

H. A. Simon 学术思想研讨会

司马贺的人工智能创新之路

史忠植

shizz@ics.ict.ac.cn

中国科学院计算技术研究所



中科院计算所

• INSTITUTE OF
COMPUTING
• TECHNOLOGY

内容提要

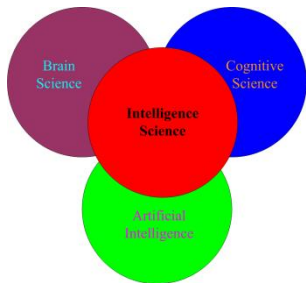
概述

符号主义创始人

决策理论开拓者

逻辑理论家发明人

科学知识发现倡导者



纪念司马贺100周年华诞

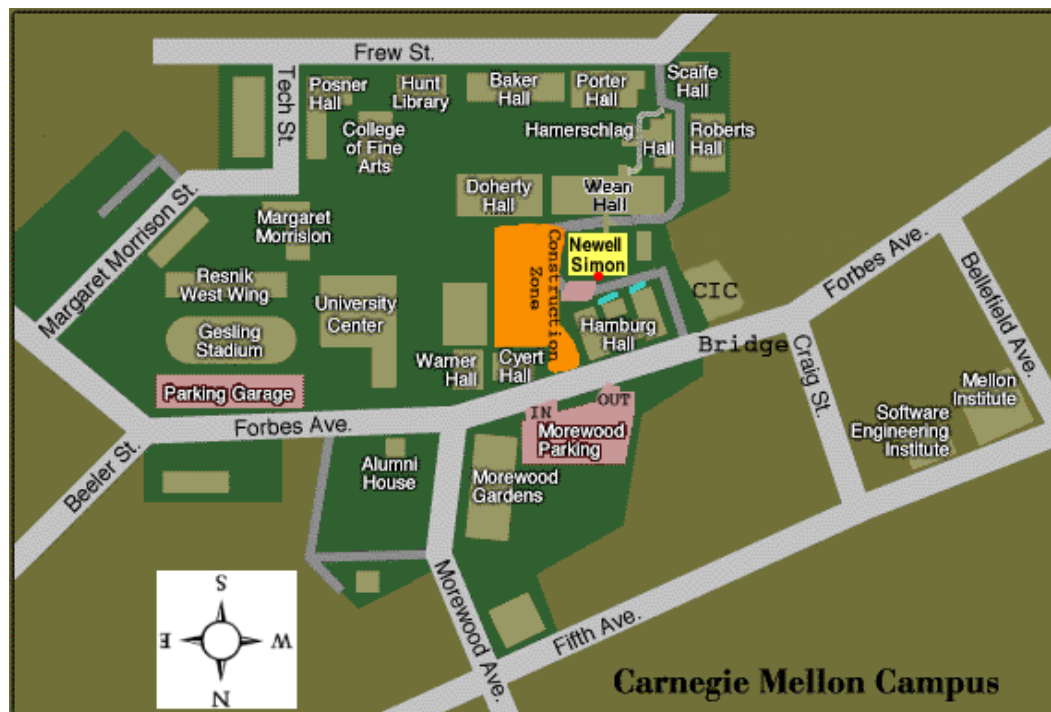
穿越人生歧途花园



学习巨擘创新之路

1916年6月15日生于美国威斯康星州密尔沃基市。

司马贺在卡内基—梅隆大学



1949–2001年在卡内基—梅隆大学（CMU，宾州的匹兹堡）任教，担任行政学和心理学教授、企管系主任、计算机系教授。

歧路花园



司马贺 (Herbert. A. Simon) 把人生比作一个“歧路花园”，在穿越这个迷宫式花园的过程中，他为自己打造了一连串令人眼花缭乱的身份，并对涉足过的每个领域都产生了革命性的影响。

博学的司马贺

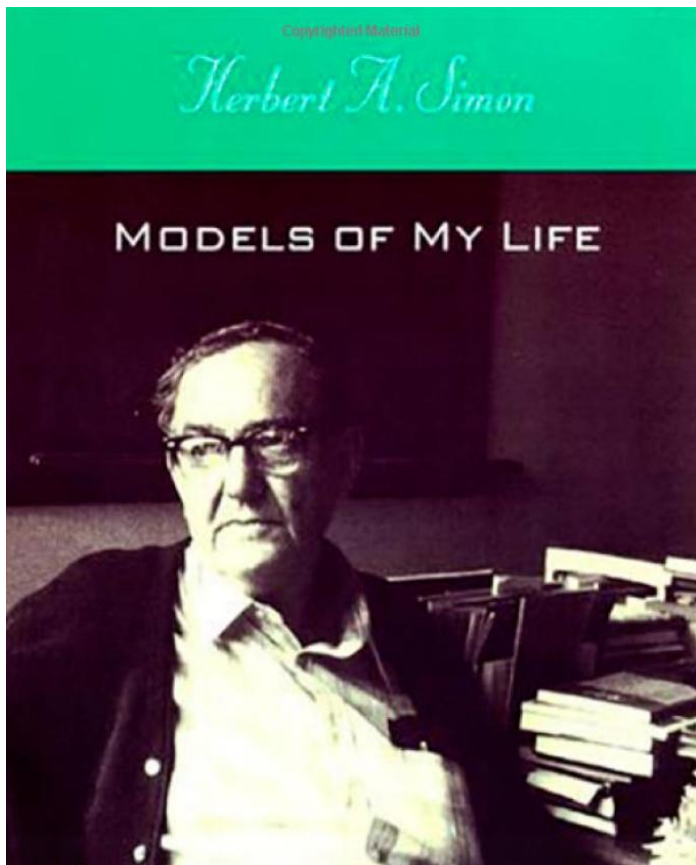
获得过9个博士头衔：

- 1943年的芝加哥大学政治学博士学位
- 1963年凯斯工学院科学博士学位
- 1963年耶鲁大学科学博士学位
- 1963年法学博士学位
- 1968年瑞典伦德大学哲学博士学位
- 1970年麦吉尔大学法学博士
- 1973年鹿特丹伊拉斯莫斯大学经济学博士
- 1978年米之根大学法学博士
- 1979年匹茨堡大学法学博士

司马贺获奖

- 1943 芝加哥大学政治学博士学位
- 1969 因心理学方面的贡献获得杰出科学贡献奖
- 1975 因认知心理学和表处理的贡献他和纽厄尔获得图林奖
- 1978 由于“经济组织内的决策过程进行的开创性的研究”，荣获诺贝尔经济学奖
- 1986 因行为学方面的成就获得美国全国科学家奖章

