

第22次统计学论坛

人工智能的不确定性研究

史忠植

shizz@ict.ac.cn

中国科学院计算技术研究所

<http://www.intsci.ac.cn/>



中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY

内容提要



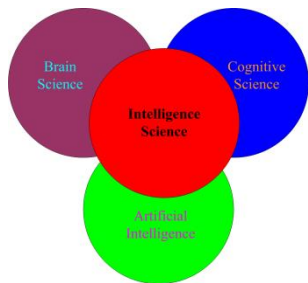
引言

不确定性推理

统计学习

大数据分析

展望



达特茅斯会议



- 1956 SUMMER, 美国达特茅斯学院, 为期2月

J. McCarthy, M. Minsky, N. Lochester (IBM), C. E. Shannon

A. Samual (IBM), H. A. Simon (CMU), A. Newell (CMU),

T. More (Princeton), R. Solomonoff (MIT), O. Selfridge (MIT)

- 洛克菲勒基金会资助, 每人1200美元, 报销往返车票

- 目标: 10个人2个月设计出具有真正智能的计算机

- 结果: 奠定了新的科学- 人工智能

什么是人工智能

- 人工智能 (Artificial Intelligence) ，英文缩写为 AI。它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。
- 它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器，该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。

人工智能

- Stuart Russell和Peter Norvig:

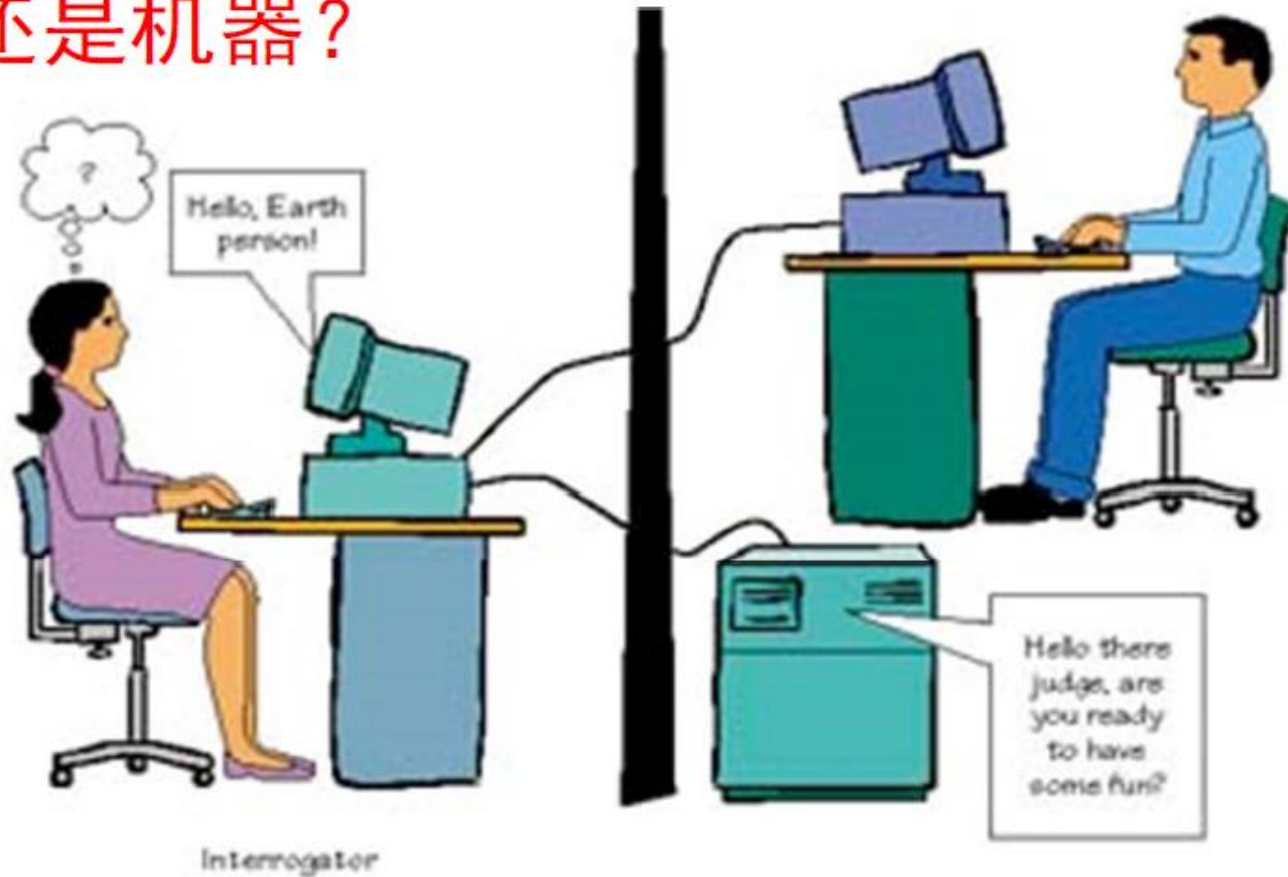
人工智能——一种现代方法(第三版)

把当前有关AI的定义分成四类：

类人思维方法	理性思维系统
类人行为方法	理性行为系统

图灵测试

是人？还是机器？



新一代人工智能

2017年7月8日，国务院印发了新一代人工智能发展规划的通知，指出

- 人工智能成为国际竞争的新焦点。人工智能是引领未来的战略性技术，世界主要发达国家把发展人工智能作为提升国家竞争力、维护国家安全的重大战略。
- 人工智能成为经济发展的新引擎。人工智能作为新一轮产业变革的核心驱动力，将进一步释放历次科技革命和产业变革积蓄的巨大能量，并创造新的强大引擎。

人工智能将引领人类第四次工业革命 - 智能化

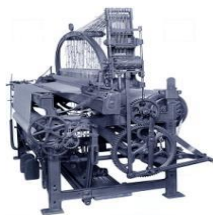
互联网时代

终正
结在

- 人工智能
- 机器人
- 交通工具 (即无人机、无人驾驶等)
- VR (虚拟现实)

来悄
临悄

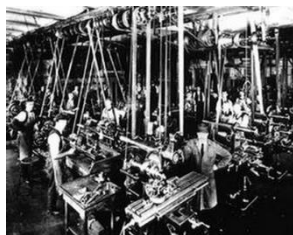
工业1.0
创造了机器工厂的
“蒸汽时代”



蒸汽机

18世纪末

工业2.0
将人类带入分工明确、大批量生产的流水线模式和“电气时代”



电力广泛应用

20世纪初

工业3.0
应用电子信息技术，进一步提高生产自动化水平



自动化、信息化

1970年代初

工业4.0
开始应用信息物理融合系统 (CPS)



信息物联系统

今天

复杂度

新一代人工智能发展的战略目标

到2020年

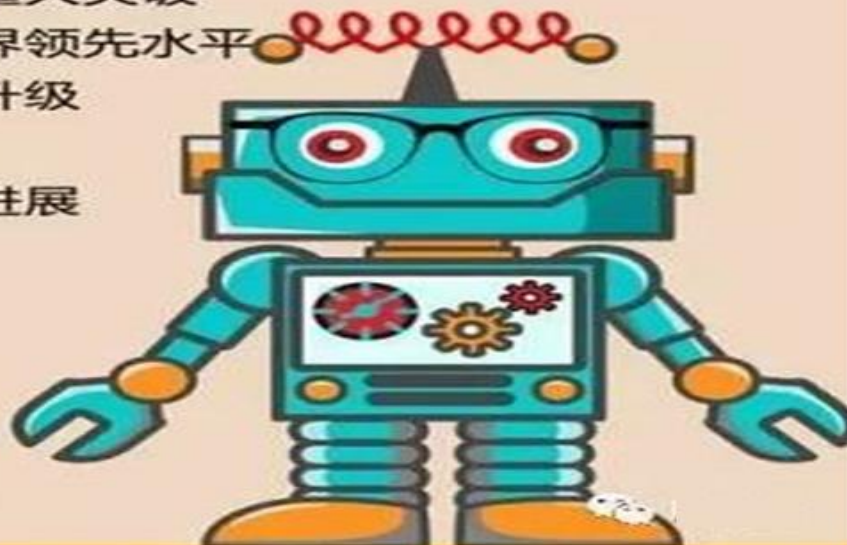
- ▶ 人工智能总体技术和应用与世界先进水平同步
- ▶ 人工智能产业成为新的重要经济增长点
- ▶ 人工智能技术应用成为改善民生的新途径

到2025年

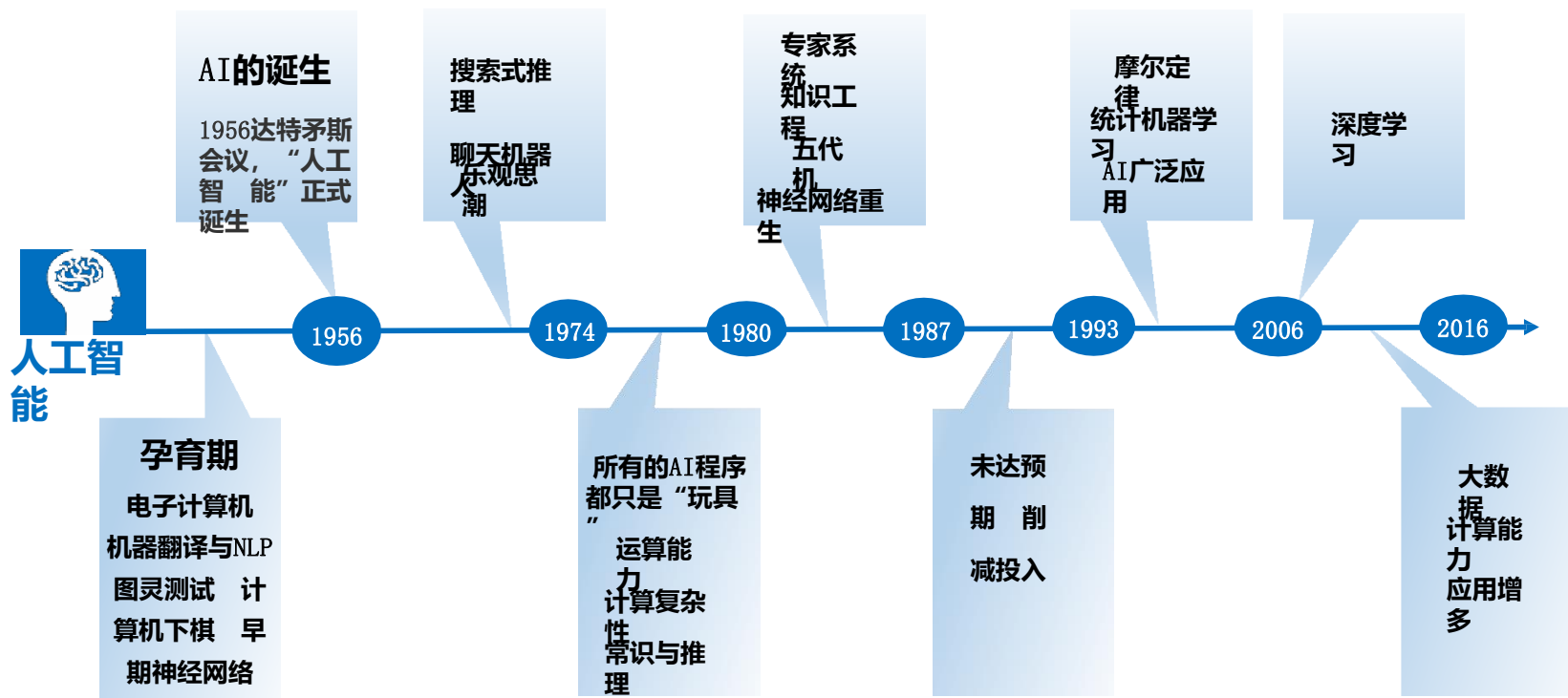
- ▶ 人工智能基础理论实现重大突破
- ▶ 部分技术与应用达到世界领先水平
- ▶ 人工智能成为我国产业升级和经济转型的主要动力
- ▶ 智能社会建设取得积极进展

