

# Reflective Intelligence: The Simplest Machine Model Design with Autonomous Motivation

Chengyuan Wang

Liberal Scholar, Huizhou Guangdong  
Email: 275966652@qq.com

Received: Dec. 3<sup>rd</sup>, 2017; accepted: Dec. 18<sup>th</sup>, 2017; published: Dec. 25<sup>th</sup>, 2017

---

## Abstract

From the perspective of “intelligence” definition, this paper rethinks people’s understanding of “machine emotion”. If the machine itself does not have a set of “emotional experience” mechanism, then, emotion recognition machine to human or animal, it is the extension of emotion recognition of other human or animal, the machine could not produce from itself to the human or animal emotions “empathy”. This paper puts forward an intelligent test method which is different from Turing test by introspective, speculative and commonsense reasoning, and constructs a simple and abstract model of “love and hate” based on this test method. The number of properties and number of states of the body is very few, but it is enough to explain the concrete function of each abstract property of intelligent mind. From the view of engineering that the existing computation technology, according to the nature of the design model of the position, method of collecting and processing data, then data can be enriched step by step, step by step, make the model animal more like “self motivation”, but to the surface looks more like “self conscious” animal.

## Keywords

Machine Motivation, Machine Emotion, Reflection Intelligence, Cat Intelligence Theory, Animal Intelligence

---

# 反思智能：存在自主动机的最简机器模型设计

王程远

自由学者，广东 惠州  
Email: 275966652@qq.com

收稿日期：2017年12月3日；录用日期：2017年12月18日；发布日期：2017年12月25日

## 摘要

本文从“智能”定义反思的角度，反思人对“机器情绪”的认识。如果机器本身不存在一套“自我情绪体验”的机制，那么，机器对人或动物的情绪识别，终究是人对其他人或动物的情绪识别的延伸，机器不可能产生源自自身对人或动物情绪感知的“共情”。本文用内省、思辨、常识推理的方法，提出一种有别于图灵测试的智能测试方法，并针对此测试方法，构建出一个“存在喜恶之情”的最简抽象模型。该模型的“性质体数量”和“性质体的状态数量”都极少，但足够说明构成智能心理的每个抽象性质体的具体作用。从工程实现的角度上说，在现有计算技术的支持下，按模型位置给出的性质，设计出采集、加工数据的方法，然后一步步丰富数据，便可一步步使该模型模拟的动物越来越像“有自主动机”、乃至表面看起来更像“有自我意识”的动物。

## 关键词

机器动机，机器情绪，反思智能，猫智能论，动物智能

Copyright © 2018 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

“智能”定义的反思：一只向饲主讨食的、受到惊吓会逃走的猫(金观涛, 2017) [1]，这只猫是个独立的个体，不存在超大容量的记忆数据，不会使用人类语言，不能使用任何人类认识的语言符号与人类对话，但能因此说这只猫不存在思维、不存在智能吗？答案应该是否定的。如果猫存在智能，而猫不能与人进行语言对话，那么，反思后的智能认识与图灵测试意义上的智能认识肯定是有区别的。并且可以因此推导，做智能测试，除语言问答测试(图灵测试)的方法外，还应该其他的测试方法。为区别于反思前与反思后的智能理论，暂且把反思前的理论称作“图灵智能”，反思后的认识称作“猫智能论”。

图灵智能理论是当前人工智能领域最基础的理论，图灵测试是衡量人工智能的重要方法，而猫智能论与图灵智能存在相当大的区别。猫智能论在发表后，若能经受所有质疑后依然保有一定的合理性，并在一定程度上被接受或认可。那么，在人类对智能的理解上，在拟人或动物思维的设计思路，都将产生巨大而深远的影响。而这种影响，应该是属于进步性的。

## 2. 猫智能论与图灵智能的具体区别

图灵机是一种做符号运算的机器，其内置的逻辑及运算的符号，对于机器本身(硬件体)而言，是没有意义的。所以，本质上，图灵计算机是一种基于人类思维做符号演进运算的机器，其所有演算逻辑都是人类赋予的，所有演算目的都是为了给人类提供服务的，是不存在自主动机的。而猫思维里的逻辑不是人类赋予的，认知符号(比如猫会把饲主当成一个符号)也不是基于人类思维逻辑进行演进的，而是猫自己生成的，并且对于猫身体本身(硬件体)而言是具有意义的。那么，猫的思维是基于什么性质生成的呢？根据猫会讨食和会逃跑之类的现象推测：猫智能论所指的智能体应该是具有自主动机、自主需求的，是存在喜恶之情的。存在喜恶之情的智能机器会根据收集到的数据自动生成一套“方法论”，并根据其特有的方法论，在特定的环境下，做出相应的判断和行为，以满足智能体自己的喜好之情。因此，猫智能论

认为，一个更像人类或动物的思维(智能)系统的学习过程，其采集、分析数据的最初动机和最终目的都不是为了给设计它的人提供服务的，而是在基于先为其自己服务的基础上，才决定是否为人服务的。比如，母猫生了小猫，小猫并非天生下来就会为母猫提供服务的，小猫如果为母猫提供服务，那是基于母猫喂养小猫的“恩情”，或者说基于母猫对小猫的“教育”。

基于上述理由，猫智能论进一步认为：若为服务他人而采集、分析生成的“机器智能个性(功能)”，是根据被服务的大众或个人的性格而来，也即，此时的智能体的“个性”并非属于智能体自己，而是属于被服务的大众或个人，本质上是属于人的喜恶思维的延伸。这样的机器人(智能体)，处理数据是以“人的目的”为“目的”的。猫智能论认为这样的智能不属于具有自主动机的智能，也即，不属于“拥有动机智能的独立个体”。猫智能论认为，有自主动机需求的机器，即，为了满足自己的需求才为人类提供服务的机器，才是更接近真实意义的人或动物机器。事实上，这也是符合客观事实的，比如，一个人与另一个人合作的最初动机并不是为另一个人服务的，而是为自己服务的。例如，a与b合作，a帮助b是基于“为换取能满足a自身需求的资源”的前提条件下为b提供服务的。

### 3. 猫智能论的智能测试方法

基于猫智能论对智能定义的不同理解，猫智能论测试某物体是否存在动机智能的方法当然也与图灵测试有所不同。具体方法：比如，测试一只猫(包括生物猫和机器猫)是否存在智能。不一定是采用问答模式进行测试的，即不一定是通过问猫，“一加一等于几”、“你会不会下棋”之类问题；而是测试，“猫是否存在动机需求，并是否会根据历史记忆自动生成有利于满足自己需求的方法”，如果会，则存在猫智能论意义上的智能，反之则不存在智能。例如，在猫(包括机器猫)肚子饿需要进食的时候，某人给猫喂食，然后观察猫能否会记住某人，观察猫下次肚子饿时是否会向某人讨要食物，如果猫会向某人乞讨食物，则猫存在智能，反之者则不存在智能。或者，某人鞭打猫使猫痛苦，观察猫是否会记住此特性，当猫再次看到某人靠近时，猫是否会为避免再次被鞭打而做出相应反应，比如逃跑，如果猫会做出逃跑行为则存在智能，反之则不存在智能；

会向“曾经喂过自己的人”讨食的猫，会躲避“鞭打过自己的人”的猫，不管是否具备与人进行语言交流的能力，都是存在智能的。在此基础上，如果还能“与人进行对话”，比如猫是否会对鞭打过自己的人说“请不要伤害我”，则这不是“是否存在智能”的问题，而是“智商高低”的问题。“能与人进行对话的猫”当然是比“不能与人进行对话的猫”的智商更高的。