人工智能的创新贡献



司马贺
H. A. Simon

- 符号主义创始人
- 决策理论开拓者
- 逻辑理论家发明人
- 科学知识发现倡导者

内容提要

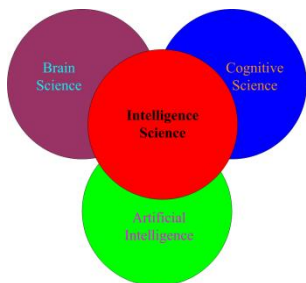
概述

符号主义创始人

决策理论开拓者

逻辑理论家发明人

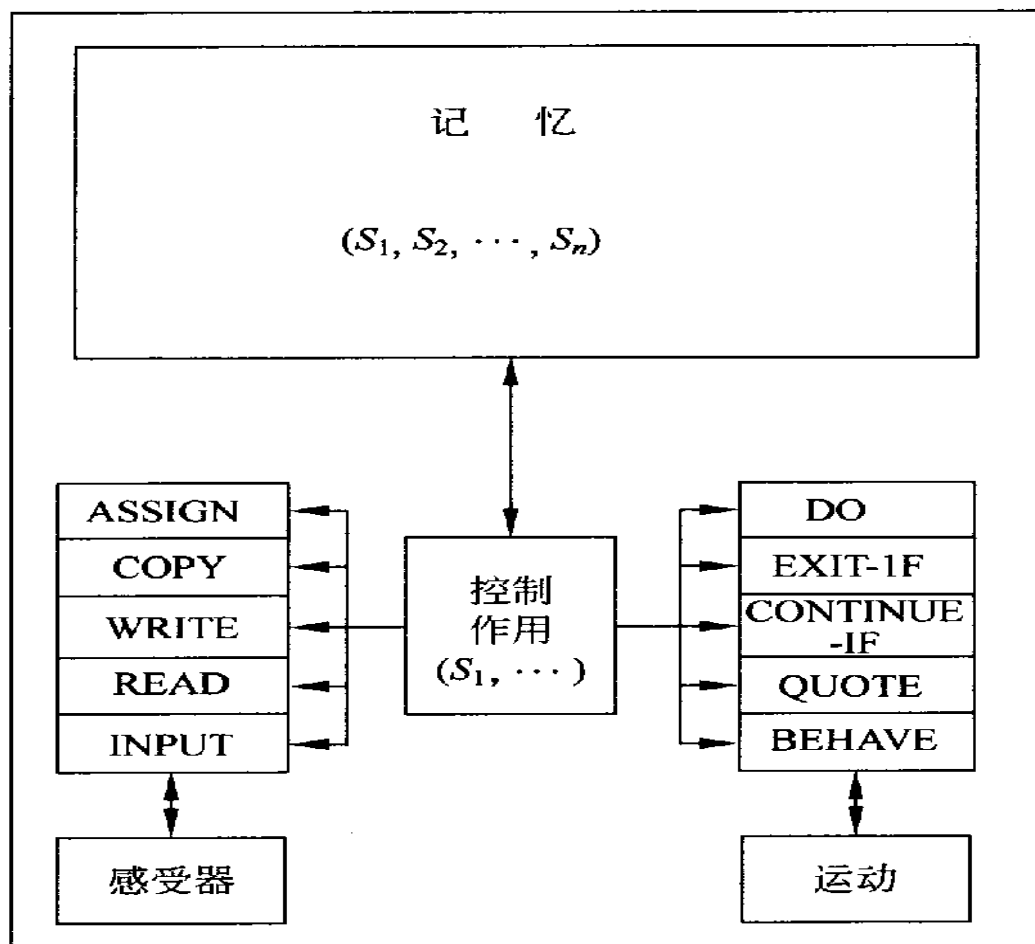
科学知识发现倡导者



物理符号系统

- 把人看成一个信息加工系统，常称作物理符号系统。用物理符号系统主要是强调所研究的对象是一个具体的物质系统，如计算机的构造系统，人的神经系统、大脑神经元等。
- 所谓符号就是模式；任何一个模式，只要它能和其它模式相区别，它就是一个符号。符号既可以是物理的符号，也可以是头脑中的抽象的符号。
- 物理符号系统的基本任务和功能就是辨认相同的符号和区分不同的符号。

物理符号系统



物理符号系统

10种操作符:

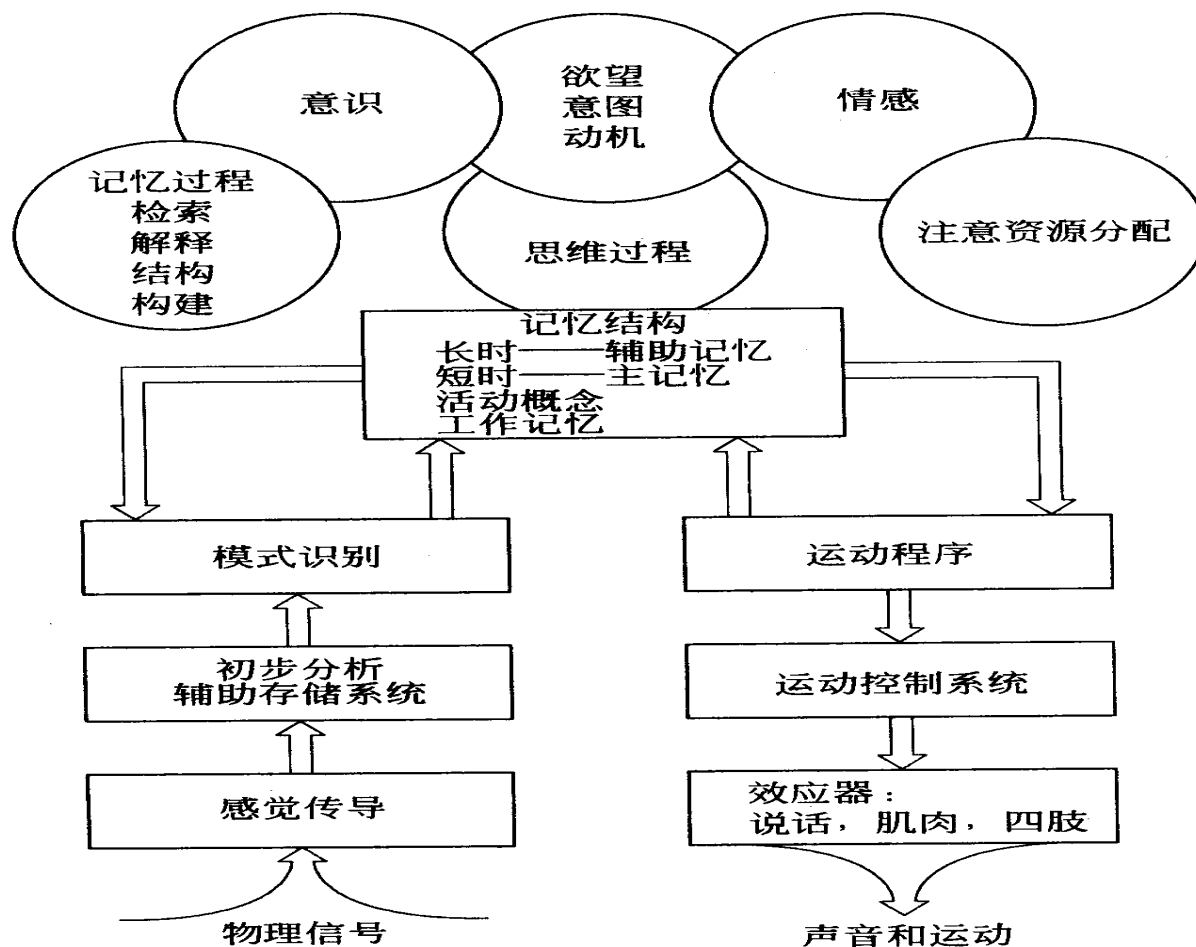
- (1) 赋值符号 (Assign a symbol):
- (2) 复制表达式。
- (3) 写表达式 (Write an expression)
- (4) 写 (Write)
- (5) 读 (Read)
- (6) 执行序列 (Do sequence)
- (7) 条件退出和条件继续 (Exit-if and Continue—if)
- (8) 引用符号 (Quote a symbol)
- (9) 外部行为 (Behve externally)
- (10) 环境输入 (Input from enuironment)

物理符号系统

物理符号系统的功能简化成6种，即：

- (1) 输入符号
- (2) 输出符号
- (3) 存储符号
- (4) 复制符号
- (5) 建立符号结构：通过找到各种符号之间的关系，在符号系统中形成符号结构。
- (6) 条件转移：如果在记忆中已经有了一定的符号系统，再加上外界的输入，就可以继续完成行为。

