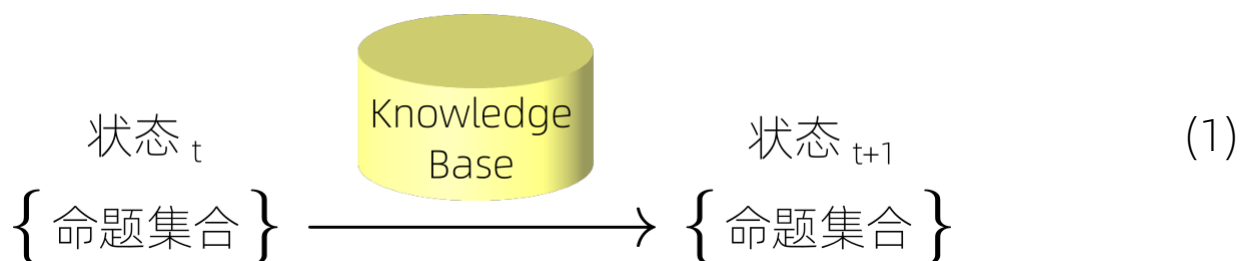


①

# 逻辑与深度学习的关系

这是经典逻辑 AI 的最基本运作模式：



它其实包含了两个算法：

- **matching** (unification):  
逻辑 rules 是包含变量的条件命题，  
例如  $\forall x. \text{是人}(x) \Rightarrow \text{会死}(x)$ .  
Unification 判定一条 rule 是否可以 apply 到某逻辑命题上，  
例如：是 人(苏格拉底) 可以跟上式的左边 unify.  
Matching 的结果是得到一推 instantiated (特例化，即不包含变量) 的命题。
- **forward- or backward-chaining** (resolution):  
由已知事实 推导出新结论，或反过来，判断某给定的新结论是否成立。  
例如：是 人(苏格拉底)  $\Rightarrow$  会死(苏格拉底)  $\wedge$  是 人(苏格拉底)  
可以推出：会死(苏格拉底)。

深度学习的特点，就是将

$$\text{状态}_t \vdash \text{状态}_{t+1} \quad (2)$$

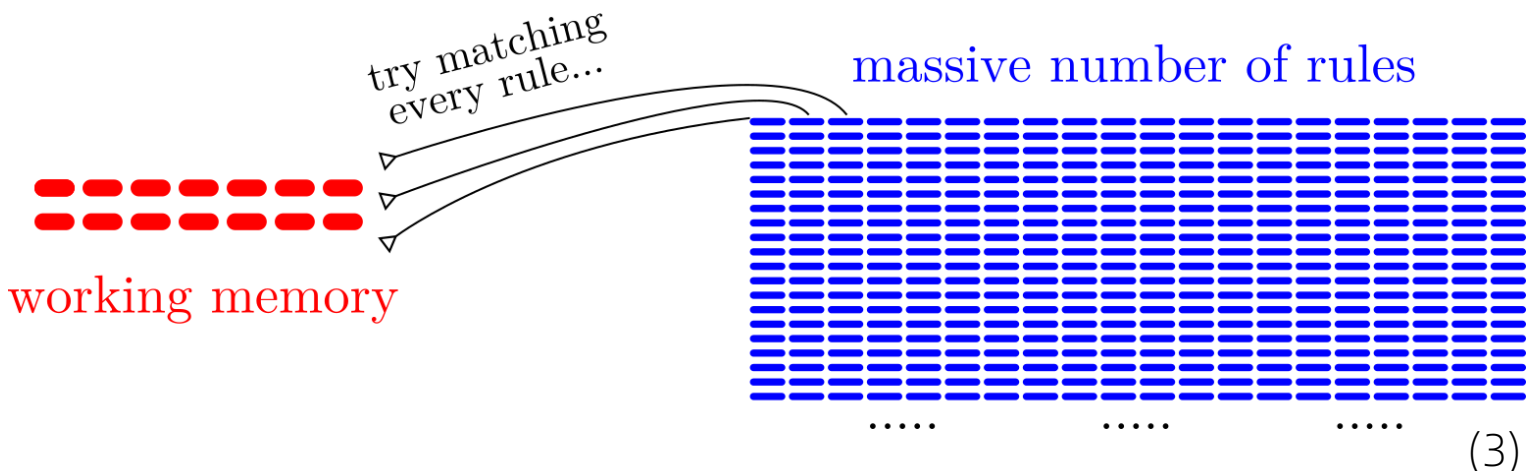
的逻辑推导过程，通通纳入进去一个非常复杂的非线性函数 (= 深度神经网络) 里面。这样做以后，上述的逻辑结构被 “mingled” 在一起，以至于很难分辨了。但也正是由于这种「大杂烩」，深度神经网络 将一套复杂的组合算法压缩成数量不算太多的一层层参数。它同时可以做 learning 和 inference 这两个动作。这种简单粗暴的方法，其实非常有效率，要超越它的速度并不容易！