到2030年

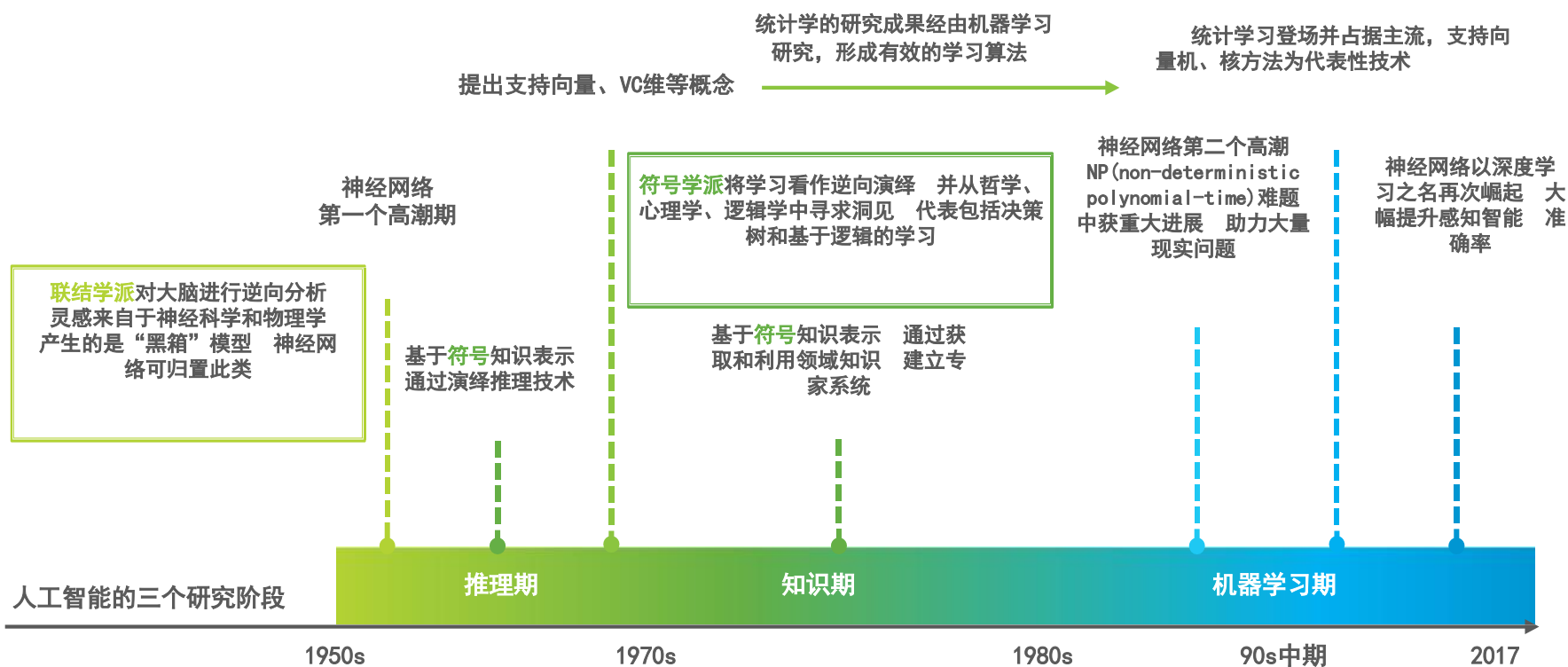
- ▶ 人工智能理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心



人工智能发展历程



人工智能发展阶段

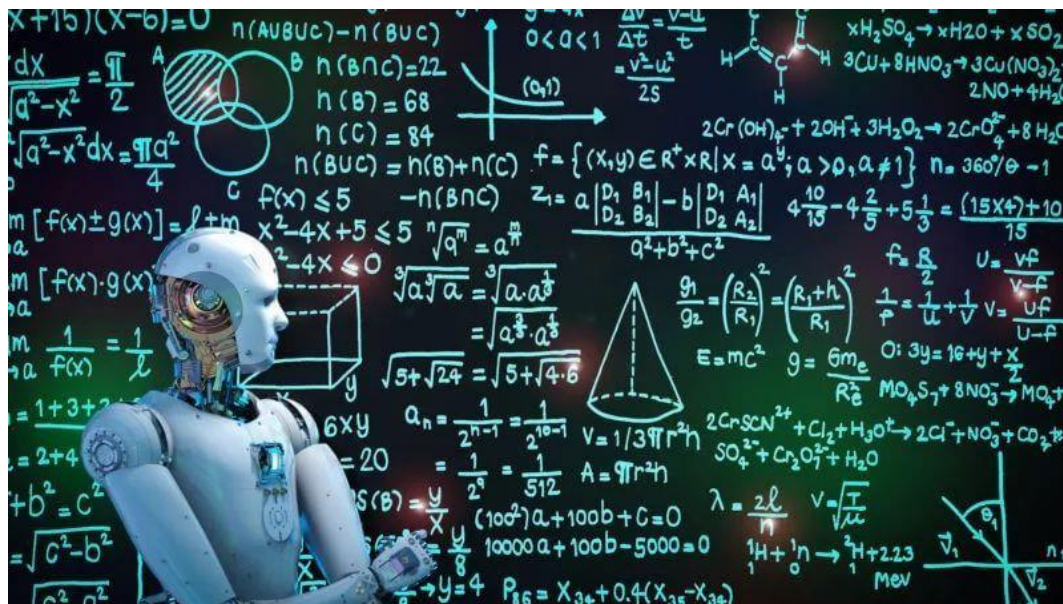


人工智能的几起几落

- **第一次浪潮(1956-1973):** 以注重逻辑推理为标志的问题求解和语言处理时代
- **第二次浪潮(1980-1992):** 依托知识积累构建模型为标志的专家系统时代
- **第三次浪潮(2006→):** 以重视数据、深度学习为标志的认知智能时代。大数据、物联网、云计算等技术为人工智能的发展打下了良好基础。

几起几落的教训

重视基础理论研究!



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。



国家最高科学技术奖



2001年2月19日上午，中共中央总书记、国家主席、中央军委主席江泽民向获得2000年度国家最高科学技术奖的中国科学院系统科学研究所研究员、中国科学院院士吴文俊和湖南杂交水稻研究中心研究员、中国工程院院士袁隆平颁奖

吴文俊人工智能科学技术奖

吴文俊人工智能科学技术奖是我国智能科学技术领域唯一以享誉海内外的杰出科学家、数学大师、人工智能先驱、中国人工智能学会名誉理事长吴文俊先生命名，依托社会力量设立的科学技术奖，具备提名推荐国家科学技术奖资格，被誉为“中国智能科学技术最高奖”。



内容提要



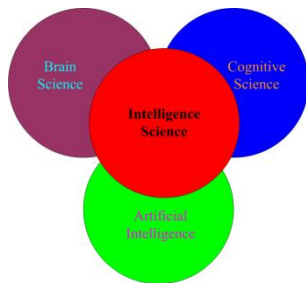
引言

不确定性推理

统计学习

大数据分析

展望



不确定性推理

- 什么是不确定性推理？
 - 不确定性推理是建立在非经典逻辑上的一种推理, 是对不确定性知识的运用与处理
 - 是从不确定性的初始证据出发, 通过运用不确定性的知识, 最终推出具有一定程度的不确定性但却合理或者近乎合理的结论的思维过程
- 为什么要研究不确定性推理？
 - 日常生活中含有大量的不确定的信息
 - ES系统中大量的领域知识和专家经验, 不可避免的包含各种不确定性。

不确定性推理

- 不确定性推理的基本问题：
 - 表示问题：即采用什么方法描述不确定性. 一般有数值表示和非数值的语义表示方法.
 - 计算问题：主要指不确定性的传播和更新, 也即获得新信息的过程. 主要包括：
 - 已知 $C(A)$, $A \rightarrow B$ $f(B, A)$, 如何计算 $C(B)$
 - 已知 $C_1(A)$, 又得到 $C_2(A)$, 如何确定 $C(A)$
 - 如何由 $C(A_1)$, $C(A_2)$ 计算 $C(A_1 \wedge A_2)$, $C(A_1 \vee A_2)$
 - 语义问题：指的是上述表示和计算的含义是什么, 如何进行解释.