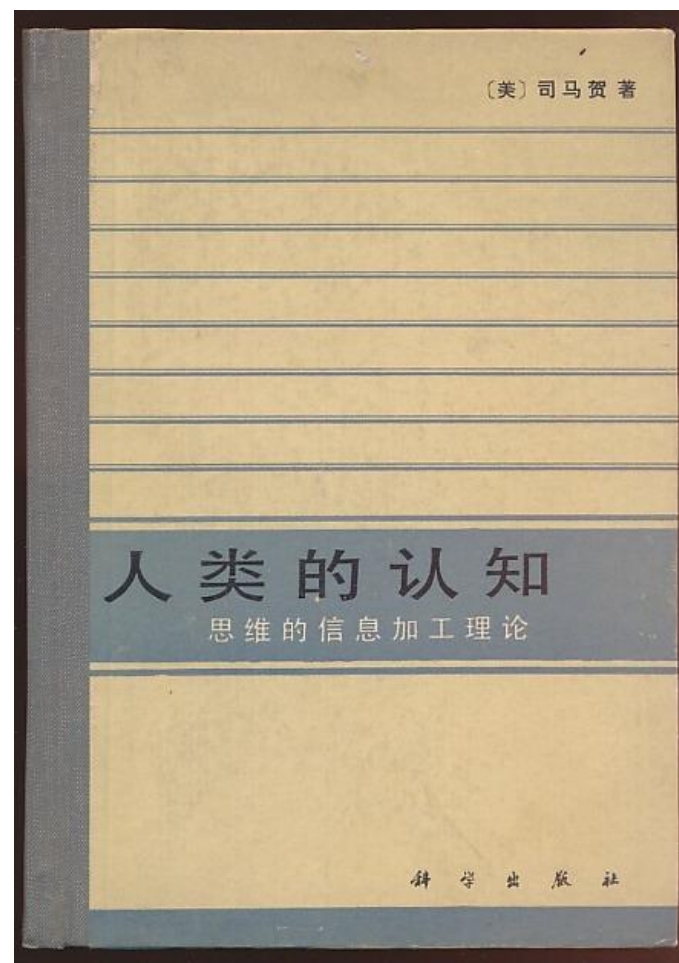
纯认知系统模型



1975年ACM图灵奖

纽威尔和司马贺的贡献

人类认知心理学 和表处理



•2016/6/16

•史忠植 司马贺的创新之路

物理符号系统假设

1976年纽威尔和司马贺提出了物理符号系统假设，说明物理符号系统的本质。主要假设内容如下：

- 物理符号系统假设，物理系统表现智能行为必要和充分的条件是它是一个物理符号系统。
- 必要性意味着表现智能的任何物理系统将是一个物理符号系统的例示。
- 充分性意味着任何物理符号系统都可以进一步组织表现智能行为。
- 智能行为就是人类所具有的那种智能：在某些物理限制下，实际上所发生的适合系统目的和适应环境要求的行为。

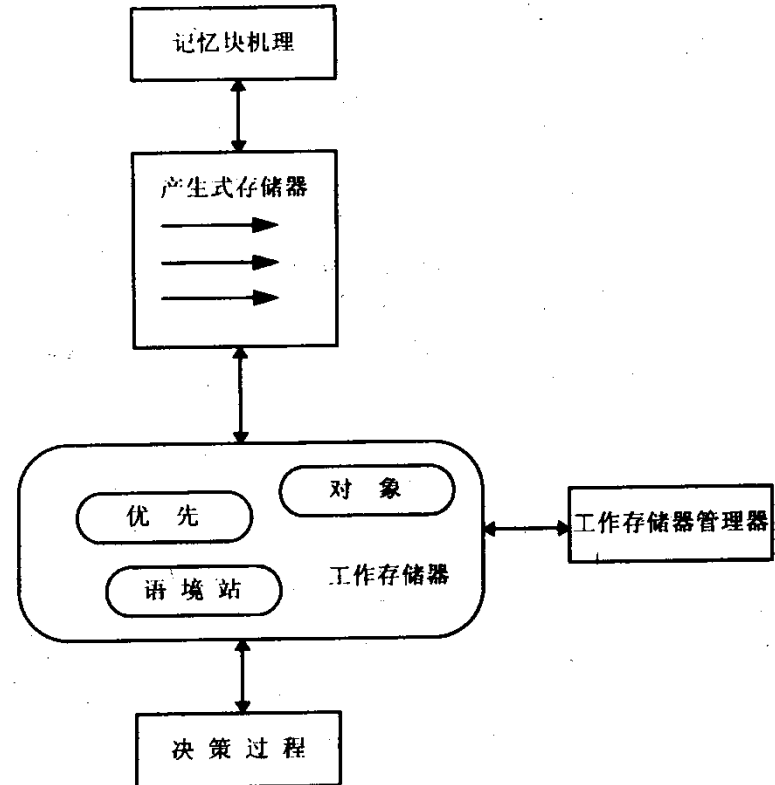
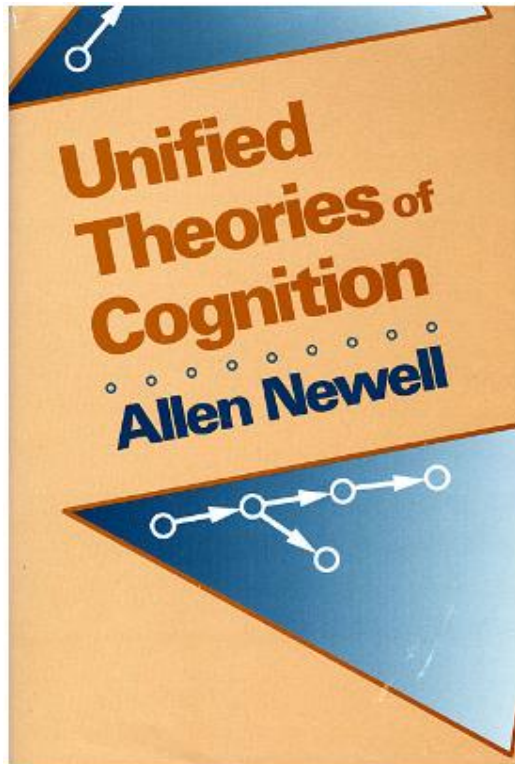
学术影响

- 物理符号系统假设为符号主义，即传统的人工智能奠定了理论基础。
- 创建了信息加工心理学，开辟了从信息加工观点研究人类思维的方向。
- 推动了认知科学和人工智能的发展。

组块理论

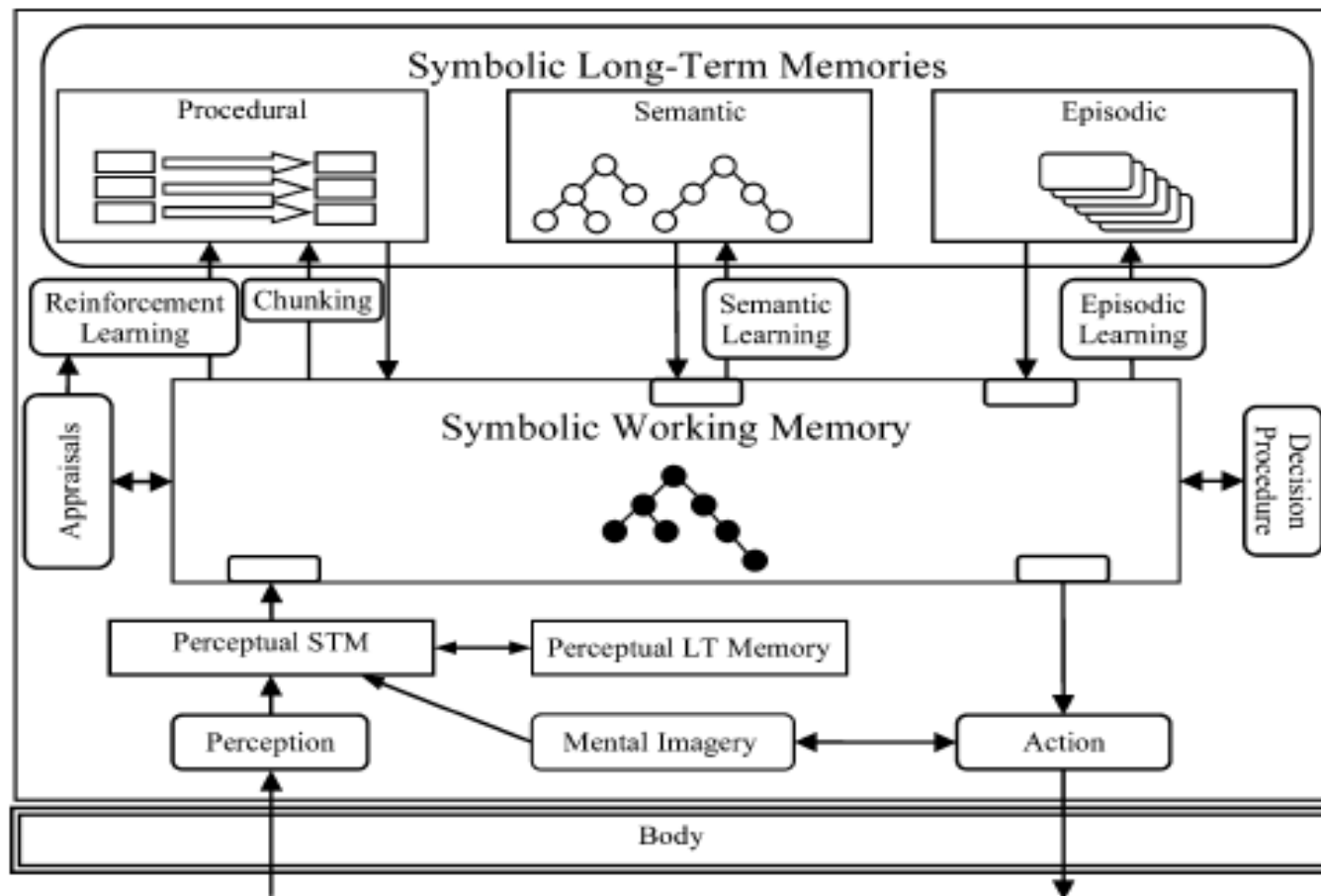
- 1956年，米勒在其著名论文“神奇数字7加减2，我们加工信息的能力的某些限制”中，明确指出短时记忆的容量为 7 ± 2 ，即一般为7，在5——9之间波动。
- 1973年，基于短时记忆的容量为 7 ± 2 个项，司马贺提出组块理论（chunking theory）