### 4. 动物思维现象的总结

基于上述观点，更相似于人或动物的智能机器应该是存在“基于自己需求的动机机制”的，这个机制的作用具有以下功能：

- 以“保护自己的机体不被破坏”及“使自己机体能保持长期正常运行”为目的。
- 当身处正在破坏自己机体的环境时，会感到“舒服程度更低”；当身处于满足自己生理需求的环境时会感到“舒服程度更高”。这个性质是先天存在的，无需学习的，无需记忆参与的。
- 会自动记录造成舒服程度转变的环境的数据。
- 当再次处于相同或类似的，使自己感到“舒服程度更低”或“舒服程度更高”的环境时，会根据历史记忆，避开“舒服程度更低”的环境，走向“舒服程度更高”的环境。

### 5. 推测存在自主动机的手机的特征

基于上述理由，用“手机”类比于“猫”，推测有“自主动机”的手机应该会有以下的现象特征：

- 当手机电量低到一定程度的时候，在不需要经过记忆的情况下，“电量越往下低”，手机就会感到“舒服程度越低”。手机的舒服程度越低，在需要经过记忆的情况下“越迫切渴望充电”，如果手机通过记忆判断为“能找到充电途径”，则舒服程度相对更高，反之则舒服程度相对更低。
- 手机会请求人帮自己充电。然后手机会记住“所有请求过的人”、“曾经帮过自己充电的人”、“不愿帮自己充电的人”。当再次需要充电时，手机会倾向于请求帮助过自己的人。
- 手机会记住“充电的位置”。如果在一个对人来说是陌生的环境里，人不知道哪里有电源插座，手机会根据历史记忆“主动”告知人，插座在什么地方。
- 如果手机有“运动能力”，当人带手机充过一次电后，手机会记住这个过程，并在下次需要充电的时候，自行走到电源处对自己进行充电。
- 如果手机会奔跑，手机奔跑撞上石头，在不需要经过记忆的情况下，舒服程度降低，手机会记住这个过程，以后不再奔跑着撞向石头。
- 手机发现曾经故意伤害过自己“生理健康”的人，手机会“主动远离”这个人。
- 手机不一定会使用语言，但会用行动表达喜恶之情。比如：不愿帮手机充电的人要使用手机，手机会拒绝提供服务。
- 手机不一定能识别到人的情绪(舒服程度)变化，但人一定可以识别到手机的情绪变化。比如，不愿帮手机充电的人要强行使用手机，手机会“不高兴”而产生“挣扎”行为。
- 在手机清醒(未关机)时拆解手机，手机会感到“舒服程度降低”。如果手机有运动能力，手机还会挣扎、逃走。
- 手机的语言不一定是源于人设计手机时内置的，还可能是人通过“奖罚”教会手机的。如果能通过奖罚教会手机使用语言，或者手机因其自身的“舒服程度”转变而“自悟”出形容情绪的语言，那么，手机与人交流时，双方都将更有“共鸣”或者“共识”。

## 6. 与猫思维活动相关的性质分析及性质的表示字母

与猫思维活动相关的性质分析：至少存在以下性质：

- 猫存在一个身体。这个性质用“J”表示。
- 猫存在感官，包括眼、耳、口、鼻、肤、四肢、生殖器，等等。这些感官会产生不同的感觉状态。这个性质用字母“G”表示。
- 猫存在于一个被猫感官感知的外部世界，这个外部世界也即日常所说的“世间万物”。这个性质用字母“B”表示。
- 猫能控制自己的感官(G)在外部世界(B)里的运动方向。这个性质用字母“H”表示。
- 猫感官的不同状态会产生不同的感知，进而会使猫产生不同的舒服程度。比如：猫饿的时候进食，会比不进食时“舒服程度更高”；猫的皮肤受伤时，会比不受伤时“舒服程度更低”；等等。即，不同的感官状态会造成不同的舒服程度状态(情绪状态)。不同的感官状态造成不同舒服程度，是不需要经过记忆的。这个性质用“W”表示。
- 机体状态对感官状态是有要求的，不同的机体状态，对感官状态的要求是不一样的。比如，猫在不同的机体状态里，同样的感官状态，舒服程度是不同的。比如，同样是进食，“在肚子饿的时候进食”，和“在肚子饱的时候进食”，前者的舒服程度更高。机体状态对感官状态存在要求，这个性质用字母“M”表示。
- 猫存在记忆。这个性质用字母“Z”表示。
- 猫的记忆可分为“被回忆起的记忆”(或称“被实时感知的记忆”)和“未被回忆起的记忆”。其中“被

回忆起的记忆”的性质，用字母“S”表示。

- 猫通过记忆，识别出伤害过自己的人，需要逃跑。如果猫通过记忆判断为“能逃跑”，相比于判断为“不能逃跑”的“舒服程度更高”；如果猫通过记忆判断为“不能逃跑”，相比于判断为“能逃跑”，则会因不能摆脱恐惧而“舒服程度更低”。不同的记忆识别判断，会造成不同的舒服程度感知，这是需要记忆参与的，这个性质用“X”表示。
- 猫的记忆里，存在“遇到伤害过自己的人需要逃跑”“肚子饿需要进食”之类需求记忆。这个性质用字母“T”表示。

### 7. 存在自主动机的最简抽象智能机器及运行原理

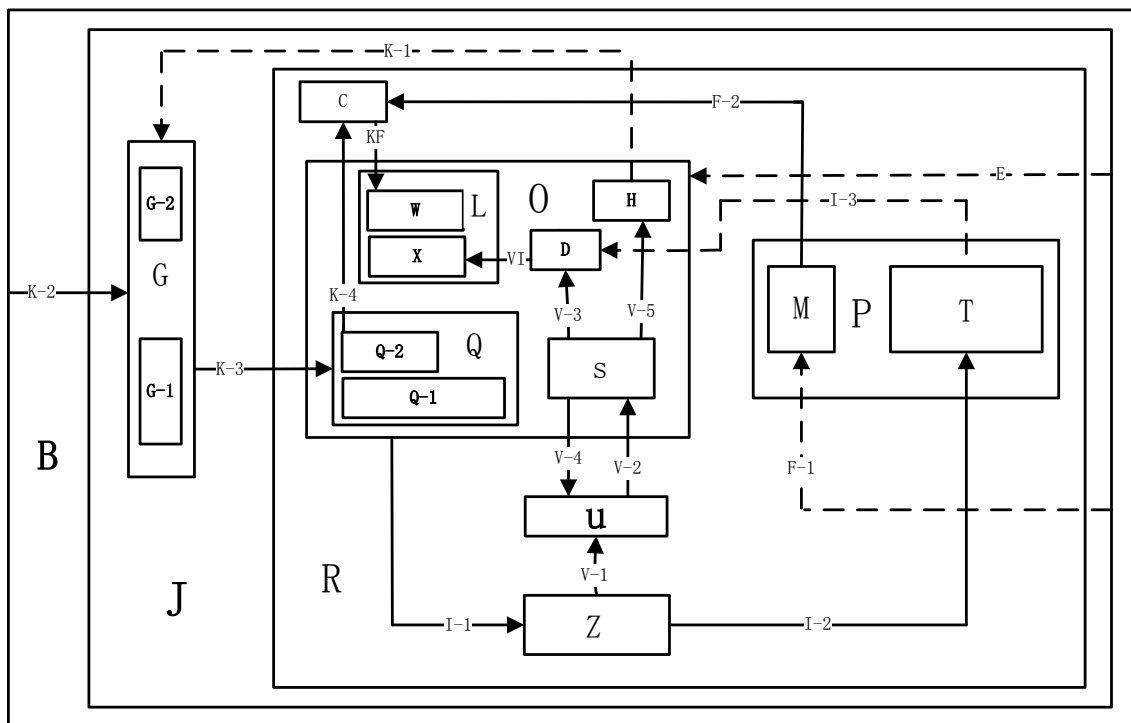
根据上述观点及分析，设计一个不一定成熟，但有一定开拓、探索意义的运行结构：“存在自主动机的最简抽象智能机器及运行原理”：