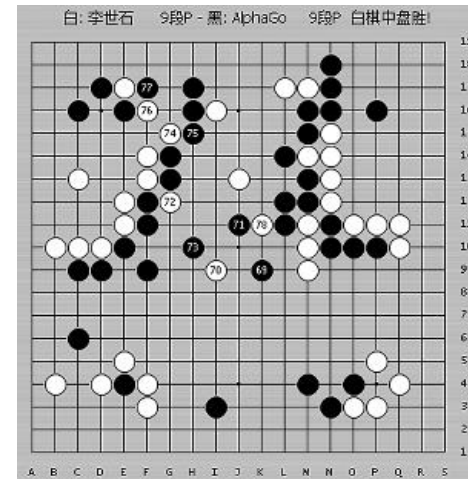
不确定性推理

不确定推理方法的分类

- 形式化方法:在推理一级扩展确定性方法.
 - 逻辑方法:是非数值方法,采用多值逻辑、非单调逻辑来处理不确定性
 - 新计算方法:认为概率方法不足以描述不确定性,出现了确定性理论,确定性因子,模糊逻辑方法等
 - 新概率方法:在传统的概率框架内,采用新的计算工具以确定不确定性描述
- 非形式化方法:在控制一级上处理不确定性
 - 如制导回溯、启发式搜索等等

阿尔法围棋 (AlphaGo)

阿尔法围棋 (AlphaGo) 是围棋人工智能程序，由谷歌旗下DeepMind公司的戴密斯·哈萨比斯、大卫·席尔瓦、黄士杰等人的团队开发。其主要工作原理是“深度学习”。2016年3月9-15日，谷歌人工智能“阿尔法围棋”与韩国棋手李世石比赛，结果4:1获胜



阿尔法围棋 (AlphaGo)

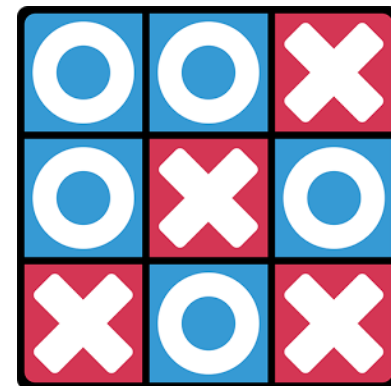
- 2017年5月23-27日，中国棋手柯洁九段和计算机围棋程序阿尔法围棋人机大战在浙江省桐乡市乌镇举行，阿尔法围棋以 3 : 0 完胜。



围棋搜索空间

■ Tic-Tac-Toe (3x3)

- 只考虑状态：黑子、白子、无子
 - 3^9 (3的9次方, 即9个3相乘) = 19,683
- 考虑落子的顺序
 - $9! = 362,880$

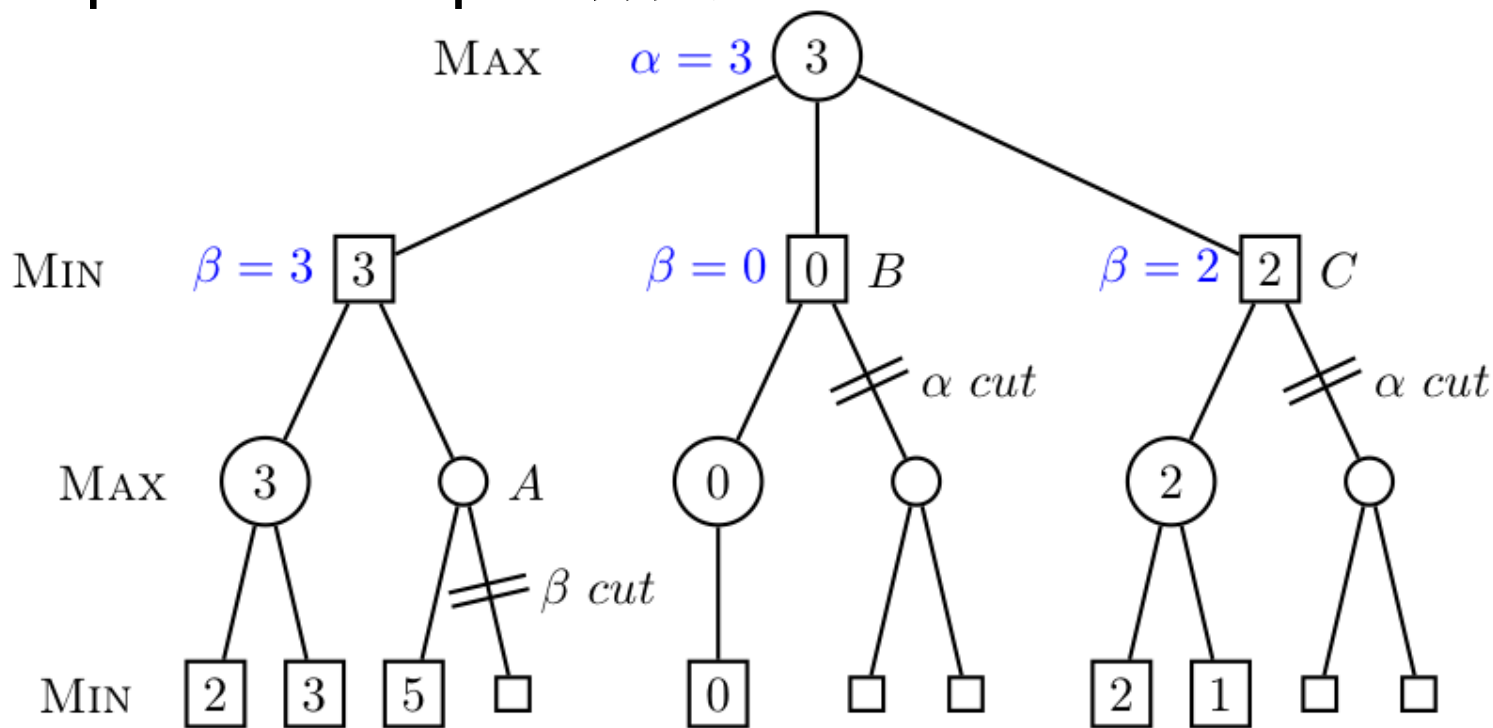


■ 围棋

- 只考虑状态： 3^{361} (约是 10^{170}) 种 ($19 \times 19 = 361$)
- 考虑行棋的顺序： $361!$ (约是 10^{768})
- 仅仅推算50步： $> 10^{120}$ 种
- 只推5步： $> 2 \times 10^{12}$ 种

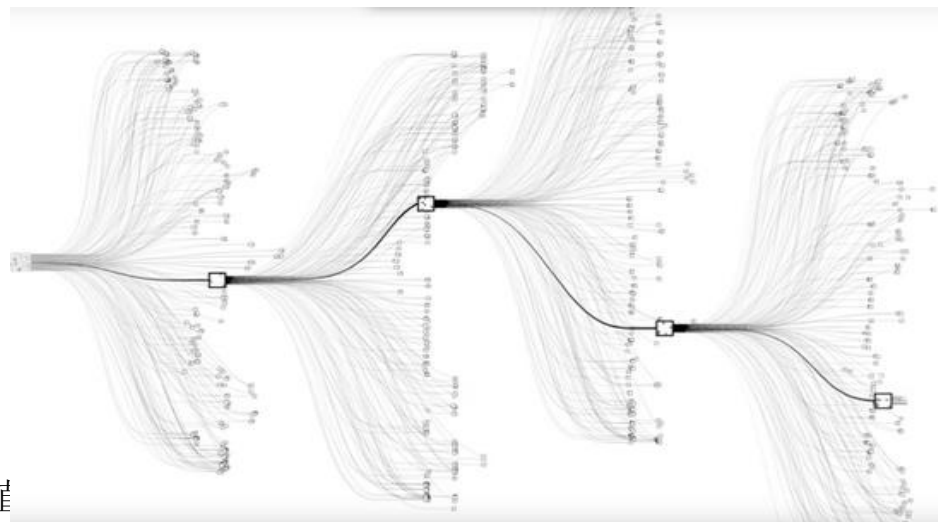
缩小搜索空间

- Alpha-Beta剪枝
- AlphaGo的Alpha含义



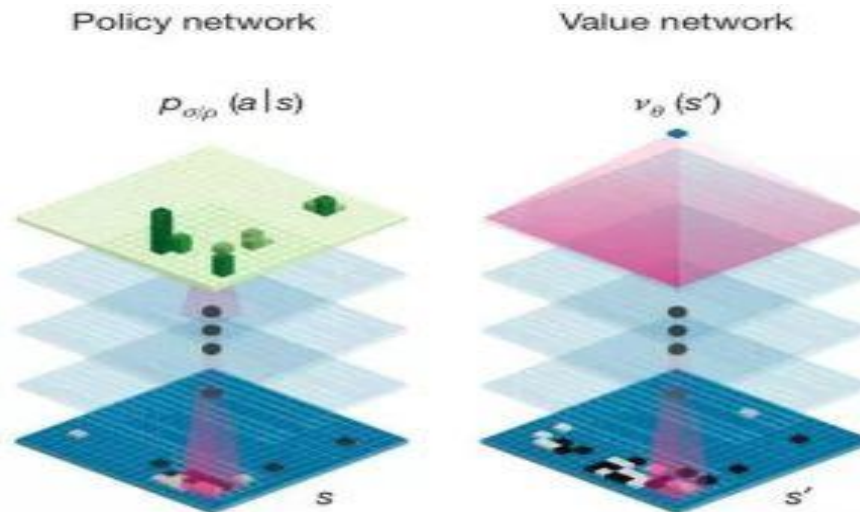
阿尔法围棋 (AlphaGo)

- 而AlphaGo以一套名为“蒙特卡罗”的运算技术为基础，搭配类神经网络，从世界棋手、人类专家的技术细节，进行像是象棋、围棋等棋局对弈的研究，然后通过两个阶段的机器学习，第一段是人工喂资料给电脑去运算，第二阶段是给予电脑在比赛过程中去学习的程序，不管是赢还是输，每次比赛过程都会把经验学进去。



阿尔法围棋 (AlphaGo)