SOAR



J. Laird, A. Newell, and P. Rosenbloom, SOAR: An Architecture for General Intelligence. Artificial Intelligence, 33, 1987

SOAR9



组块理论

- 组块理论对自然语言处理具有重要的理论意义和实用价值，在自然语言理解和机器翻译中发挥作用。
- 组块策略就是将零散的构件组成有意义的单元。从信息加工的角度来看，组块是人对信息进行组织或再编码。

内容提要

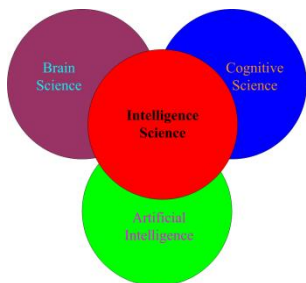
概述

符号主义创始人

决策理论开拓者

逻辑理论家发明人

科学知识发现倡导者

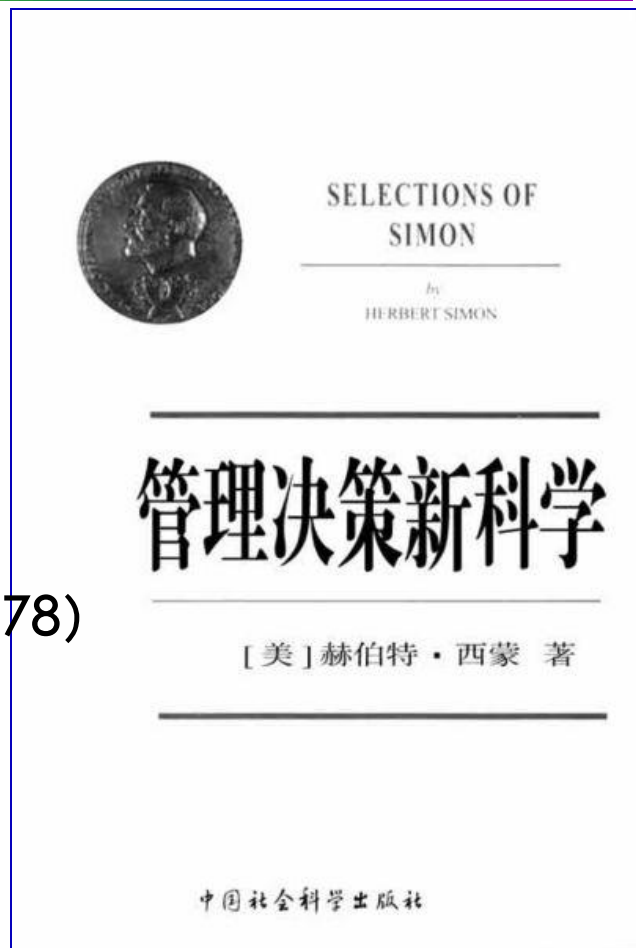


决策理论开拓者

提出了行为主义行政学理论，特别是他提出的组织决策理论，奠定了决策在行政学研究中重要地位。受其影响，在管理学领域形成了一个新的管理学派——决策理论学派。正是由于他“对经济组织内的决策程序所进行开创性研究”，获得了1978年度诺贝尔经济学奖。

管理决策著作

《管理行为》（1947，1957，1976）
《公共管理》（与斯密斯伯格合作，1950）
《组织理论的比较》（经济研究评论，1952）
《理性抉择的行为模型》经济学季刊，1955）
《理性抉择与环境结构》心理学评论，1956）
《组织》（与马奇合著，1958）
《经济学与行为科学中的决策模型》（1959）
《管理决策新科学》（1960，1975，1977）
《求解难题过程中的事物搜索》（1962）
《论如何决定做什么》（贝尔经济学杂志，1978）
《思维模型》（1979）
《有限理性模型》（1982）
《我的生活模型》（1991）
《人工智能科学》（第3版，1996）
《行政行为—行政决策研究》



司马贺决策理论

一、决策贯穿管理的全过程，决策是管理的核心。

组织中经理人员的重要职能就是作决策。他认为，任何作业开始之前都要先做决策，制定计划就是决策，组织、领导和控制也都离不开决策。

二、系统阐述了决策原理。

提出决策过程包括4个阶段：搜集情况阶段；拟定计划阶段；选定计划阶段；评价计划阶段。这四个阶段中的每一个阶段本身就是一个复杂的决策过程。

司马贺决策理论

三、用“令人满意”的准则代替“最优化”准则。

以往的管理学家往往把人看成是以“绝对的理性”为指导，按最优化准则行动的理性人。司马贺认为事实上这是做不到的，应该用“管理人”假设代替“理性人”假设，“管理人”不考虑一切可能的复杂情况，只考虑与问题有关的情况，采用“令人满意”的决策准则，从而可以做出令人满意的决策。

四、决策可分为程序化决策和非程序决策。

经常性的活动的决策应程序化以降低决策过程的成本，只有非经常性的活动，才需要进行非程序化的决策。

决策的四个阶段

收集情报

① 找出制定决策的根据，即收集情报

拟定计划

② 找到可能的行动方案

选定计划

③ 根据当时的情况和对未来发展的预测，从各个备选方案中选定一个方案

评价计划

④ 对已选择的方案及其实施进行评价

有限理性理论

- 理性的含义：

指非感情的一种计算、思考 的心智方法。

- 有限理性理论

“管理人”的价值取向和目标往往是多元的，不仅受到多方面因素的制约，而且处于变动之中乃至彼此矛盾状态；

“管理人”的知识、信息、经验和能力都是有限的，他不可能也不企望达到绝对的最优解，而只以找到满意解为满足。在实际决策中，“有限理性”表现为：决策者无法寻找到全部备选方案，也无法完全预测全部备选方案的后果，还不具有一套明确的、完全一致的偏好体系，以使它能在多种多样的决策环境中选择最优的决策方案。