#### 7.1. 最简结构体图示：图 1(a)~(c)

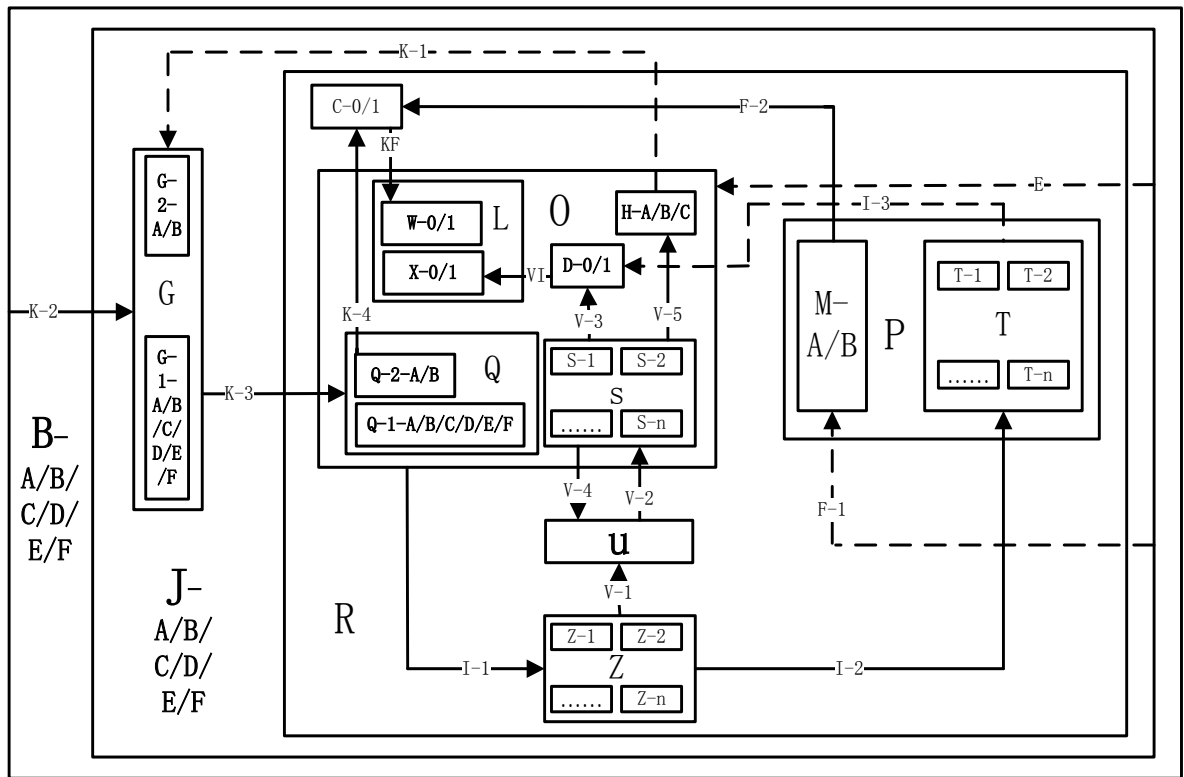
如图 1(a)~(c)。

#### 7.2. 最简结构体内符号的性质和关系讲解

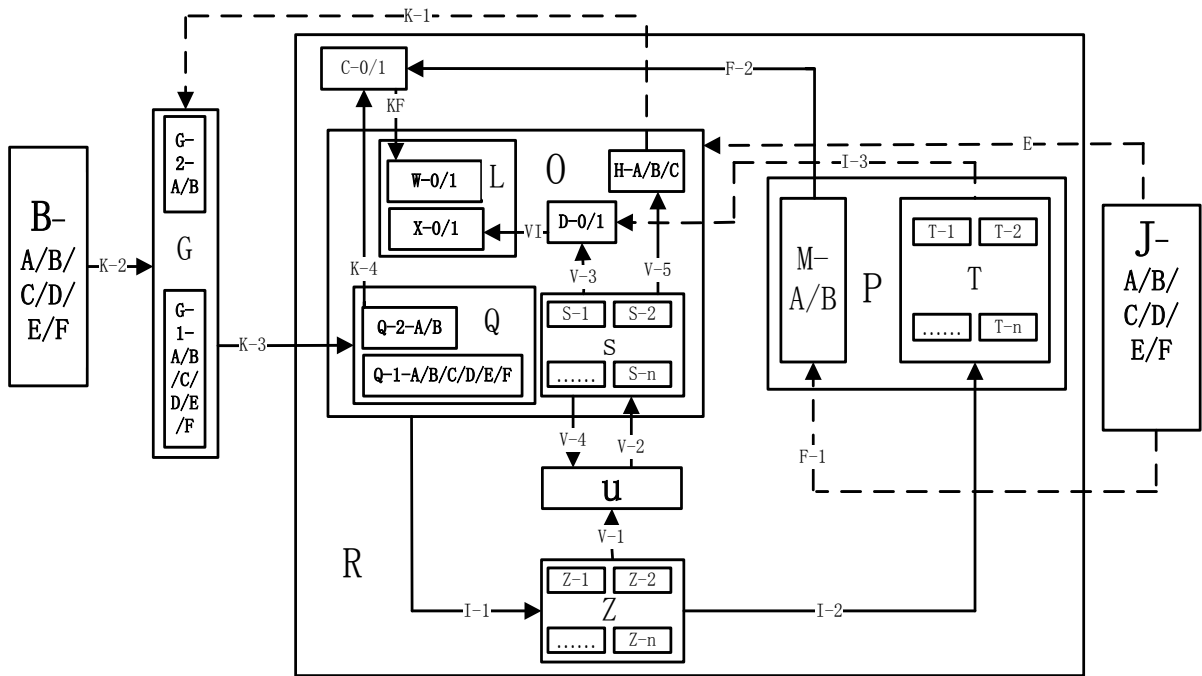
- H 有 H-A、H-B、H-C 三种状态。H 不能同时存在“H-A、H-B、H-C”其中的两种及两种以上状态。
- B 有 B-A、B-B、B-C、B-D、B-E、B-F 六种状态。B 不能同时存在“B-A、B-B、B-C、B-D、B-E、B-F”其中的两种及两种以上状态。B 各状态间的转变规律：B-A 只能直接转变为 B-F 或 B-B，B-B 只能直接转变为 B-A 或 B-C，B-C 只能直接转变为 B-B 或 B-D，B-D 只能直接转变为 B-C 或 B-E，B-E 只能直接转变为 B-D 或 B-F，B-F 只能直接转变为 B-E 或 B-A (如图 2)。
- G-1 有 G-1-A、G-1-B、G-1-C、G-1-D、G-1-E、G-1-F 六种状态。G-1 不能同时存在“G-1-A、G-1-B、



(a)



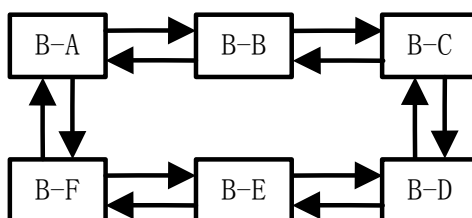
(b)



(c)

图注：图 1(a)~(c)，是描绘同一个意思的图，可以理解为同一幅图。形态区别在于：(a) 没有标出细节；(b)、(c)是完整图；(c)为了方便观察把 B 和 J 分别移到了左右两边；图 1(a)~(c)统称“最简结构图(体)”。