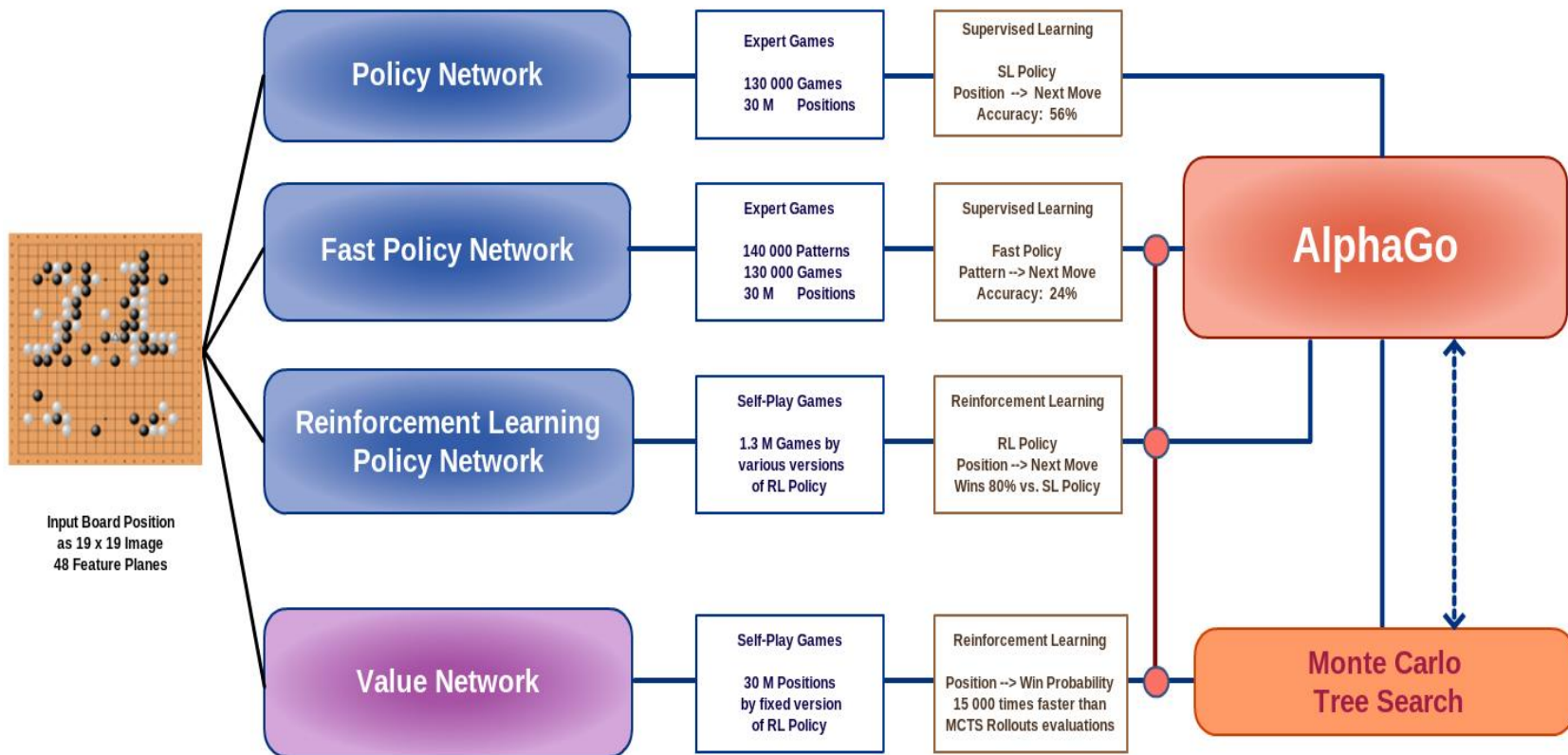
- 第二阶段采用的技术，叫做深度神经网络，主要是通过云端运算的方式，进行深度学习，然后进行“策略网络 (Policy network)”与“评价网络 (Value network)”的分析，程序就像人类的大脑一般，会有自己的思考模式，这也是为什么AlphaGo可以持续进化。



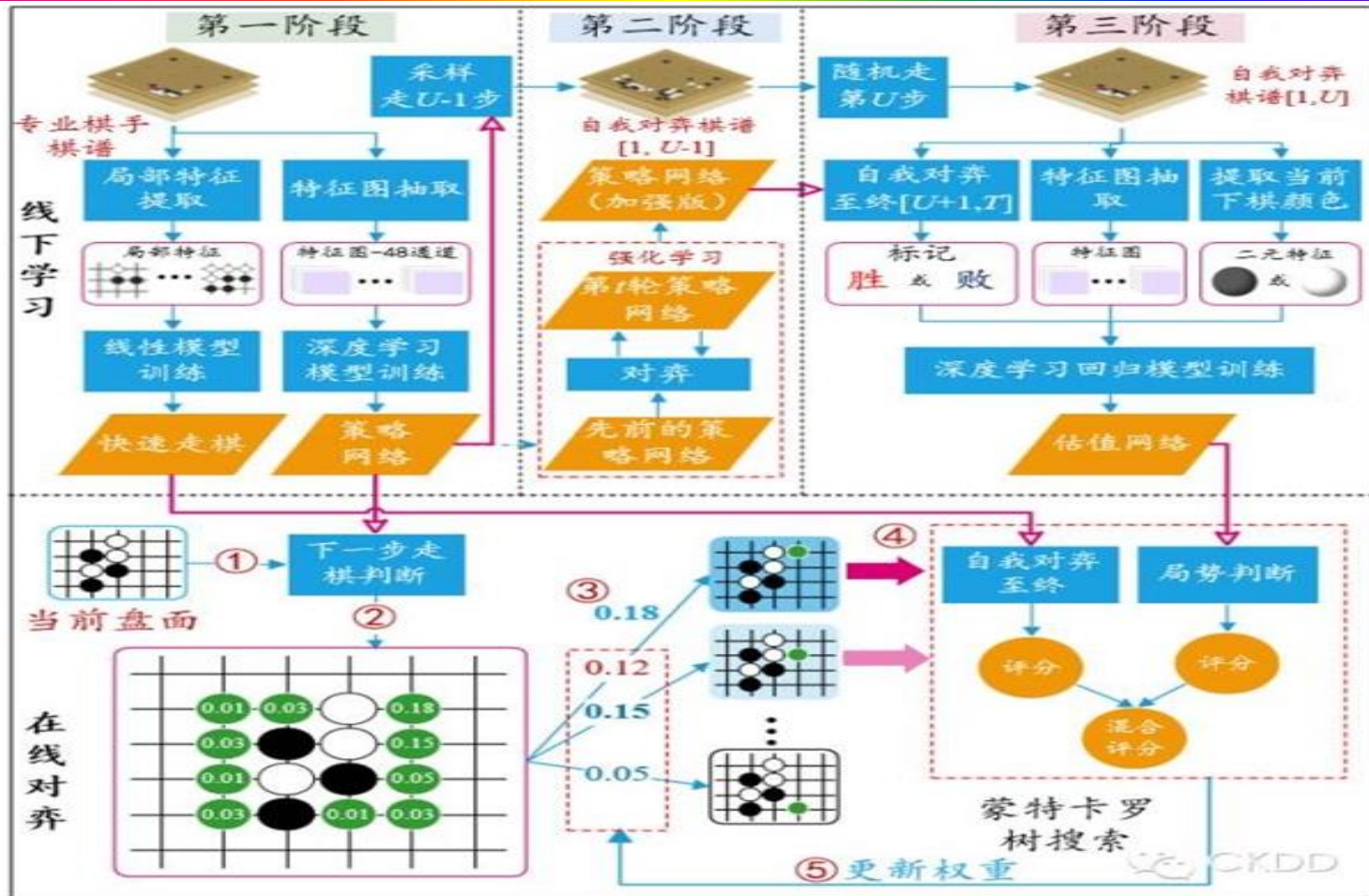
AlphaGo系统

AlphaGo Overview

based on: Silver, D. et al. Nature Vol 529, 2016
copyright: Bob van den Hoek, 2016



AlphaGo系统



蒙特卡罗方法

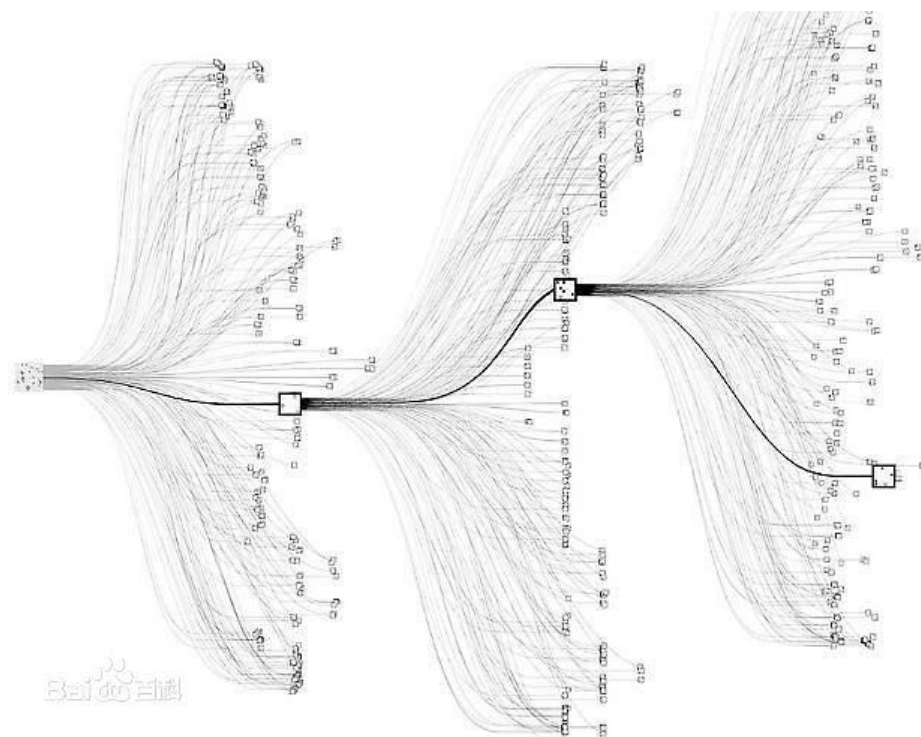
- 蒙特卡洛树搜索又称随机抽样或统计试验方法，属于计算数学的一个分支，它是在上世纪四十年代中期为了适应当时原子能事业的发展而发展起来的。传统的经验方法由于不能逼近真实的物理过程，很难得到满意的结果，而蒙特卡洛树搜索方法由于能够真实地模拟实际物理过程，故解决问题与实际非常符合，可以得到很圆满的结果。

蒙特卡罗方法

- 蒙特卡罗方法是一类随机方法的统称，是**随机抽样技术或统计试验方法**
 - **采样越多，越近似最优解**
 - 这类方法的特点是，可以在随机采样上计算得到近似结果，随着采样的增多，得到的结果是正确结果的概率逐渐加大，但在（放弃随机采样，而采用类似全采样这样的确定性方法）获得真正的结果之前，无法知道目前得到的结果是不是真正的结果。
 - **尽量找好的，但不保证是最好的**

蒙特卡罗方法

- 2006年，雷米•库洛姆（Remi Coulom）描述了蒙特卡罗方法在游戏树搜索的应用并命名为蒙特卡罗树搜索
- Multi-armed bandit Problem.