满意型决策

- 所谓满意，是指决策只需要满足两个条件：

- ✓有相应的最低满意标准

- ✓策略选择能够超过最低满意标准

人工智能处理问题时，经常通过采用启发式搜索，找到问题的满意解即可。

- 满意原则的优点：它在满足要求的情况下，极大地减少搜寻成本、计算成本，简化了决策程序。因此，满意标准是绝大多数的世纪决策所遵循的基本原则。

内容提要

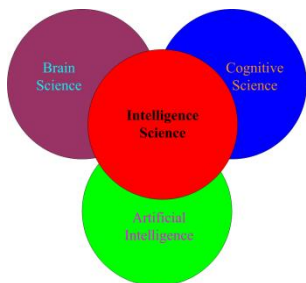
概述

符号主义创始人

决策理论开拓者

逻辑理论家发明人

科学知识发现倡导者



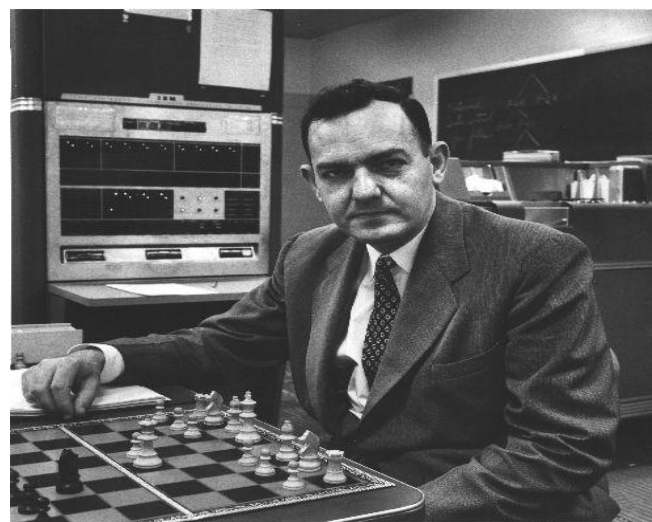
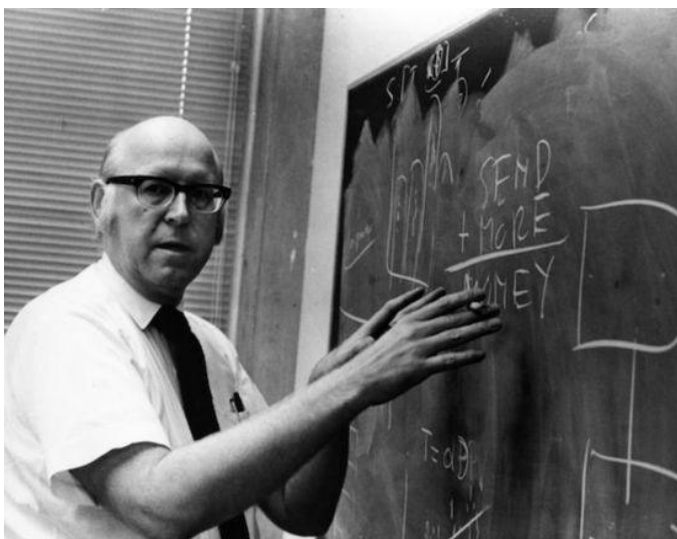
达特茅斯会议

1956年，美国达特茅斯大学助教约翰·麦卡锡、哈佛大学马文·明斯基、贝尔实验室香农、IBM信息研究中心内森·罗切斯特一起发起了著名的“达特茅斯会议”，正式确定了人工智能的概念，并因此使之成为一个崭新的学科。AI一词成了人工智能的名称，麦卡锡也因作为第一个人工智能的系统语言的发明者而被尊称为“人工智能之父”。



逻辑理论家

- 1956年人工智能研究一个重大的突破，是司马贺和纽厄尔合作编制的《逻辑理论家(Logic Theorist)》，即数学定理证明程序。



逻辑理论家

- 《逻辑理论家(Logic Theorist)》程序, 可以证明怀特海和罗素的数学名著《数学原理》第二章中的38个定理, 即命题逻辑部分的一个很大子集。1963年, 经过改进的LT程序在一部更大的电脑上, 最终完成了第二章全部52条数学定理的证明。

逻辑理论家

- 纽厄尔和司马贺认识到人类证明数学定理也有类似的思维规律，通过“分解”（把一个复杂问题分解为几个简单的子问题）和“代入”（利用已知常量代入未知的变量）等方法，用已知的定理、公理或解题规则进行试探性推理，直到所有的子问题最终都变成已知的定理或公理，从而解决整个问题。

数学机械化



笛卡尔



莱布尼茨



希尔伯特



哥德尔

- 数学问题的机械化，就要求在运算或证明过程中，每前进一步之后，都有一个确定的、必须选择的下一步，这样沿着一条有规律的、刻板的道路，一直达到结论。

数学机械化



塔斯基



司马贺



王浩



吴文俊

使用一种机械化方法证明一类定理，才真正体现了机械化定理证明。

自动定理证明

- 美籍华人学者、洛克菲勒大学教授王浩在“自动定理证明”上获得了更大的成就。1959年，王浩用他首创的“王氏算法”，在一台速度不高的IBM704电脑上再次向《数学原理》发起挑战。不到9分钟，王浩的机器把这本数学史上视为里程碑的著作中全部（350条以上）的定理，统统证明了一遍。

吴方法

- 1977 年，吴文俊在《中国科学》上发表论文《初等几何判定问题与机械化问题》。
- 1984 年，吴文俊的学术专著《几何定理机器证明的基本原理》由科学出版社出版。
- 1985 年，吴文俊的论文《关于代数方程组的零点》发表，具体讨论了多项式方程组所确定的零点集。



通用问题求解

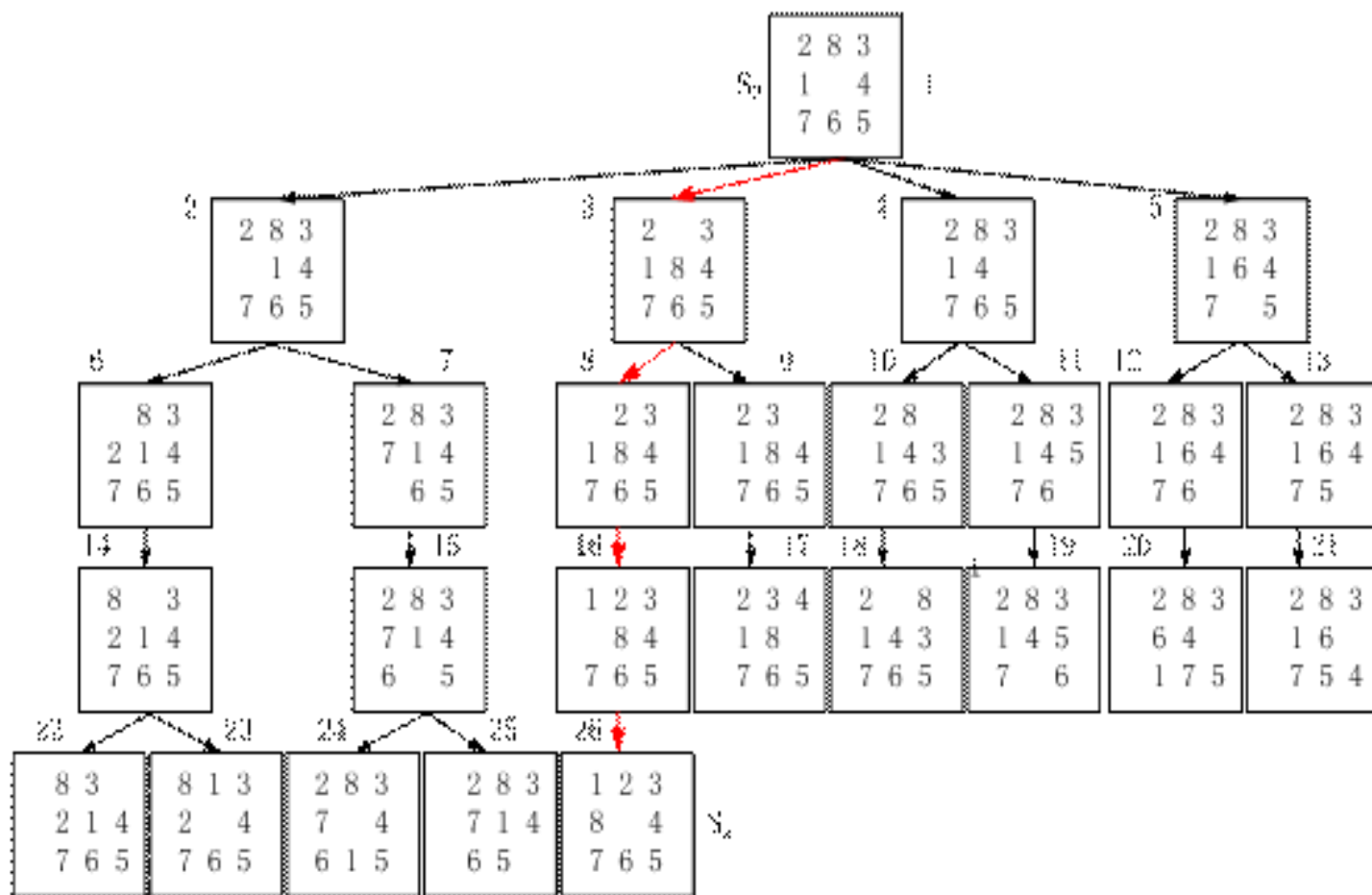
从 1957 年 开 始 ，
Newell ， Shaw 和
Simon 等 人 就 开 始 研
究 一 种 不 依 赖 于 具 体
领 域 的 通 用 问 题 求 解
程 序 ， 称 为 GPS
(General Problem
Solver)。