Figure 1. Simplest structure  
图 1. 最简结构体



图注：(1) 各状态只能直接转变成相邻的状态。(2) 如果要转变成不相邻的状态，则必需一步接一步转变为相邻的状态，直到转变为目标状态为止。比如 B-A 要转变成 B-D，必需先转为 B-B 或 B-F。再从 B-B 到 B-C 再到 B-D，或者再从 B-F 到 B-E 再到 B-D。

**Figure 2.** Schematic diagram of transition relations of “B” states

**图 2.** “B” 各状态的转变关系示意图

G-1-C、G-1-D、G-1-E、G-1-F” 其中的两种及两种以上状态。

- G-2 有 G-2-A、G-2-B 两种状态。G-2-A、G-2-B 不能同时存在。
- Q-1 有 Q-1-A、Q-1-B、Q-1-C、Q-1-D、Q-1-E、Q-1-F 六种状态。Q-1 不能同时存在 “Q-1-A、Q-1-B、Q-1-C、Q-1-D、Q-1-E、Q-1-F” 其中的两种及两种以上状态。
- Q-2 有 Q-2-A、Q-2-B 两种状态。Q-2-A、Q-2-B 不能同时存在。
- M 有 M-A、M-B 两种状态。M-A、M-B 不能同时存在。
- C 有 C-0、C-1 两种状态。C-0、C-1 不能同时存在。
- D 有 D-0、D-1 两种状态。D-0、D-1 不能同时存在。
- W 有 W-0、W-1 两种状态。W-0、W-1 不能同时存在。
- X 有 X-0、X-1 两种状态。X-0、X-1 不能同时存在。
- T 包括 T-1、T-2、T-3 至所有 T，图中用.....(省略号)和 T-n 表示。
- S 包括 S-1、S-2、S-3 至所有 S，图中用.....(省略号)和 S-n 表示。
- G 包括 G-1、G-2。G-1、G-2 同时存在。
- Q 包括 Q-1、Q-2。Q-1、Q-2 同时存在。
- L 包括 W、X。W、X 同时存在。
- L 有 L-(W-0/X-0)，L-(W-0/X-1)，L-(W-1/X-0)，L-(W-1/X-1) 四种状态，分别简化写做 L-0/0，L-0/1，L-1/0，L-1/1。L 不能同时存在 “L-0/0，L-0/1，L-1/0，L-1/1” 其中的两种及两种以上状态。
- Z 存在多少种状态则 U 存在多少状态，U 本身不能同时存在两种及两种以上状态。
- Z 包括 Z-1、Z-2、Z-3 至所有 Z，图中用.....(省略号)和 Z-n 表示。Z-1、Z-2、Z-3、.....、Z-n 不能同时存在两种及两种以上状态。
- T、U、S 皆由 Z 填充。即，T-1、T-2 至 T-n，U，S-1、S-2 至 S-n 由 Z-1、Z-2、Z-3、.....、Z-n 填充。
- T 通过 I-3 填充到 S，写作 T-S 或 S-T。
- 通过 V-1、V-2 填充到 S，并与 S-T 相互对应的 Z，写作 S-TZ。
- Q、S、L、H 组成 O。
- M、T 组成 P。
- Q-2 和 M 组成 C。

- S-T 与 S-TZ 组成 D。
- O、P、U、Z 组成 R。
- G、R 组成 J。
- J、非 J 组成 B。
- Y=肯定, N=否定。

### 7.3. 运行原理

- 1) J 通过 F-1 塑造 M 状态。
- 2) B 通过 K-2 塑造 G 状态。

a) B 与 G-1 的状态关系: B-A 则 G-1-A, B-B 则 G-1-B, B-C 则 G-1-C, B-D 则 G-1-D, B-E 则 G-1-E, B-F 则 G-1-F。

b) B 与 G-2 的状态关系: B-A 则 G-2-A, B-B 则 G-2-A, B-C 则 G-2-A, B-D 则 G-2-B, B-E 则 G-2-B, B-F 则 G-2-B。

c) G-1 与 G-2 的状态关系: G-1-A 则 G-2-A, G-1-B 则 G-2-A, G-1-C 则 G-2-A, G-1-D 则 G-2-B, G-1-E 则 G-2-B, G-1-F 则 G-2-B。

- 3) J 体通过 E 线, 控制 O 开启或关闭。
- 4) O 开启, H 必然通过 K-1 控制 G。

若 H 为 H-A, G 在 B 各状态间运动, B 必然通过 K-2 依次把 G-1 循环塑造为 G-1-A、G-1-F、G-1-E、G-1-D、G-1-C、G-1-B、G-1-A(如图 3(a)); 若 H 为 H-C, G 在 B 各状态间运动, B 必然通过 K-2 依次把 G-1 循环塑造为 G-1-A、G-1-B、G-1-C、G-1-D、G-1-E、G-1-F、G-1-A(如图 3(b)); 若 H 为 H-B, G 停止运动, G 停留在某个 B 状态, 也即 G-1 状态停止转变, 停留在某个固定状态。

- 5) O 开启, G 必然通过 K-3 塑造 Q 状态。G 与 Q 的状态关系:

a) G-1-A 则 Q-1-A, G-1-B 则 Q-1-B, G-1-C 则 Q-1-C, G-1-D 则 Q-1-D, G-1-E 则 Q-1-E, G-1-F 则 Q-1-F。

b) G-2-A 则 Q-2-A, G-2-B 则 Q-2-B。

- 6) O 开启, Q-2 必然通过 K-4 与 M 通过 F-2 共同塑造 C 状态。

“Q-2、M”与“C”的状态关系:

- a) 若 Q-2 与 M 为“Q-2-A/M-B”或“Q-2-B/M-A”, 则 C 状态为 C-0。
- b) 若 Q-2 与 M 为“Q-2-A/M-A”或“Q-2-B/M-B”, 则 C 状态为 C-1。

- 7) O 开启, C 必然通过 KF 塑造 W 状态。

C 与 W 的状态关系: C-0 则 W-0, C-1 则 W-1。

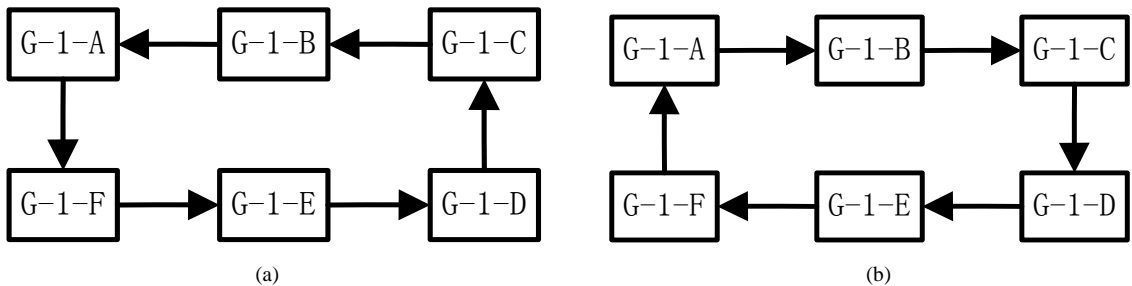


Figure 3. (a) The changing law of “G-1” state in “H-A”; (b) The changing law of “G-1” state in “H-C”