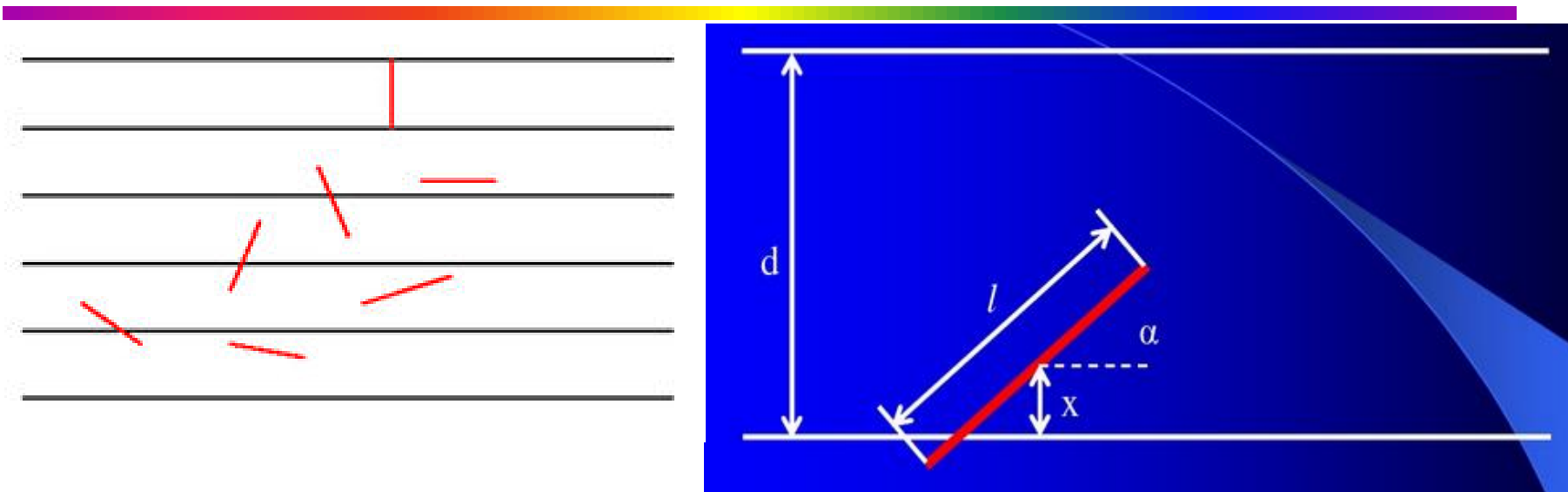
Buffon投针

- Monte Carlo方法的基本思想很早以前就被人们所发现和利用。
 - 早在17世纪，人们就知道用事件发生的“频率”来决定事件的“概率”。
 - 1777年，蒲丰（Buffon）提出著名的Buffon投针试验来近似计算圆周率 π

Buffon投针

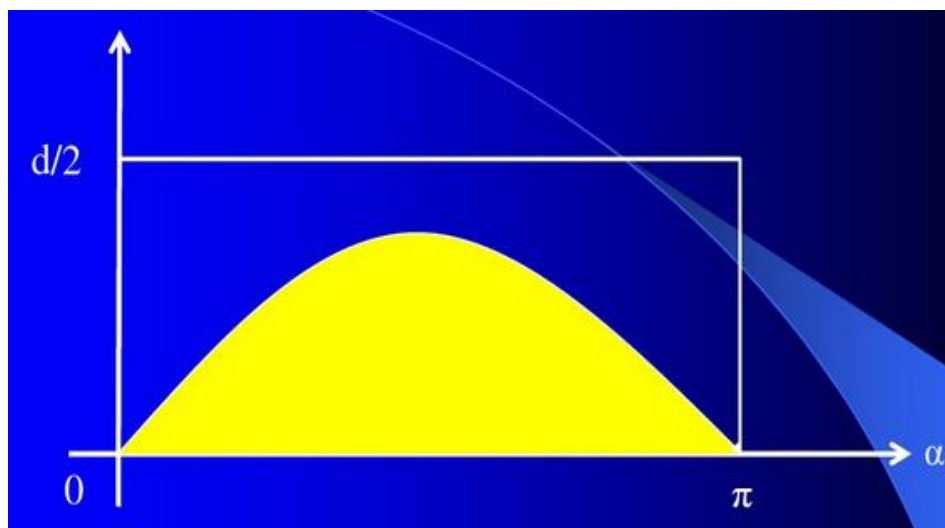
- Buffon 投针
 - 在一张白纸上，画上许多条间距为 d 的等距的平行线，另有一根长度为 l ($l < d$) 的针，随机地向纸上投掷针，并记录投掷针的次数 n ，以及针与直线相交的次数 m 。据此计算 π 值。

Buffon投针



- 随机投针可以理解为针的中心点与最近的平行线的距离 x 是均匀地分布在区间 $[0, d/2]$ 上的随机变量
- 针与平行线的夹角 α 是均匀地分布在 $[0, \pi]$ 上的随机变量，且 x 与 α 相互独立
- 于是针与平行线相交的充要条件： $x < (1/2) \sin \alpha$

Buffon投针



- 根据 $0 \leq x \leq d/2$ 和 $0 \leq \alpha \leq \pi$, 从而全概率空间为 x - α 平面上的一个矩形, 其中满足相交性的部分为 $x \leq (d/2) \sin \alpha$ 部分, 如黄颜色所示。
- 因而黄颜色部分与长方形面积之比即为针与直线相交的概率

蒙特卡罗方法

- 像投针实验一样，用通过概率实验所求的概率来估计我们感兴趣的一个量，这样的方法称为蒙特卡罗方法（Monte Carlo method）。
- 当模型含有不确定的随机因素，分析起来通常比确定性的模型困难，模型难以作定量分析，得不到解析的结果
- 或者是虽有解析结果，但计算代价太大以至不能使用
- 这时可以考虑采用 Monte Carlo 方法

蒙特卡罗评估

- 基于蒙特卡洛模拟的局势评估方法，可以与传统的静态局面评估方法相比对。
- 在计算机围棋中，可以利用掷点法完成对围棋盘面的评估。
 - 从当前局面的所有可落子点中随机选择一个落子点
 - 重复上述过程
 - 直到胜负可判断为止
 - 经过多次模拟后，选择胜率最大的点落子

蒙特卡罗规划

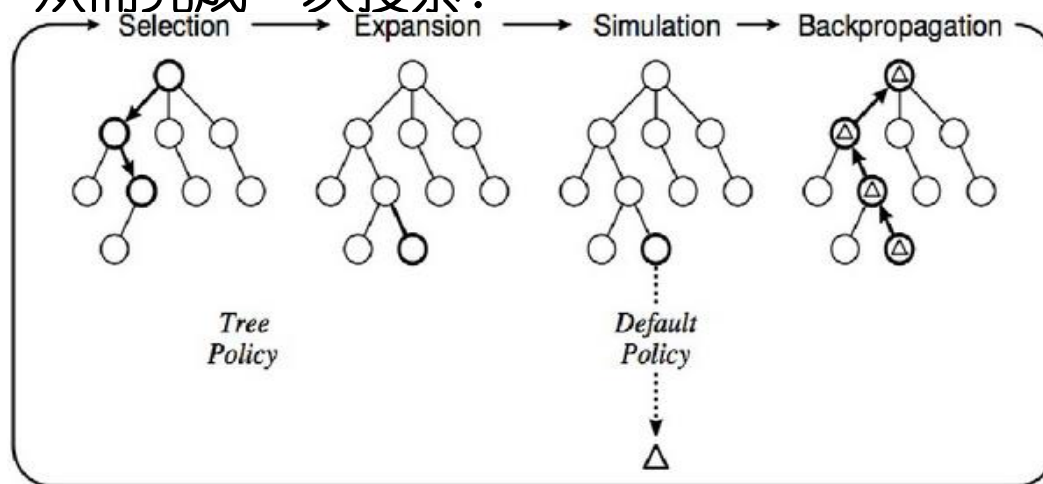
- 蒙特卡罗规划是解决马尔科夫决策问题的有效方法之一，它是以**蒙特卡罗方法**的思想为基础的一种规划方法。
 - 将可能出现的状态转移过程用**状态树**表示
 - 从初始状态开始重新抽样，逐步扩展树中的节点
 - 某个状态再次被访问时，可以利用已有的结果，提高效率
 - 在抽样过程中，可以随时得到行为的评价

蒙特卡罗规划

- 不同于普通搜索树的建立，蒙特卡罗规划算法通过从初始状态开始重复给出随机模拟事件，并逐步扩展搜索树中的每个节点，而每一次模拟事件都是“**状态-行为-回报**”的三元序列。
- 因此，相对于在评估之初就将（隐含）博弈树进行展开的静态方法而言，蒙特卡罗规划的过程是一种**动态的搜索过程**。

蒙特卡罗规划

- 更具体来讲，蒙特卡罗规划算法共包含四个基本步骤，分别为选择 (selection)、扩展 (Expansion)、模拟 (Simulation) 和回溯 (Back propagation)。
- 图中的每个节点表示博弈过程中的一个盘面状态，每一条边表示在父节点上采取一个行动，并得到子节点所对应的状态。这四个基本步骤依次执行，从而完成一次搜索：



蒙特卡罗规划

1. **选择**：从根节点出发，在搜索树上自上而下迭代式执行一个子节点选择策略，直至找到当前最为紧迫的可扩展节点为止，一个节点是可扩展的，当且仅当其所对应的状态是非停止状态，且拥有未被访问过的子状态；
2. **扩展**：根据当前可执行的行动，向选定的节点上添加一个（或多个）子节点以扩展搜索树；
3. **模拟**：根据默认策略在扩展出来的一个（或多个）子节点上执行蒙特卡罗棋局模拟，并确定节点的估计值；
4. **回溯**：根据模拟结果向上依次更新祖先节点的估计值，并更新其状态

蒙特卡罗规划

- 在该过程中，算法首先根据一定的策略在搜索树上选择一个节点用于扩展，该策略被称为**搜索树策略** (Tree Policy)。
- 当**确定了扩展节点之后**，算法以该节点为根进行大量的随机模拟，并根据模拟的结果确定在该根节点下如何落子，并**更新祖先节点的估计值**，该策略一般称为**模拟策略或默认策略** (Default Policy)。

蒙特卡罗规划

- 算法1：蒙特卡罗规划（MCTS approach）

function MCTSSEARCH(s_0)

以状态 s_0 创建根节点 v_0 ;

while 尚未用完计算时长 **do**:

$v_l \leftarrow$ TREEPOLICY(v_0);

$\Delta \leftarrow$ DEFAULTPOLICY($s(v_l)$);