通用问题求解

- 基于这一思想，他们于1960年编制了能解10种类型不同课题的通用问题求解程序GPS。通用问题求解研究前后持续了10年，最后的版本发表于1969年。
- 通过通用问题求解程序的研究，司马贺等发展了手段-目的的分析方法，提出MEA算法，得到广泛应用。
- 司马贺对决策过程中合理选择和环境影响行为理论、启发式求解方法进行了探讨。

启发式搜索



手段一目的分析法（MEA）

首先确定目标，然后考虑要达到这一目标，先要实现什么子目标，再考虑要达到该子目标，又必须先实现什么子目标，由此类推下去。

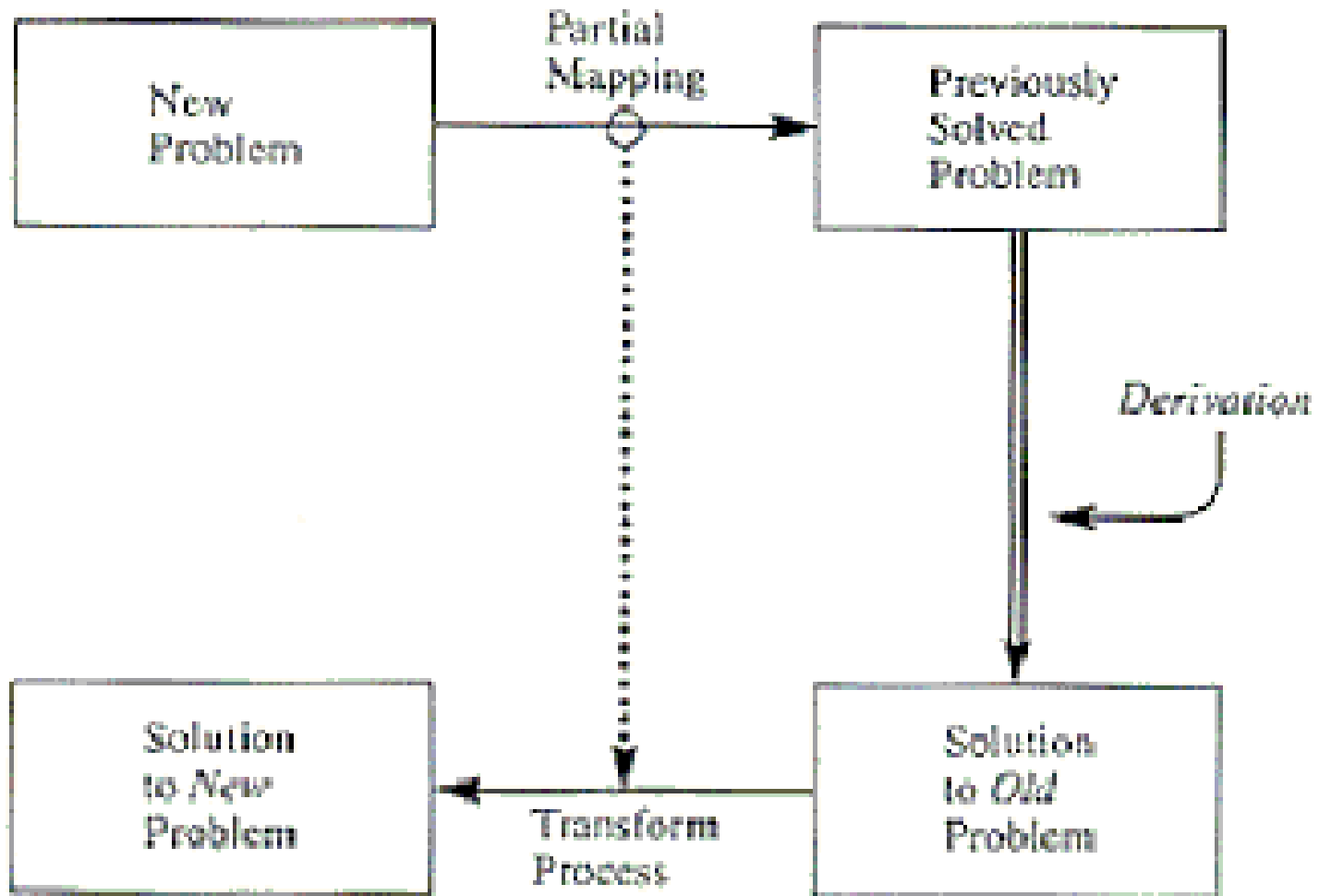
第一步：明确解决问题的目标

第二步：确定妨碍目标的障碍

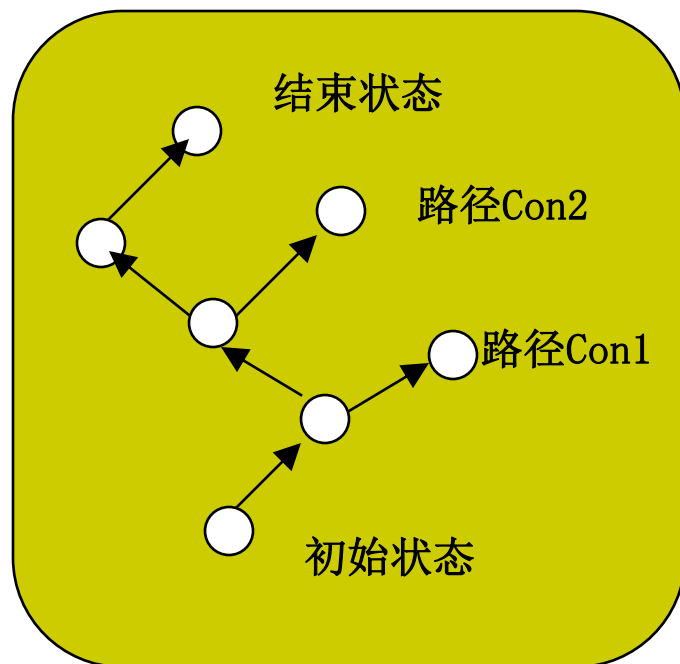
第三步：分析造成障碍的原因

第四步：对症下药，排除障碍

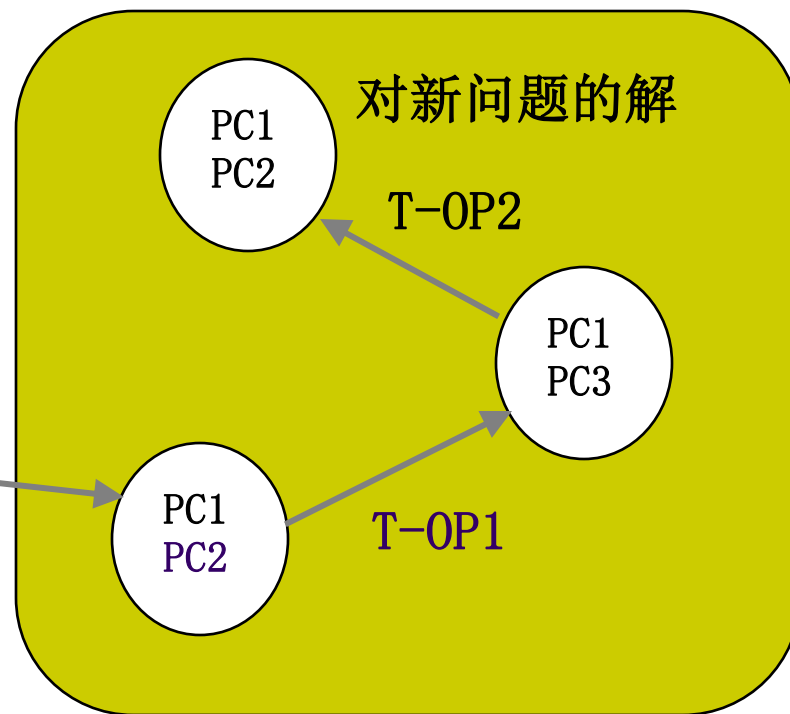
转换类比



EMEA类比问题求解



原空间



T-空间

常用T-操作符

- 1) 通用插入
- 2) 通用删除
- 3) 子序列拼接
- 4) 子目标保持替换
- 5) 终结段连结
- 6) 初始段连结
- 7) 序列合并
- 8) 操作符记录
- 9) 用参数替换
- 10) 解序列截断
- 11) 序列倒置

差别表

入口形式:

“为简少<差别>, 应采用<T-操作符集>中的一个操作”

T-空间的差别测度 D_r （差异函数）

D_r 的值是四维向量：

- 新旧问题初态的差别
- 新旧问题终态的差别
- 新旧问题路径限制的差别
- 新旧问题方法可应用度的差别。

内容提要

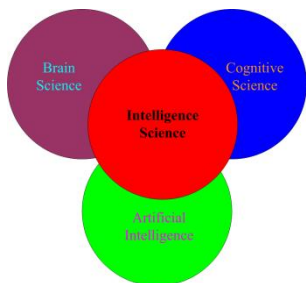
概述

符号主义创始人

决策理论开拓者

逻辑理论家发明人

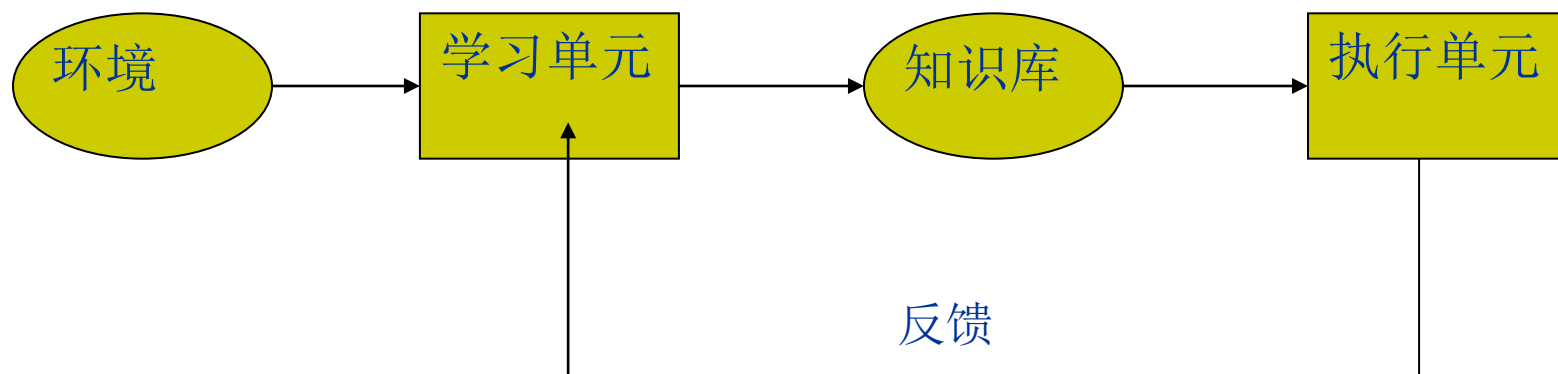
科学知识发现倡导者



学习模型

司马贺对学习的定义是：“如果一个系统能够通过执行某种过程而改进它的性能，这就是学习”。

- 学习是一个过程，
- 学习是对一个系统而言，
- 学习改变系统性能。

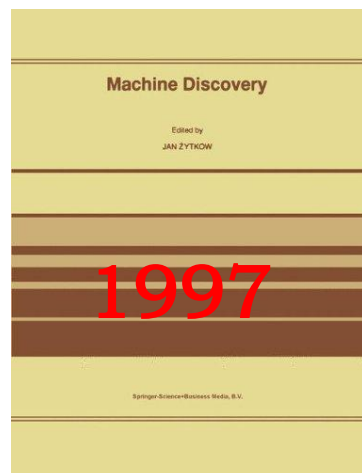
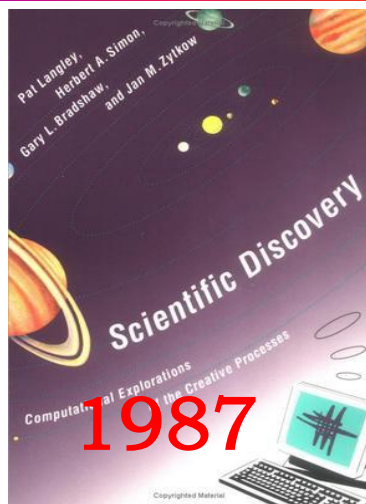


科学知识发现