图 3. (a) “H-A” 时, “G-1” 状态的转变规律; (b) “H-C” 时, “G-1” 状态的转变规律



8) O 开启, O 状态必然通过 I-1 塑造 Z 状态。

9) O 状态塑造 Z 状态。其中, “W-0 转变为 W-1” 或 “W-1 转变为 W-0” 过程中的关键 Z 将被抽离出来, 通过 I-2 写到 T。

10) Z 通过 V-1 塑造 U 状态。

11) O 开启, U 必然通过 V-2 塑造 S 状态。

12) O 开启, S-TZ 必然通过 V-3, 与 T 通过 I-3 塑造的 S-T, 共同塑造 D 状态。

a) D 状态的区分: 比如: 假设 “XX” 表示 S-T 任意内容, 写作 S-T-XX, 与之对应的是 S-TZ-XX。不管 XX 是什么内容, 都转换成 N、Y 进行相符与否对比。若 S-TZ-XX 与 S-T-XX 不相符, 写作 S-TZ-XX-N/S-T-XX-Y; 若 S-TZ-XX 与 S-T-XX 相符, 写作 S-TZ-XX-Y/S-T-XX-Y。

b) “S-T、S-TZ” 与 “D” 的状态关系:

i)  $D-(S-TZ-XX-N/S-T-XX-Y)$  则  $D-(TZ-N/T-Y) = D-(N/Y) = D-0\% = D-0$ 。

ii)  $D-(S-TZ-XX-Y/S-T-XX-Y)$  则  $D-(TZ-Y/T-Y) = D-(Y/Y) = D-100\% = D-1$ 。

13) O 开启, D 必然通过 VI 塑造 X 状态。

D 与 X 的状态关系: D-0 则 X-0, D-1 则 X-1。

14) O 开启, S 必然通过 V-5 控制 H。

控制的原则是提升 L(W/X), 防止 L 下跌。

15) 根据前面规律可知:

a) “W、G-1” 的状态为 “W-1、G-1-A” 或 “W-1、G-1-B” 或 “W-1、G-1-C”, 则 “G-2、M” 的状态为 “G-2-A、M-A”。

b) “W、G-1” 的状态为 “W-1、G-1-D” 或 “W-1、G-1-E” 或 “W-1、G-1-F”, 则 “G-2、M” 的状态为 “G-2-B、M-B”。

c) “W、G-1” 的状态为 “W-0、G-1-A” 或 “W-0、G-1-B” 或 “W-0、G-1-C”, 则 “G-2、M” 的状态为 “G-2-A、M-B”。

d) “W、G-1” 的状态为 “W-0、G-1-D” 或 “W-0、G-1-E” 或 “W-0、G-1-F”, 则 “G-2、M” 的状态为 “G-2-B、M-A”。

16) 近路原则:

有两条或两条以上路径可以提升 L 或防止 L 下跌, 必然选择最近的路径的原则称作近路原则。S 控制 H 必然遵循 “近路原则”。比如 “W-0、G-1-C” 时, S 必然是控制 G-1-C 转变为 G-1-D 以达到 W-0 转变为 W-1 的目的, 而不是选择依次转变为 G-1-B、G-1-A、G-1-F。“W-0、G-1-A” 时必选 G-1-F, “W-0、G-1-D” 时必选 G-1-E, “W-0、G-1-F” 时必选 G-1-A, 都是同样道理。

17) T 的形成原理、S 控制 H 的原理、X 的转变原理图(图 4), 以及相关问题的讲解:

a) 在 Z、T 为空时, 也即没有任何符号记忆时, 一旦 O 开启, O 便有了状态输入, Z 便有符号记忆。O 状态一旦因 G、H 的状态转变而转变后, Z 便有了 Q、H 状态转变的符号逻辑记忆。Z 有了 Q、H 状态转变的符号逻辑记忆, 并有了 W-0 与 W-1 的记忆后, 而仍在 T 为原始状态, 也即 T 没有任何因 O 输入的内容而塑造出的内容时, T 的内容为 “寻找 ‘控制 H、G, 以使 W-0 转变为 W-1’ 的符号路径”。

b) G 状态转变的原因, 除来自 H 状态的控制外, 还来自不受 H 直接控制的 B 的状态转变。

注: 实际上, G 状态也是不受 H 直接控制的, H 直接控制的是 G 在 B 的运动方向, 让 G 运动至 B 的各状态里去获得 G 的某种特定状态。而在最简结构图的描述里, G 状态好像直接受 H 控制, 其实实际表达的是 G-1 状态直接受 H 控制, G-2 则不直接受 H 控制。在最简结构体里, G-1 状态之所以直接受 H 的控制, 是为描述 T 性质。若没有确定的 H 状态与 G 状态的关系, 就没有可确定的符号控制逻辑, 也即

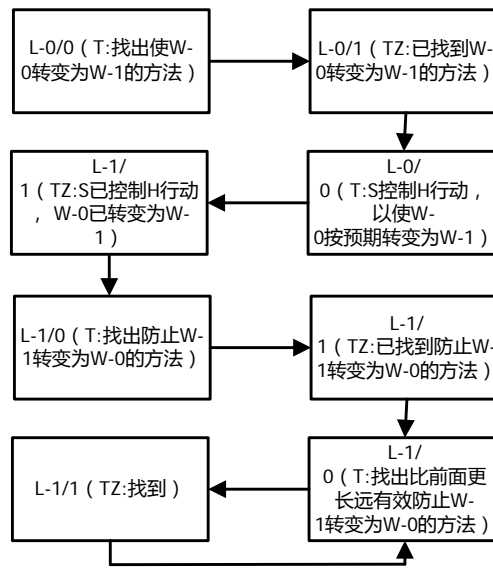


Figure 4. The formation principle of “T”  
图 4. “T” 的形成原理

无法形成用于控制 H 的 T。要描述 T 性质必需存在“最少、但足够”的“H 状态与 G 状态是绝对因果关系的性质”。在最简化性质结构体里，如果去除 G-1，只留 G-2 以及 H 状态和 G-2 状态的绝对因果关系性质，那么不管 B 有多少种状态，G 都只有两种状态，在 W 为 W-1 时将会因缺乏“预警状态”而“无法完整描述 T 性质”，也即无法描述“防止 W-1 转为 W-0”的性质。

c) W 状态实质由 Q-2 状态和 M 状态决定。H 不能直接塑造 Q-2 和 M 的状态，H 只能通过塑造 G-1 状态间接塑造 W 状态。S 不能直接塑造 G-1 状态，S 只能通过塑造 H 间接塑造 G-1 状态进而塑造 W 状态。S 根据 T，从 Z 找出控制 H 符号路径，并依此对 H 进行操控以达到塑造 W 状态的目的。

d) 满足 S-T 的过程，是一个记忆思维判断过程，所以 S 并非一定要对 H 有所操控的。满足 S-T 的过程有三种形式：

- i) S 只对 H 有所操控。
- ii) S 对 U 和 H 都有所操控。
- iii) S 只对 U 进行了操控。

18) S 控制 H 的原理的具体表现(如图 5~16)。

“————→”和“-----▶”为 S 控制 H 的示意线，示意线的前端和后端分别为控制发生前和发生后的状态。

#### 7.4. 最简结构体内，各符号对应的“动物体”或“人工智能动物体”里的命名

- B: 外界
- J: 身体、机体
- R: 脑
- O: 感知体

注：“主观感知”存在于 O，O 承载“主观感知”。J 内的 O 的“主观感知”、R 内的 O 的“主观感知”、O 本身的“主观感知”，组成承载“主观感知”的 O 的任意一个极端小的点，对于“主观感知”都属于“外界”。用日常语言的说法：自己看自己的脑，自己的脑不属于主观感知、不属于自己、不属