我们知道 (或推测) 一个智能系统 应该具有 符号逻辑的结构。这点知识可不可以用来 约束/加速 深度神经网络？答案似乎是有可能的。现时 state-of-the-art 处理 视觉的 CNN 和 处理文字的 BERT，它们都有内部结构，而不是 fully-connected，而且 这内部结构 对应于 被处理的资料的结构。因此我们有理由相信，逻辑结构 可以用来约束 深度神经网络的结构，达到加速。

②

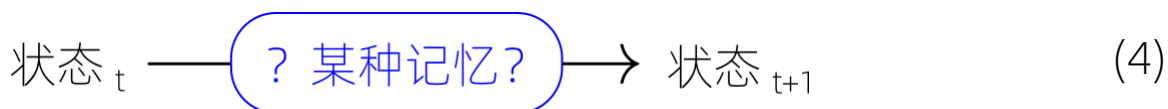
接下来我们详细一点看逻辑系统的结构：

Knowledge Base 里面有很多 rules，系统要将这些 rules 逐一 match with 系统状态 (= working memory) 里面的命题：

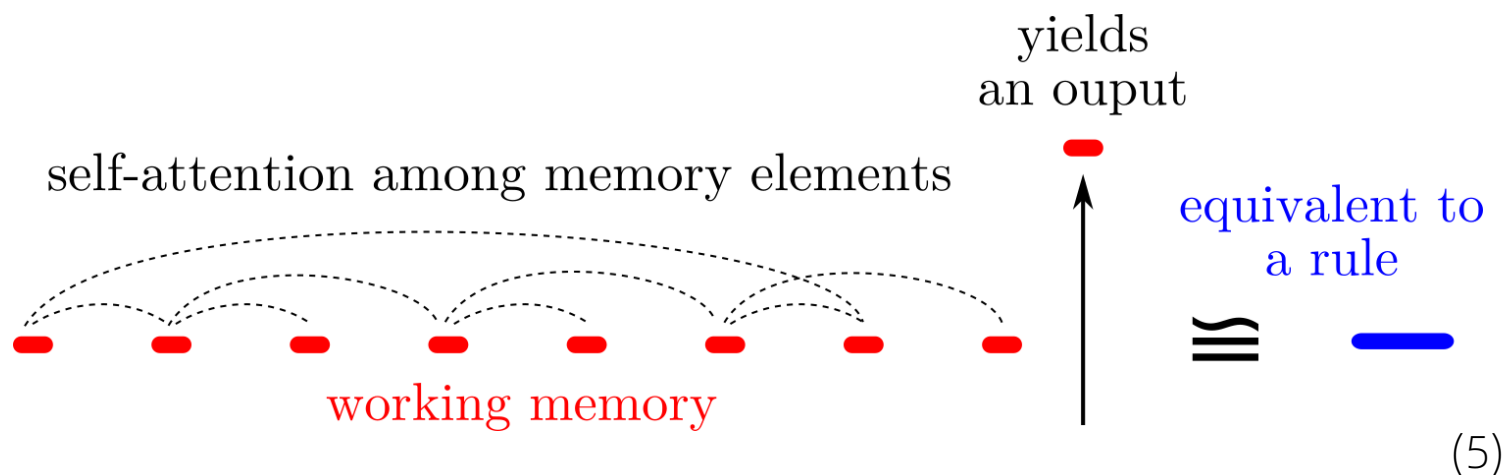


成功 matched 的 rules 可以导出新的结论，加进 working memory 的状态里面。

这个复杂的操作，完全被一个神经网络取代。或者可以更抽象地说：



而以 Transformer 来说，它是一种 输入元 之间 的记忆体（这记忆就储存在 Q, K, V 矩阵里），而它 **implicitly** 做到了 rules 的作用：



换句话说，Transformer 内部有 逻辑 rules 的结构。那么很自然的问题就是，能否发掘更多 逻辑/逻辑系统 的结构？这个问题 已部分地被 范畴逻辑 理论解决了。

那么逻辑结构 必然带来一些 代数结构，我们希望