司马贺倡导科学知识发现，研制的机器发现程序有

- BACON系列系统：从数据中推导数学定律
- GLAUBER系统：化学反应的定性关系
- STAHL系统：结论能够解释一些物质的结构
- DALTON系统：从假定出发，结合相关化学反应的输入资料进行推导
- KEKADA系统：德国生物化学家克雷布斯发现尿素循环的活动

BACON系统



1976年到1983年，司马贺指导兰利（Langley）、席特考（Zytkow）、布拉德肖（Bradshaw）研制了6个版本的BACON系统。它们的共同特点：采用数据驱动，通过启发式约束搜索，依赖理论数据项，递归应用一些通用的知识发现方法。

BACON系列

BACON系统版本	主要启发式方法
BACON.1	探测趋势和常数
BACON.2	寻找常数差异的特殊方法
BACON.3	探测趋势和常数 更高层次描述的递归
BACON.4	探测趋势和常数 更高层次描述的递归 固有属性方法 公约数方法
BACON.5	寻找常数差异的普通方法 固有属性方法 公约数方法 基于期望的方法
BACON.6	处理噪音的爬山法 更高层次描述的递归 固有属性方法 公约数方法 基于期望的方法

BACON系列

BACON系统版本	物理定律名称	获得的定律形式
BACON.1	波义耳定律	$PV=c$
	开普勒第三定律	$D^3P^2=k$
	欧姆定律	$IL=-rl+v$
	伽利略定律	$D/T^2=g$
BACON.3	理想气体定律	$PV= aNT + bN$
	开普勒第三定律	$D^3 [(A-k)/t]^2=j$
	伽利略定律	$dP^2/Lt^2=k_5$
	库仑定律	$FD^2/Q_1Q_2=c$
	欧姆定律	$TD^2/(LI-rl)=v$
BACON.4	欧姆定律	$I=v/r$
	阿基米德定律	$C=V+i$
	斯内耳折射定律	$\sin i/\sin r= n_1/n_2$
	动量守恒定律	$m_1V_1=m_2V_2$
	引力定律	$F=Gm_1m_2/d^2$
	布莱克比热定律	$c_1m_1T_1+c_2m_2T_2=(c_1m_1+c_2m_2)T_f$
BACON.5	斯内耳折射定律	$n_1\sin\theta_1/n_2\sin\theta_2=1.0$
	动量守恒定律	$(U_1-aV_1)/k_1=-(U_2-aV_2)/k_2$
	布莱克比热定律	$c_1M_1(F_1-aT_1)=-c_2M_2(F_2-$