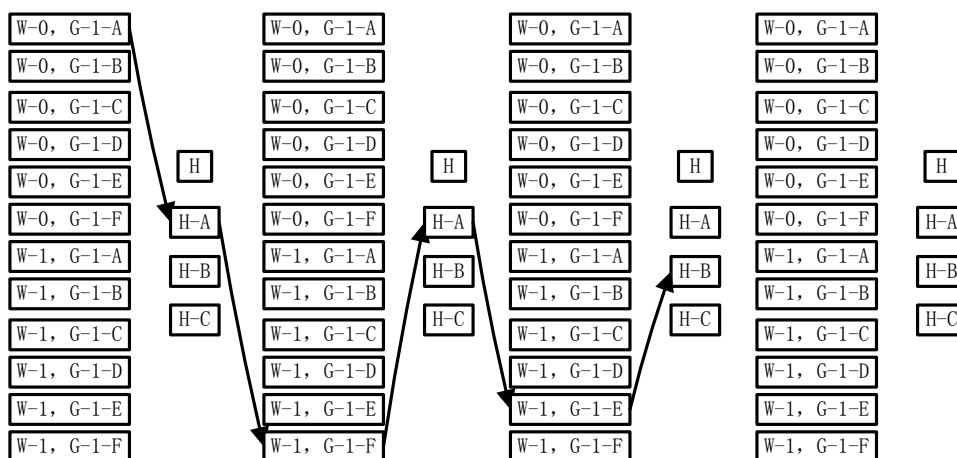


Figure 5. The transition process of “H” state in “W-0, G-1-A”

图 5. “W-0, G-1-A” 时, “H” 状态的转变过程

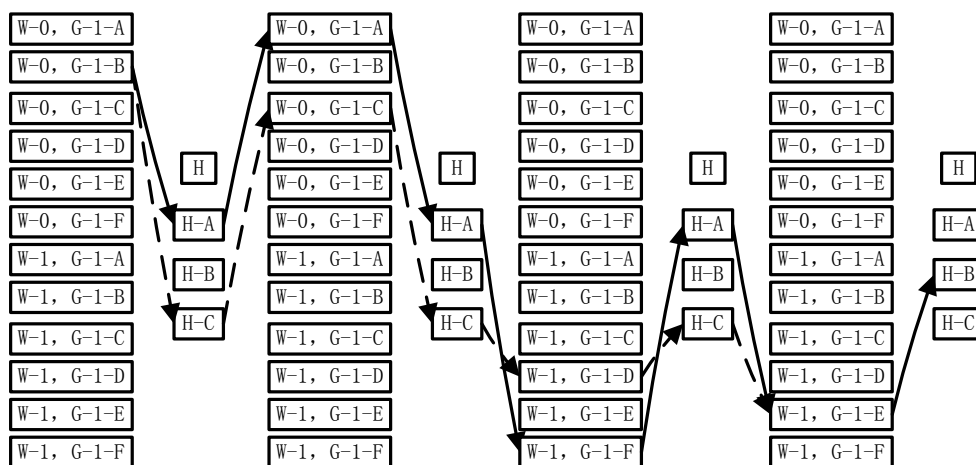


Figure 6. The transition process of “H” state in “W-0, G-1-B”

图 6. “W-0, G-1-B” 时, “H” 状态的转变过程

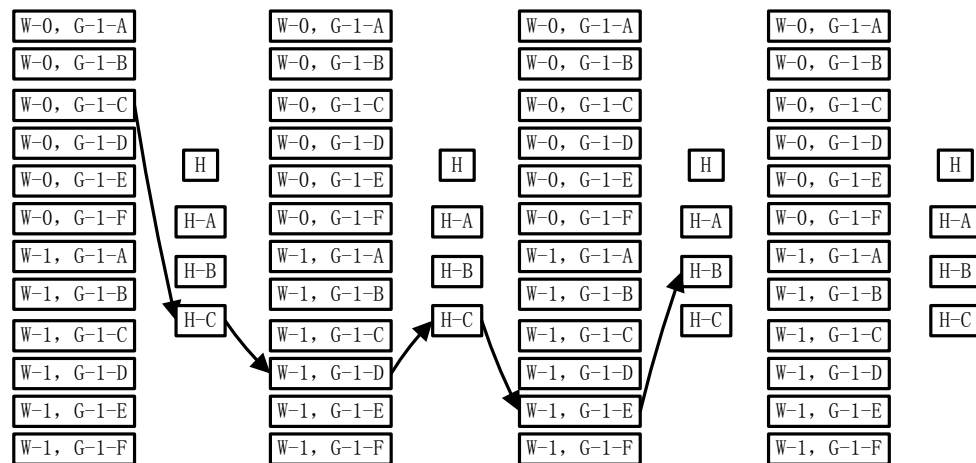


Figure 7. The transition process of “H” state in “W-0, G-1-C”

图 7. “W-0, G-1-C” 时, “H” 状态的转变过程

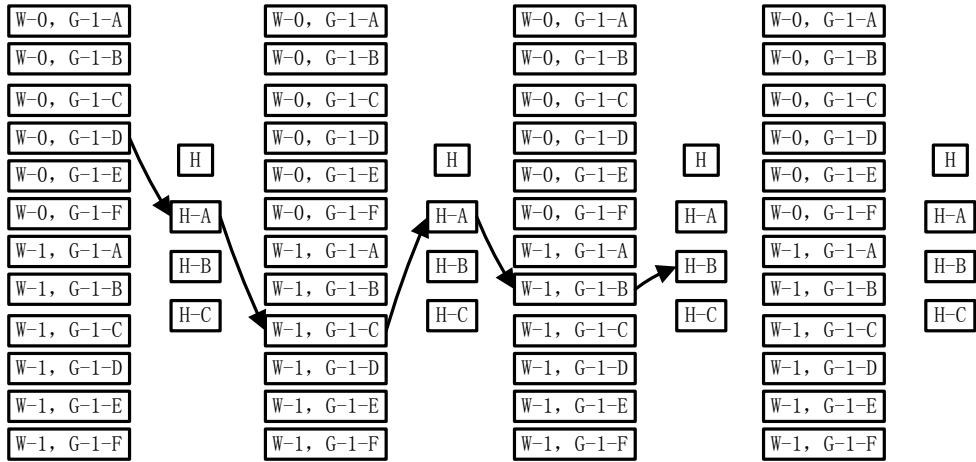


Figure 8. The transition process of “H” state in “W-0, G-1-D”

图 8. “W-0, G-1-D” 时, “H” 状态的转变过程

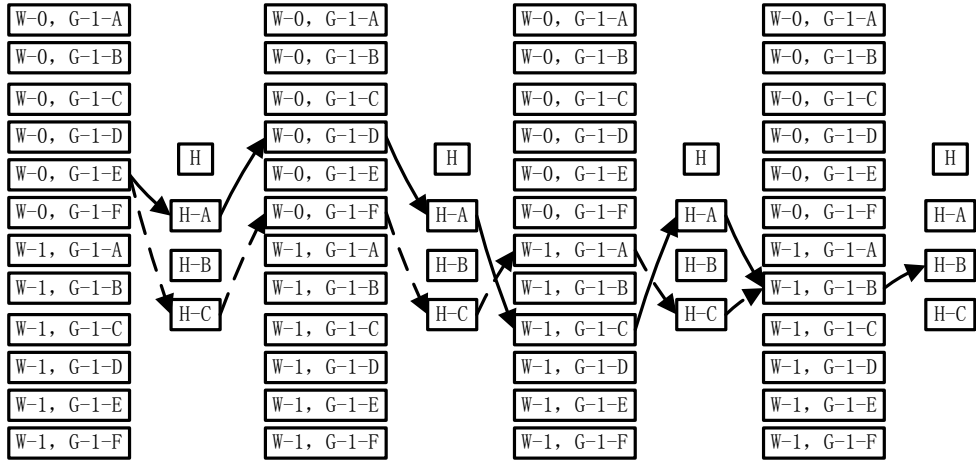


Figure 9. The transition process of “H” state in “W-0, G-1-E”

