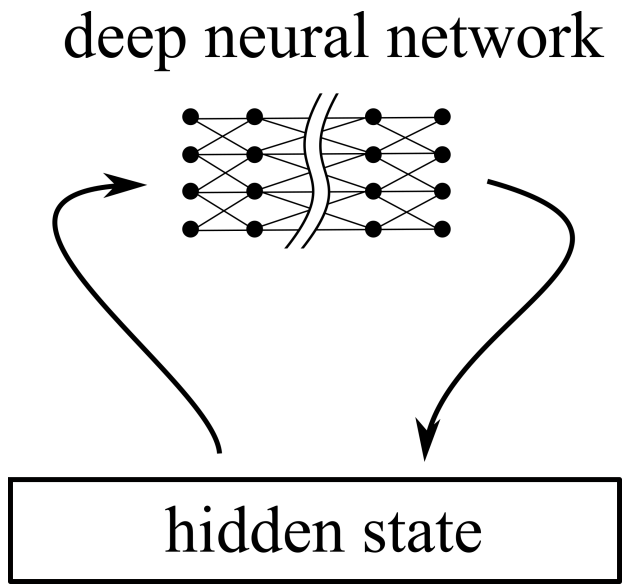
- 亦即是说：有限自动机 + 无限读写带 = 可以计算任何函数 (Church-Turing 假设)
- Turing 机 等价於 λ-calculus、combinatory logic、cellular automata、game of life、recurrent 神经网络、等

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)