KEKADA系统

KEKADA系统以考察实验过程的设计，解释观察数据产生的科学发现活动为目的的设计。该系统模拟的是**1953**年诺贝尔医学与生理学奖奖金获得者德国生物化学家克雷布斯发现尿素循环的活动，以克雷布斯的名字命名。该系统是司马贺团队研制的系列程序中最贴近发现的历史过程的程序。**“KEKADA”**这个单词在印度语中是**“Krebs”**的同义词。

KEKADA系统

设计者将从资料中获取的数据划分为产生式系统中工作记忆的**5类要素**：过程，物质，实验，附加事实和假设。过程要素描述了化学反应中的输入、输出等属性。物质要素提供例如酸或碱等物质的属性。实验要素包括了实验的输入、执行条件等信息。附加事实包括了对发生的过程确认的量度。假设是对现象或过程可能如何发生的描述，与确认相关。编入系统中的这些要素，可以视为提供给系统的，与克雷布斯当时所具有的相似的背景知识

KEKADA系统

KEKADA系统有**9**组一共**64**个启发式。这些启发式与产生式的形式是完全相同的，即采用“如果……，那么……”的形式，由**OPS₅**（**Official Production System**）产生式系统语言表达。这**9**组启发式分别是：

- | | |
|--------------|-----------|
| 1) 问题选择者； | 2) 问题产生者； |
| 3) 决策制订者； | 4) 实验提议者； |
| 5) 预期设置者； | 6) 假设产生者； |
| 7) 假设修改者； | 8) 信心修改者； |
| 9) 假设/策略选择者。 | |

科学知识发现

1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Bacon.1–Bacon.5						Abacus, Coper		Fahreheit, E*, Tetrad, IDS_N			Hume, ARC		DST, GP _N LaGrange			SDS		SSF, RF5, LaGrange			
←AM			Glauber		NGlauber			IDS _Q , Live							RL, Progol		HR				
←Dendral			Dalton, Stahl		Stahlp, Revolver		Gell-Mann		BR-3, Mendel		Pauli		BR-4								
						IE			Coast, Phineas, AbE, Kekada				Mechem, CDP					Astra, GP _M			

Legend

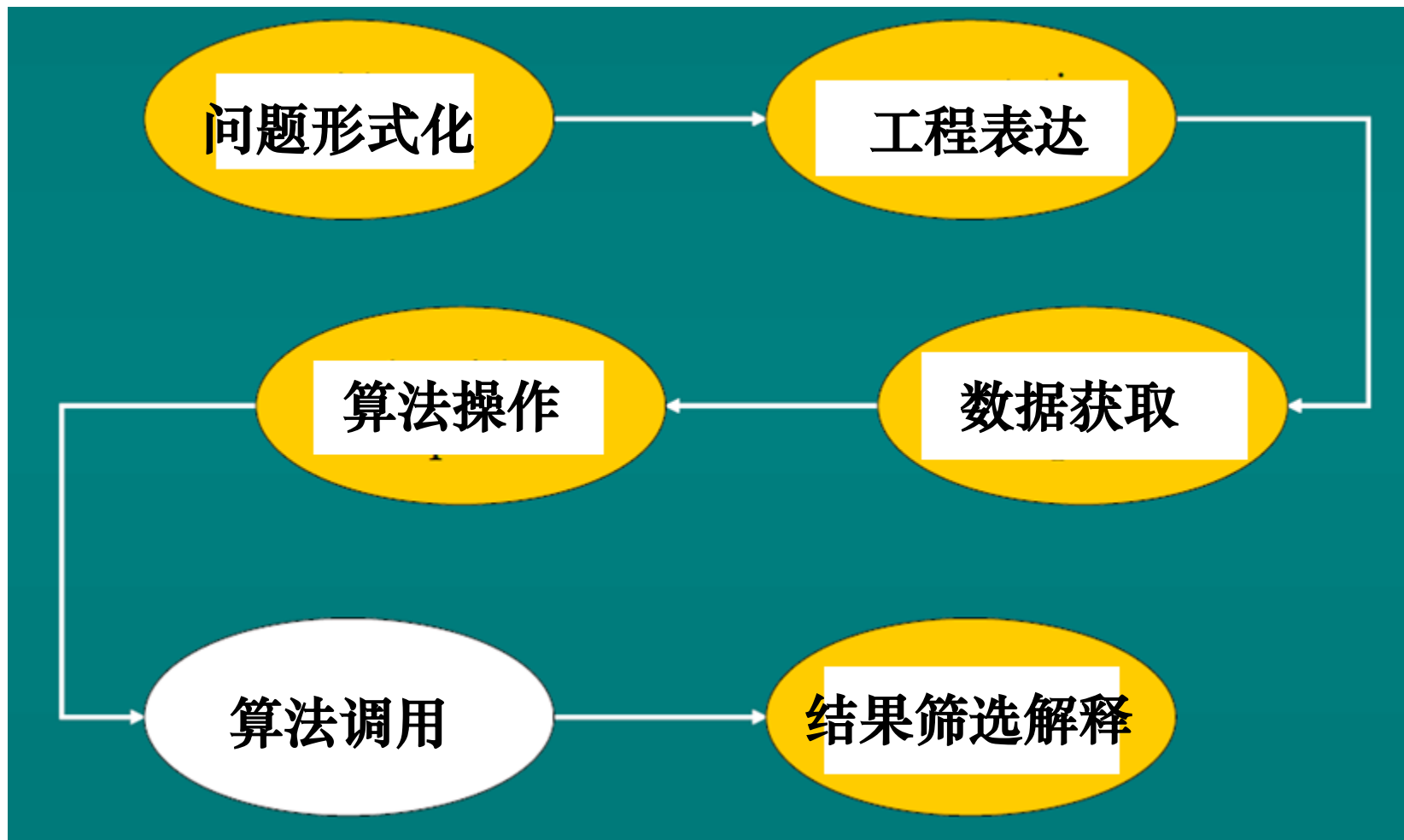
Numeric laws

Qualitative laws

Structural models

Process models

科学知识发现步骤



科学知识发现应用

气候变化

气候变化和
碳排放



地球生态系统
模型和解释

生物信息学

基因表达和
异形组织

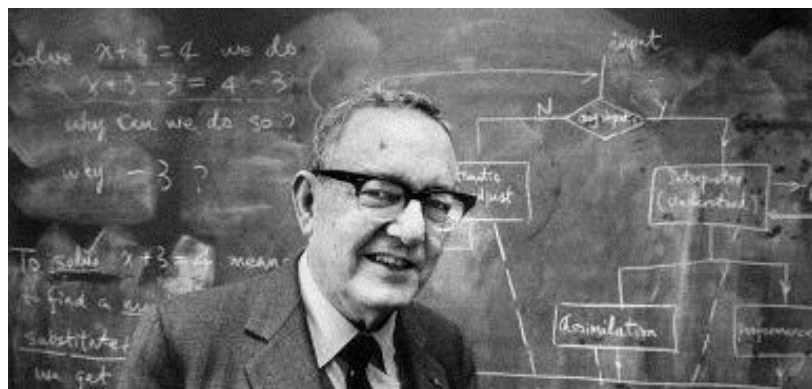


基因调控
模型和解释

大数据



对中国的深情厚谊



"Engineering, medicine, business, architecture and painting are concerned not with the necessary but with the contingent - not with how things are but with how they might be - in short, with design."

Herbert Simon

- 1980年起，中美学术交流委员会成员，并于1983年至1987年担任美方主席。
- 1985年，中国科学院心理研究所授予西蒙教授名誉研究员称号
- 1994年，他获得中国科学院外籍院士称号。

对中国的深情厚谊



- 1988年在东京召开FGCS国际会议，司马贺应邀在大会上作主题报告，指出与中国合作的示例演练学习“对认知心理学和学习理论做出了重要贡献”

司马贺之路

Simon road in CMU



致 谢

智能科学<http://www.intsci.ac.cn>