BACKUP(v_l, Δ);

end while

return $a(\text{BESTCHILD}(v_0))$;

蒙特卡罗规划

- 算法1描述了蒙特卡罗规划的伪代码，在该算法中，节点 v_0 是与初始状态 s_0 相对应的根节点， $s(v)$ 表示节点 v 所对应的状态。
- 更新策略为：
 - 设 v_l 为当前要模拟的节点， Δ 为报酬
 - 对 v_l 及祖先的模拟次数加1
 - v_l 的收益加 Δ
 - 更新 v_l 的祖先节点的收益，同类节点加 Δ ，非同类节点减 Δ

UCB和UCT算法

- 多臂老虎机模型
- UCB算法
- UCT算法

多臂老虎机模型



多臂老虎机模型

- 1952年，Robbins提出了一个统计决策模型
- 多臂老虎机：
 - 多臂老虎机拥有 k 个手臂，拉动每个手臂所能产生的回报遵循一定的概率，且互不相关，如何找到一个策略，使得拉动手臂的人获得最大的收益
 - 用于解决蒙特卡洛规划中选择落子点的问题

多臂老虎机模型

- 在游戏开始之前所有的手臂对玩家来说都是等价的，若想知道哪个手臂最好，就需要不断的去试探，并通过不断的试探发现规律，以此来推断出哪个手臂可能获得最大的回报。
- 由于玩家所拥有的试探次数是有限的，因此不得不在探索和利用间寻求一个平衡点，
 - **探索**就是通过进行更多的试探以获得更多的知识，这包括尽可能排除收益低的手臂和尽可能发现收益高的手臂，
 - 而**利用**则是充分利用当前已经获得的知识（对每个手臂所对应的收益高低的判断）获得尽可能大的回报

多臂老虎机模型

- 该模型所蕴含的在搜索过程权衡探索和利用的基本思想是对蒙特卡罗规划方法进行改进的核心内容之一。
 - 在该思想下的两个典型算法：
 - 信心上限算法UCB
(Upper Confidence Bound)
 - 信心上限树算法UCT
(Upper Confidence Bounds for Trees)
- 它们构成了现代蒙特卡罗博弈理论的基础算法

UCB算法

- 一般来说，多臂老虎机玩家开始动手玩时，会依照目前积累的经验来决定下一次要选择哪一台机器玩，这称之为利用 (exploitation)。
- 相对地，如果玩家不断地依照目前所获得的经验来决定，而不试图尝试其它的机器，则可能会忽略收益率更高的机器，因此适度地尝试其他机器是必须的，这称之为探索 (exploration)。
- 如何在利用与探索之间保持平衡，就是UCB试图解决的 ExE (exploitation vs. exploration) 问题。

UCT算法

- 虽然利用构造蒙特卡罗规划算法可以解决围棋博弈问题，
 - 但是由于蒙特卡罗规划方法在没有知识的指导时树的扩展层数较少，不利于最优解的获取，
- 因此可将UCB1算法加入蒙特卡罗规划树的构建过程中，形成了**信心上限树算法 (UCT)**

UCT算法

- 由于围棋的复杂度高，且极具欺骗性，对计算机程序提出了巨大的挑战。为了处理如此众多的可能情况，人工智能专家已经设计出一些算法，来限制搜索的范围，但它们都无法在大棋盘的比赛中战胜实力稍强的人类棋手。
- 2006年秋季，两位匈牙利研究人员报告了一种新算法，它的胜率比现有最佳算法提高了5%，能够在小棋盘的比赛与人类职业棋手抗衡，这种被称为UCT算法

UCT算法

- 法国两位学者将**UCT**（Upper Confidence Bound Apply to Tree）集成到一个他们称之为**MoGo**的程序中。该程序的胜率竟然比先前最先进的蒙特卡罗扩展算法几乎高出了一倍。
- **2007**年春季，**MoGo**在小棋盘的比赛中击败了实力强劲的业余棋手，在大棋盘比赛中也击败了实力稍弱的业余棋手，充分展示了能力
- 他们认为**UCT**易于实现，并有进一步完善的空间

UCT应用

- 2006年第一个基于UCT算法的围棋程序MoGo在 9×9 棋盘上达到了专业棋手的水平。
- 2009年8月基于UCT的开源程序Fuego在 9×9 棋盘上战胜了周俊勋九段。
- 2012年6月计算机围棋程序Zwn19S在 19×19 棋盘上达到了KGS的6段评级。
- 2013年3月围棋软件CrazyStone在受让四子的情况下。战胜日本棋手石田芳夫九段，其棋力已达到业余五、六段的水平
- 2016年……

朱迪亚·珀尔

Judea Pearl (1937 -):

- 1986年提出概率因果逻辑推理
- 2003年 获Allen Newell奖
- 2008年 获富兰克林奖
- 2011年 获得图灵奖。
- 2011年 获英国经济和政治科学学院的Lakatos奖。



概率推理 + 图论 + 认知 + 因果关系 +

因果推理

Judea Pearl: 要建立真正的人工智能，少不了因果推理



因果推理

在20世纪80年代，他带头推动机器以概率的方式进行推理。但是现在，他却成为了该领域最尖锐的批评者之一。他在自己的新书《The Book of Why: The New Science of Cause and Effect》中指出，人工智能的发展已经受到阻碍，因为它不能完全理解智能的真正含义。



模糊推理

拉特飛·扎德（**Lotfi A. Zadeh**，**1921**～）美國自動控制專家，美國工程科學院院士。1921年2月生於蘇聯巴庫。1949年獲哥倫比亞大學電機工程博士。1965年，扎德在《信息與控制》雜誌第8期上發表《模糊集》的論文，引起了各國數學家 and 自動控制專家們的注意。因發展模糊集理論的先驅性工作而獲電氣與電子工程師學會(IEEE)的教育勳章。1997年提出粒度計算。



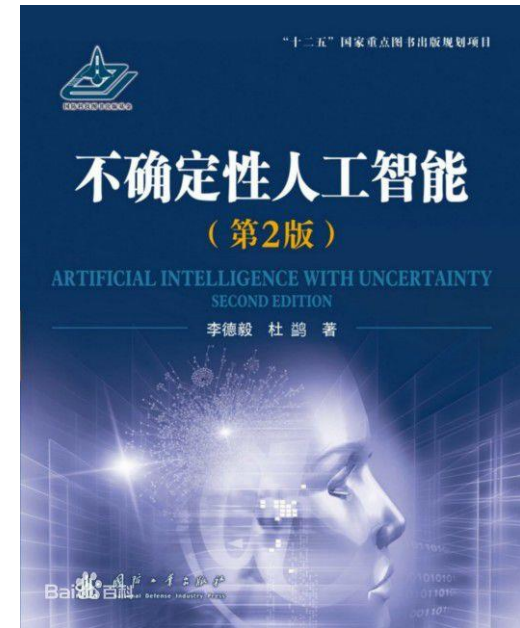
模糊推理

汪培庄，1936年生，湖北黄冈人，博士生导师，1957年毕业于北京师范大学数学系。提出了模糊随机落影空间理论的模型、因素空间理论。著有《模糊集合论及其应用》、《模糊集与随机集落影》，合著《可逆马尔可夫过程》。1988年，汪培庄教授指导的几位博士研制成功一台模糊推理机——分立元件样机。现为辽宁工程技术大学教授。



模糊推理

李德毅，指挥自动化和人工智能专家。1944年11月出生于江苏省泰县。中国工程院院士、欧亚科学院院士。1967年毕业于南京工学院（现东南大学），1983年获英国爱丁堡海里奥特·瓦特大学博士学位。中国人工智能学会理事长。提出云模型、自动驾驶汽车智慧大脑。著作《不确定性人工智能》，利用认知的物理学方法，寻找不确定性知识和智能处理中的规律性。



新的DUCG平台

原中国科协党组副书记、副主席、中国人工智能学会不确定性人工智能专委会主任张勤教授开发DUCG (Dynamic Uncertain Causality Graph) 平台，在医疗和核电站方面成功应用。



动态不确定因果图DUCG



动态不确定因果图DUCG Dynamic Uncertain Causality Graph 在智能辅助临床诊断 和分诊中的应用

张勤

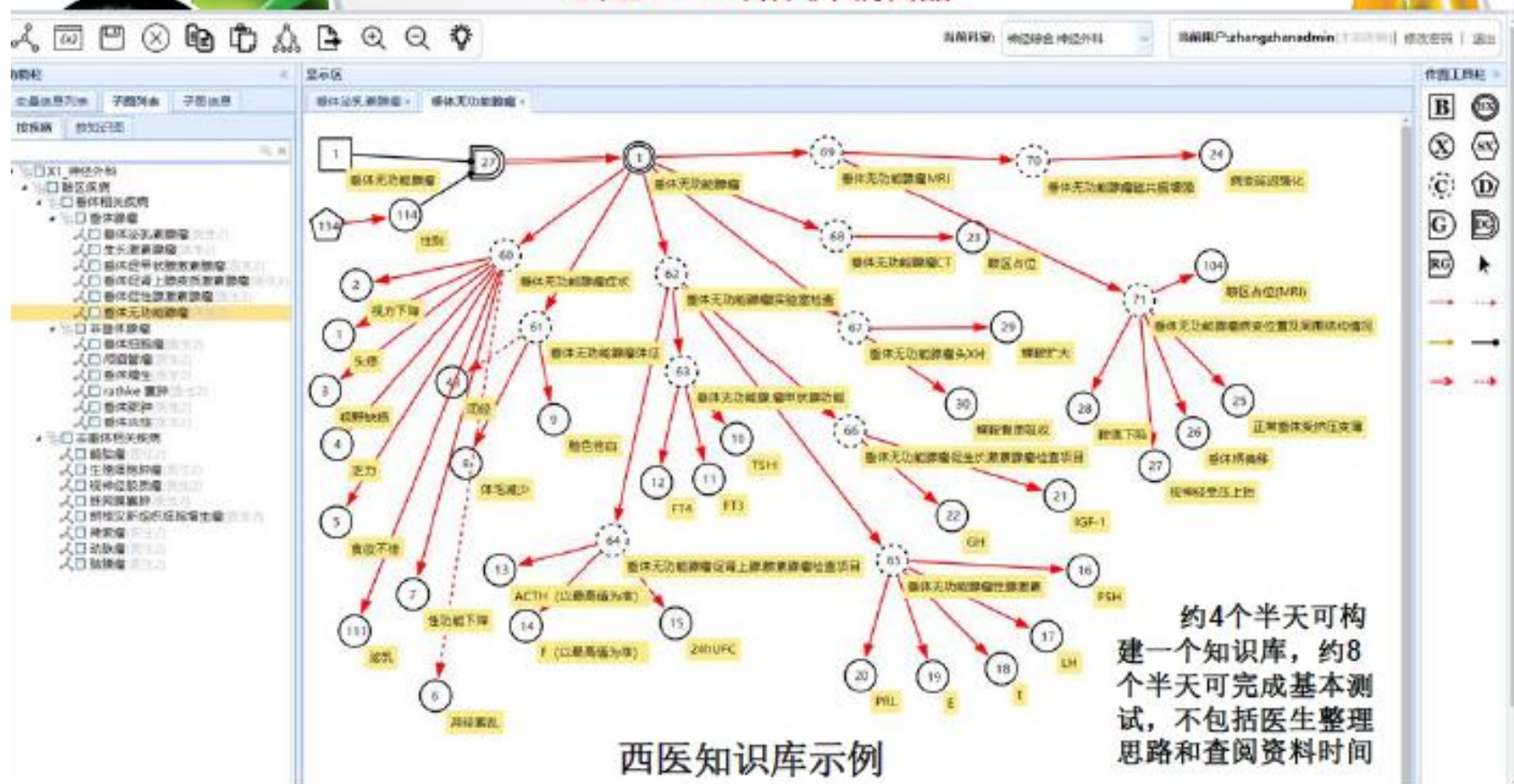
清华大学核研院、计算机系教授、博导
中国人工智能学会不确定性人工智能专委会主任
国际核能院 (INEA) 院士、全国政协常委、中国科协荣誉委员