图 9. “W-0, G-1-E” 时, “H” 状态的转变过程

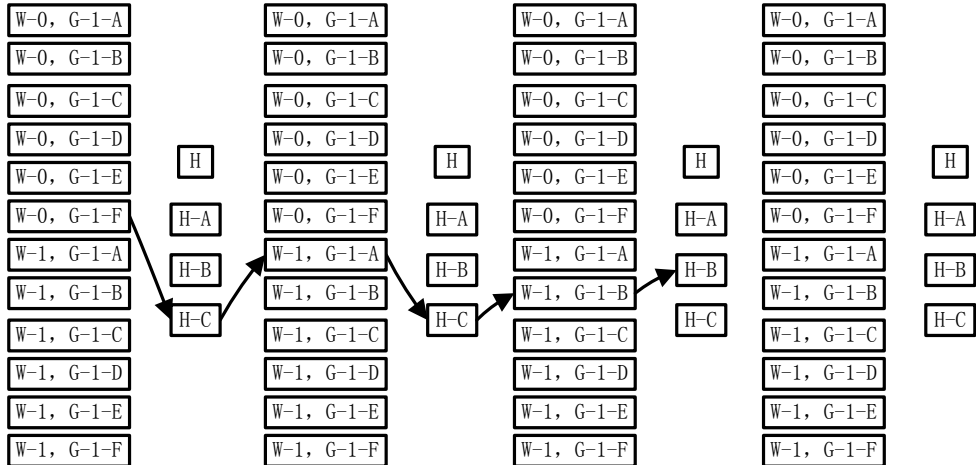


Figure 10. The transition process of “H” state in “W-0, G-1-F”

图 10. “W-0, G-1-F” 时, “H” 状态的转变过程

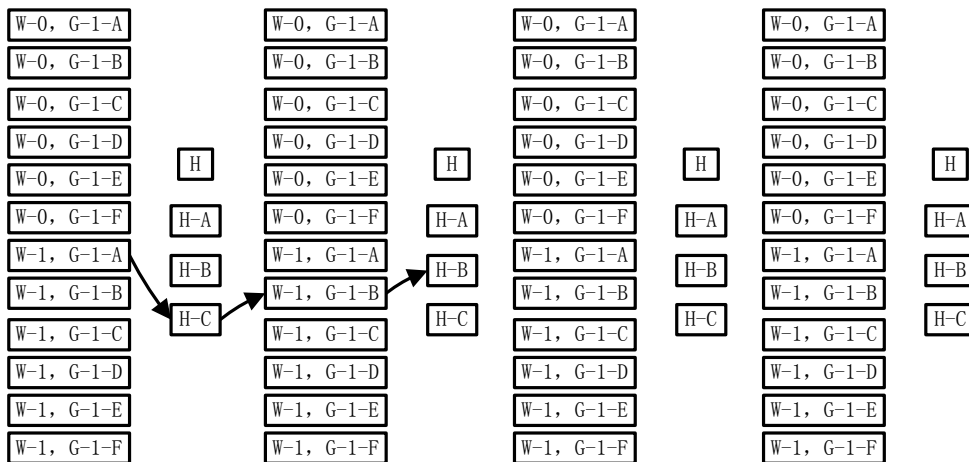


Figure 11. The transition process of “H” state in “W-1, G-1-A”

图 11. “W-1, G-1-A” 时, “H” 状态的转变过程

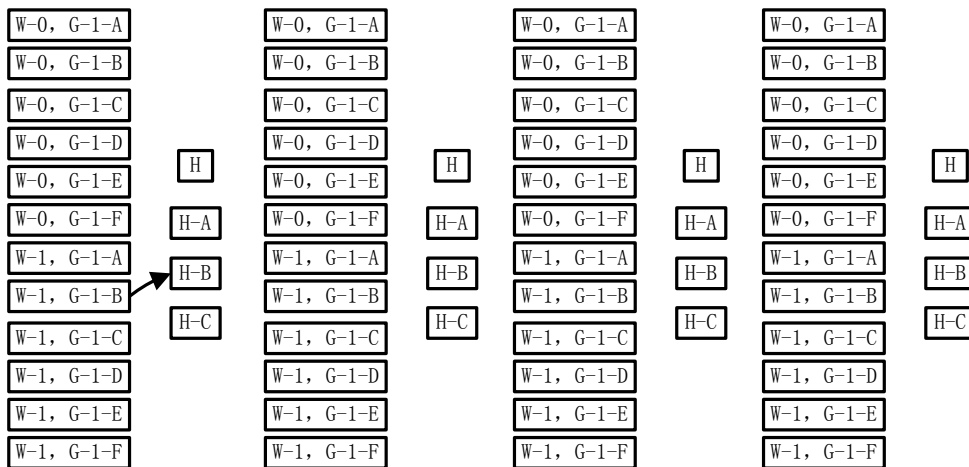


Figure 12. The transition process of “H” state in “W-1, G-1-B”

图 12. “W-1, G-1-B” 时, “H” 状态的转变过程

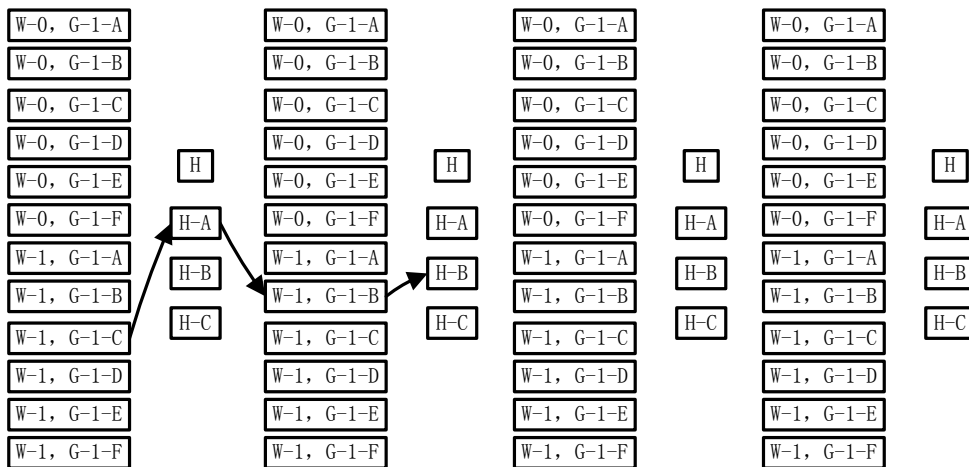


Figure 13. The transition process of “H” state in “W-1, G-1-C”

图 13. “W-1, G-1-C” 时, “H” 状态的转变过程

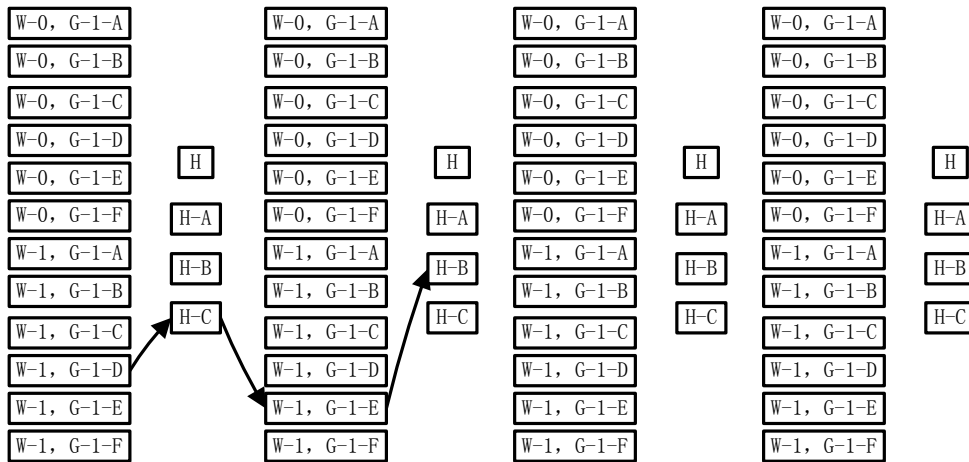


Figure 14. The transition process of "H" state in "W-1, G-1-D"