所以 RNN 也是一种 Turing 机

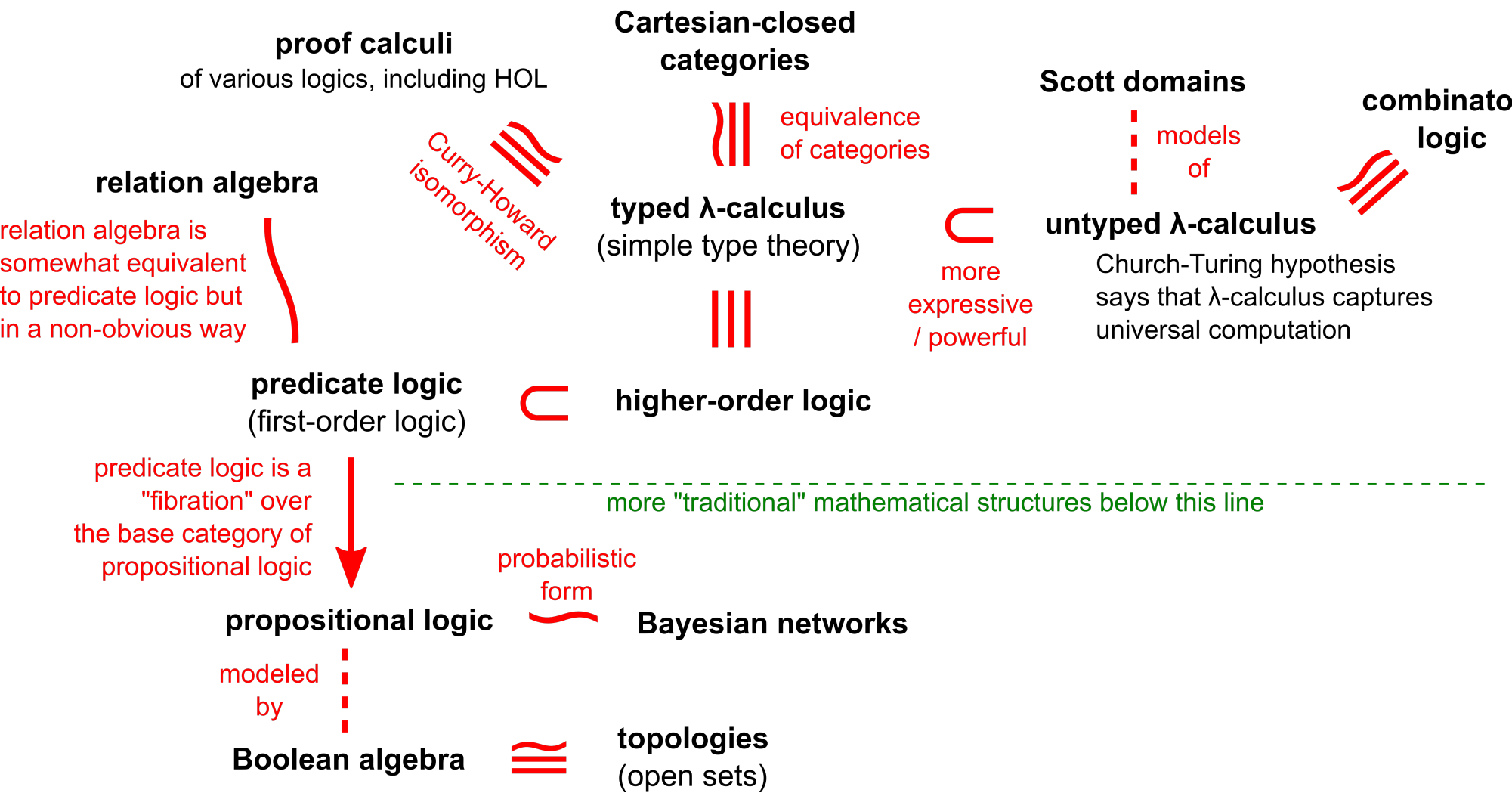
Section 4

经典逻辑 AI 系统 的基本结构

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** (rewriting systems), 是一种更一般的逻辑

逻辑 的 种类繁多

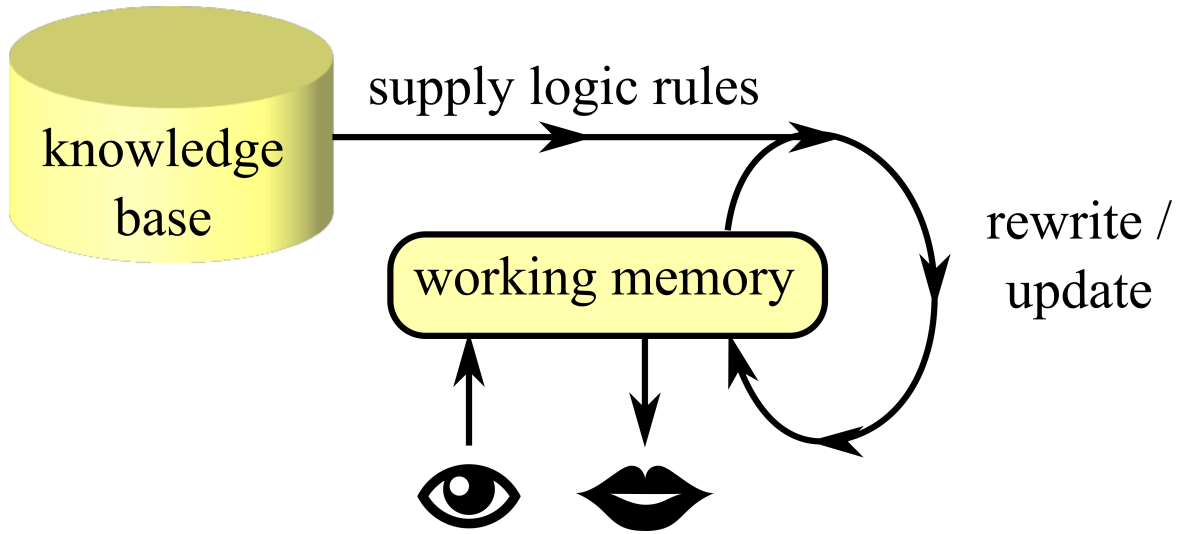

proof calculi
of various logics, including HOL
Cartesian-closed categories
equivalence of categories
Curry-Howard isomorphism
relation algebra
relation algebra is somewhat equivalent to predicate logic but in a non-obvious way
typed λ-calculus
(simple type theory)
Scott domains
models of
combinato logic
more expressive / powerful
untyped λ-calculus
Church-Turing hypothesis says that λ-calculus captures universal computation
predicate logic
(first-order logic)
higher-order logic
predicate logic is a "fibration" over the base category of propositional logic
more "traditional" mathematical structures below this line
probabilistic form
propositional logic
Bayesian networks
modeled by
Boolean algebra
topologies
(open sets)

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: $P_1 =$ 「昨天下雨」
  $P_2 =$「今天下雨」
  命题 没有 **内部结构**
- 谓词逻辑: $P_3 =$ 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）：



(12)

何谓 logic rule ?

- 举例：爱一个人 但他不爱你 则失恋

$$\heartsuit(x,y) \wedge \neg\heartsuit(y,x) \Rightarrow \text{☹}(x) \tag{13}$$

- 这是一条 rule，变量 x, y 需要 **代入** 适当的个体，例如 $\{x/\text{John}, y/\text{Mary}\}$
- 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，如图 (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」）
- 很多中国 AI 后起之秀，只熟悉当下最流行的算法，稍为扯远一点就不懂，视野不够广阔

我的理论

- 在我的理论里，rules 和 matching 机制，都 **纳入** 到 神经网路 里
- 神经网络 这件武器，它的优点是 可以近似 很复杂的 mappings，它是现时最强的 机器学习 方法
- 我将 逻辑结构 放松 (relax)，务求做到足够的 归纳偏好 即可
- 这理论中最重要的元件，是 symmetric 神经网络

We're looking for developers to implement a prototype.

Thank you