动态不确定因果图DUCG

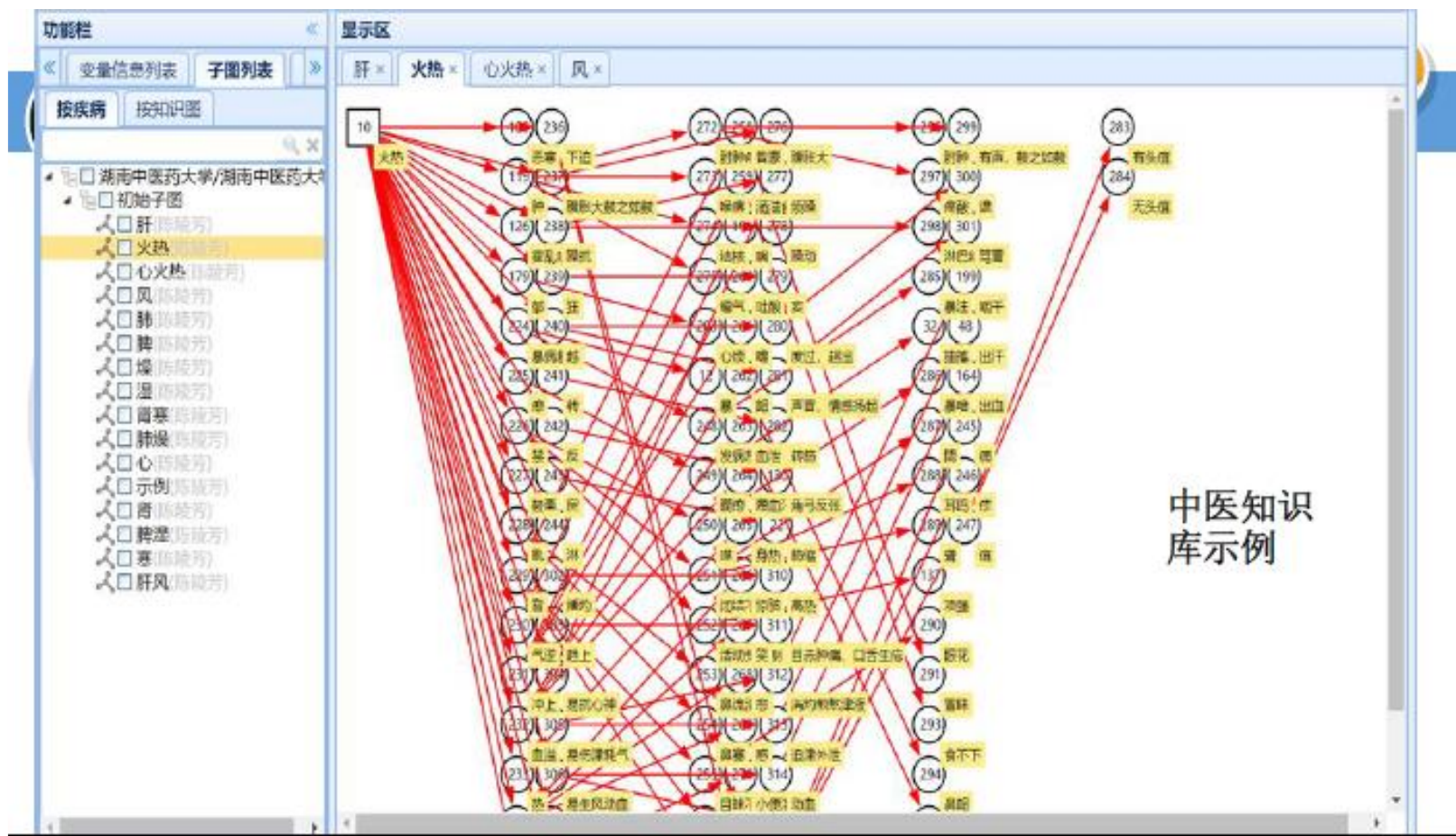


动态不确定因果图DUCG

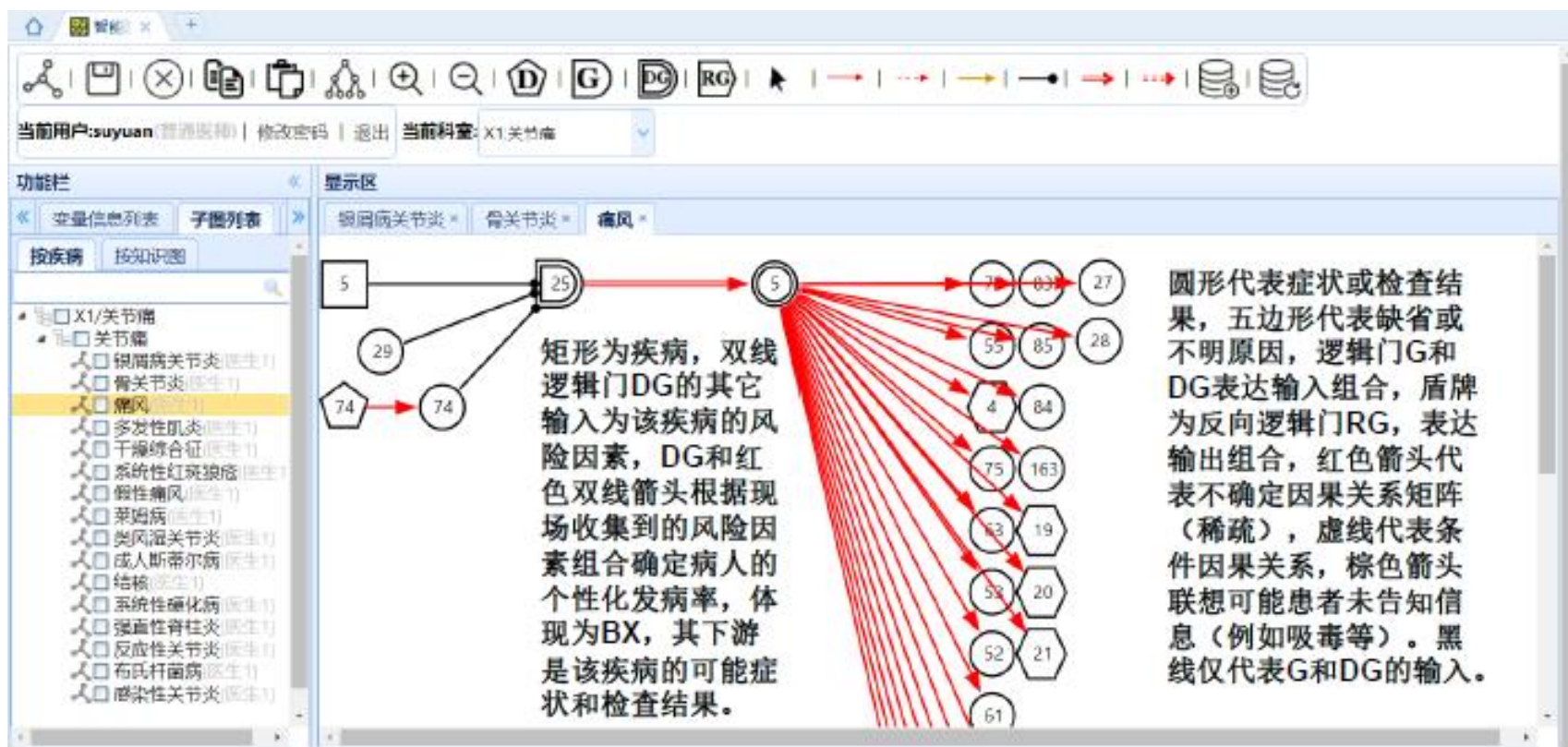
云上DUCG知识库编辑器



动态不确定因果图DUCG



动态不确定因果图DUCG



模块化知识库构建，通常按疾病（矩形）进行正向（由因到果）表达

动态不确定因果图DUCG

清睿智能医学辅助诊断平台

张老师(主任医师)

选择患者 常规诊断 精细诊断 历史记录 推荐检测 生成病历 暂存证据 读取证据

未知设正常 正常设未知 全部设未知 已知疾病

患者信息

显示文本记录 修改记录 辅助诊断结果

请输入需要查找的信息,支持拼音及首字母搜索...