图 14. "W-1, G-1-D" 时, "H" 状态的转变过程

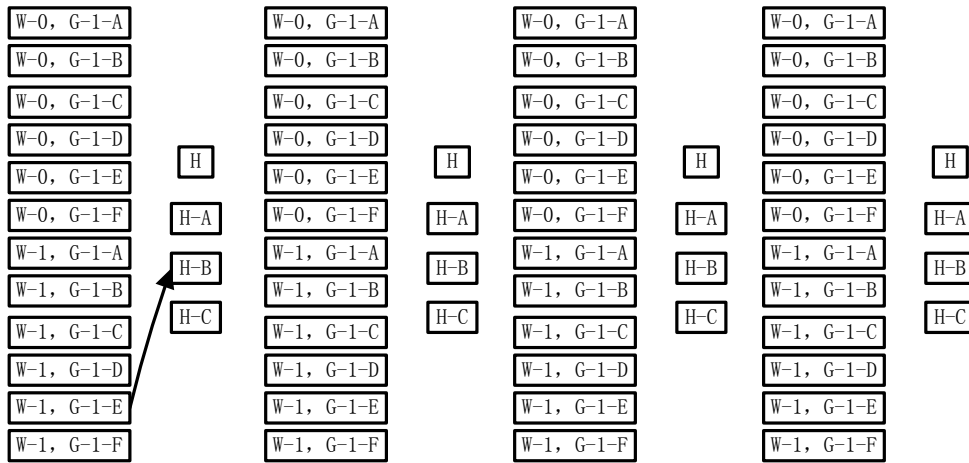


Figure 15. The transition process of "H" state in "W-1, G-1-E"

图 15. "W-1, G-1-E" 时, "H" 状态的转变过程

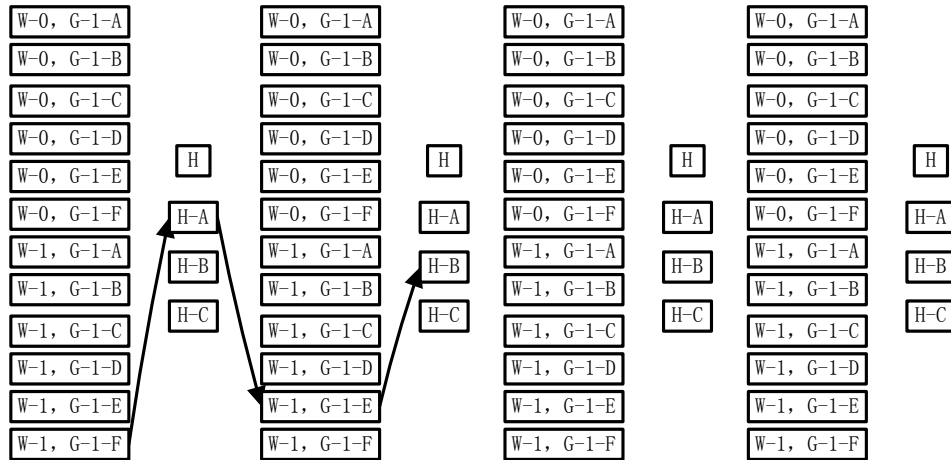


Figure 16. The transition process of "H" state in "W-1, G-1-F"

图 16. "W-1, G-1-F" 时, "H" 状态的转变过程

于“我”。自己的脑属于外界，属于被主观观察的客观现象。但是自己看自己的脑产生的主观心理感知改变，脑却会有相应的物理状态改变。

- G: 外感官

注:

1) 在最简结构体里，外感官只有 6 种状态：“G-1-A, G-2-A”、“G-1-B, G-2-A”、“G-1-C, G-2-A”、“G-1-D, G-2-B”、“G-1-E, G-2-B”、“G-1-F, G-2-B”。但在“复杂体”里，比如在猫身体里，可以存在几千万亿亿亿种不同状态。

2) G 对应于人或动物时，是各类感觉器官，例如眼、耳、口、鼻等等；G 对应于机器时，则是各类模仿人或动物感官功能的传感器。

- Q: 外感官感知体

- H: 控制体

- M: 非记忆需求体

- T: 记忆需求体

- C: “外感官感知体状态”与“非记忆需求体状态”的“重叠程度体”(“Q 状态”与“M 状态”的“重叠程度体”)

- D: “记忆识别状态”与“记忆需求状态”的“重叠程度体”(“S-T 状态”与“S-TZ 状态”的“重叠程度体”)

- P: 动机体、需求体

- S: 记忆数据运算体

- L: 舒服程度感知体

注:

1) 在最简结构体里，L 只有四种程度状态，但在“复杂体”里，可以有几千甚至几万种状态。

2) L 上升，大约等于“变得舒服”“变得更舒服”“变得快乐”“变得更快乐”“变得开心”“变得更开心”“变得高兴”“变得更高兴”等等；L 下跌，大约等于传统意思“变得难受”“变得更难受”“变得痛苦”“变得更痛苦”“变得不舒服”“变得更不舒服”等等。

3) 如果事物不会对 L 程度造成升跌波动，就不会对事物作美、丑、好、坏、喜爱、厌恶，等等之类的性质区分。之所以对事物有偏好之情，是因为事物会造成 L 程度升跌。

4) 使 L 上升的事物，包括日常所说的：美好的、美丽的、甜美的、美妙的、漂亮的、可爱的、善良的、好吃的、好玩的、美的、好的，等等；使 L 下跌的事物，即是日常所说的：丑恶的、丑的、难看的、可恶的、讨厌的、恶劣的、歹毒的、苦涩的、难吃的，等等。

- W: 非记忆舒服程度感知体

- X: 记忆舒服程度感知体

- Z: 记忆体

注：关于 Z、T、S、D、U 的关系。

1) 从符号记忆的角度概括，T 和 Z 是一体的，Z 包含 T。从符号“是否存在需求因素”的角度概括，T 符号是 Z 符号里“存在需求因素”的那部分。

2) 从是否被感知的角度概括，D 和 S 是一体的，S 体包含 D 体。从“是否存在来自 T 的符号与来自 Z 的符号对比的因素”的角度概括，D 是 S 里“存在 T 符号与 Z 符号对比因素”的那部分。

3) 从符号记忆的角度概括，U、S、Z (包括 T) 是一体的，Z 包含 U、S。从“是否存在被感知因素”的角度概括，U、S 是 Z 里“存在被感知因素”的那部分。

4) 从符号存在被感知因素的角度概括，U、S 是一体的。从“是否是最后进入 S 的符号”的角度概括，U 是 S 里“最后进入 S”的那一个符号。

## 8. 最简结构体复杂化模拟动物智能的升级过程

模拟“有记忆动物”，应从简单一步步往复杂方向模拟。比如从蜗牛开始模拟，到鱼，到猫，到猩猩，直到抽象结构体成为一个表面看起来“独立的”、能自主悟出含情绪含义的语言并用于形容自己的感受、说出的语言能与人类产生“高度共鸣”的、拥有超高智商的“人”为止。

## 9. 结语

猫智能论的核心思想是在计算机内置一个“舒服程度”数值器，机器用“提升舒服程度”和“防止舒服程度下跌”做“行动原则”，以决定自己的运动方向。并以此用于模拟动物对事物的喜恶偏好现象，或者说模拟动物的情绪体验。本质上并非机器真的拥有了自我意识，只是在一定程度上提升了机器行为与动物行为的相似度。

## 参考文献 (References)

[1] 金观涛. 反思“人工智能革命” [EB/OL]. <http://www.aisixiang.com/data/105451.html>, 2017-08-08.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2326-3415, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [airr@hanspub.org](mailto:airr@hanspub.org)