医院: 某某医院 科别: 某某科室 就诊时间: 2018-09-27 10:58:54
 患者: 遂宁_系统性 性别: 女 年龄: 26 职业:
 红斑狼疮 婚姻: 未婚 患者编号:
 1534753231392
 住址:
 过敏史: 无

风险因素 症状 体征 实验室检查 影像学检查 病理检查

主诉症状 关节痛

必选

切换必选症状

症状选择区域，包含多个彩色按钮，如：年龄-青年、指甲顶针样改变-阴性、脱发-阴性、多汗-阳性、牛羊痘感染-无、乏力-阳性、食欲不振-阴性、痛风石-阳性、既往病史或家族过敏史-阴性、游走性红斑-阳性、神经根病(如带状疱疹)、前驱史或山区进入史-无、晕倒-阳性、性别-女、结膜炎-阳性、尿蛋白-阳性、偶(发)发、腹痛-阳性、皮疹-阳性、类风湿因子-阳性、关节畸形-阳性、关节僵硬或肿胀(非疼痛原因引起)、关节疼痛(肿胀)-慢性、关节疼痛(大小)-小、关节疼痛(中轴外周)-外周、关节疼痛(四肢远端)-非远端对称、恶心-阳性、呕吐-阳性、精神状态-阳性、高钾血症(高血钾)-阴性、肺动脉高压-阳性、...

其他

推荐检测结果

待查类别	归类推荐度	分项描述	分项推荐度
血常规	64.87%	血常规(白细胞)	61.84%
		血色素	2.66%
		血小板	0.36%
自身免疫性抗体	8.98%	抗dsDNA抗体	3.65%
		抗CCP	3.61%
		抗SM抗体	1.47%
		抗SSA	0.14%
		抗SSB	0.10%
病变关节X光检查	7.89%	X光检查发现存在典型软骨或关节囊钙化	3.92%
		X线检查发现关节破坏	2.09%

共 72 条记录 每页 10 条

1 2 3 4 5 ... 8

推荐检测界面：
动态优化定制个性化的临床路径，不再依赖临床指南。

完成

75%

动态不确定因果图DUCG




关于DUCG医疗诊断的总结和展望



- ◆ 已能诊断单发和并发病，正在开发继发病推理机和推荐处方功能。
- ◆ DUCG为大数据医学影像识别提供做影像检查的推荐，智能识别后将结果回传DUCG，由DUCG综合给出诊断结论，形成医疗诊断的全流程全覆盖。
- ◆ 基层医生只需要进行如何使用DUCG和准确获取信息的培训即可达到专家看病的水平，大大减少基层医生的培训和使用成本。
- ◆ DUCG增加处方功能后，可减少用药乱象，提高全科精准治疗水平。
- ◆ 减少漏诊误诊和漏检误检，实现医保个性化精准检测、治疗和控费。
- ◆ 颠覆医生职业，除少数专家型医生外，一般医生在DUCG帮助下，其诊断和处方水平即可达到专家级，大大缩短学习时间，解决优质医疗资源供需矛盾，增加基层医生处方权，切实实现医疗资源下沉。
- ◆ 未来多数医生将主要关注如何做好手术、手工检查、以及使用仪器，需要大量经验积累和传承的逆向诊断环节由DUCG算法完成。



核电站安全运维专家系统

 核电站安全运维智能专家系统

DUCG Intelligent System for Safe Operation of Nuclear Power Plants

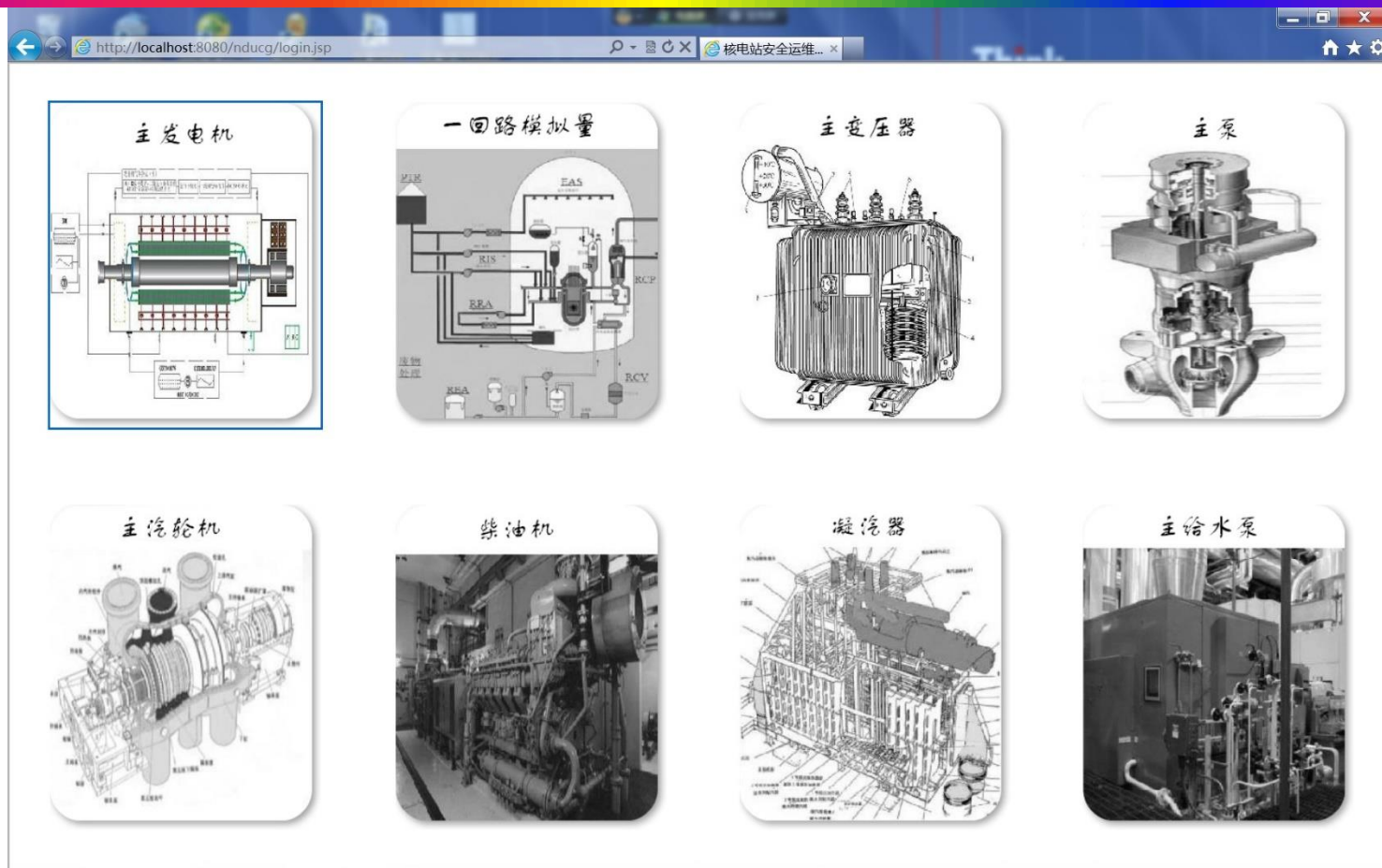
English

用户名:

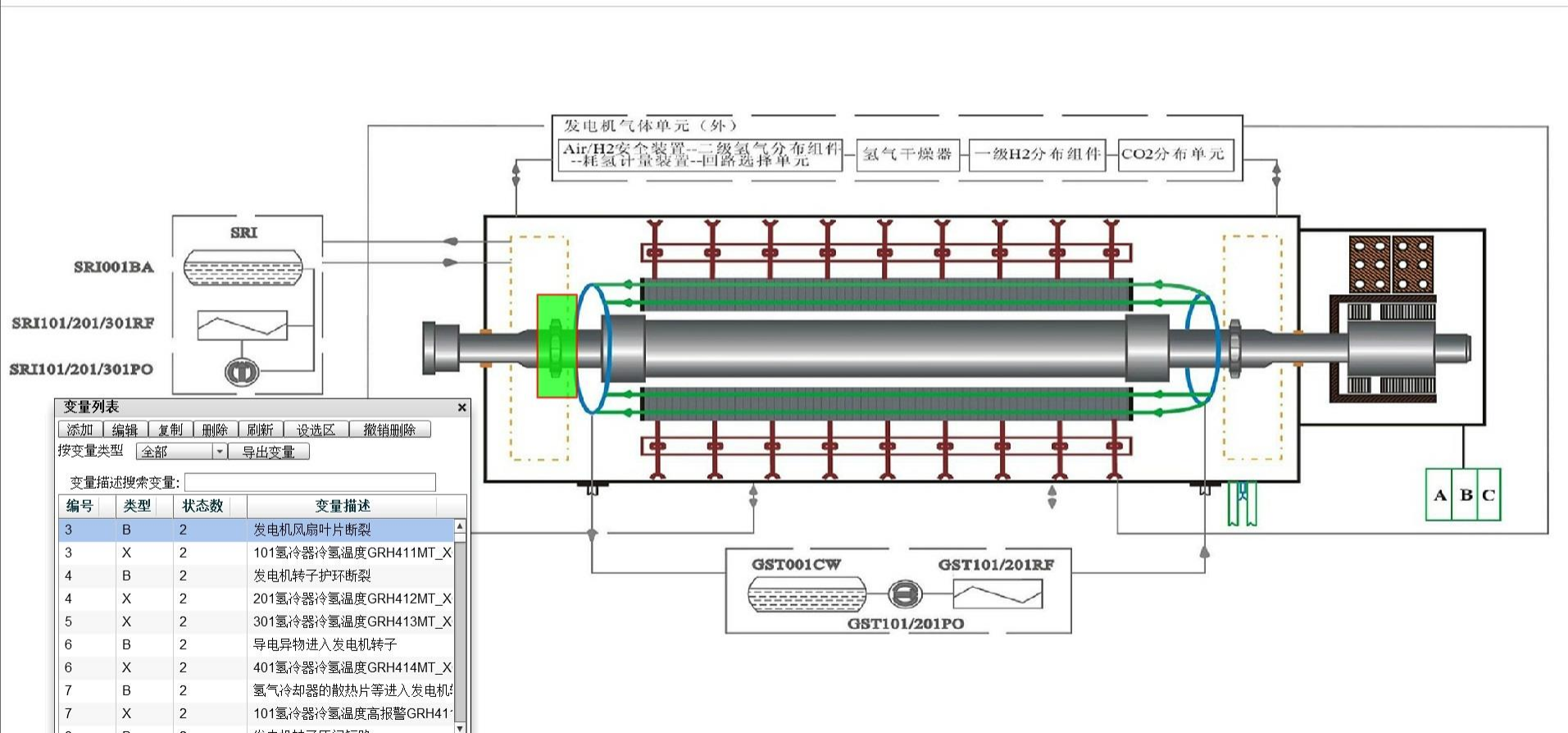
密码:

平台: 应用平台

岭东核电站8个拟监测系统



岭东核电站发电机图形知识库编辑界面



变量列表

添加 编辑 复制 删除 刷新 设选区 撤销删除

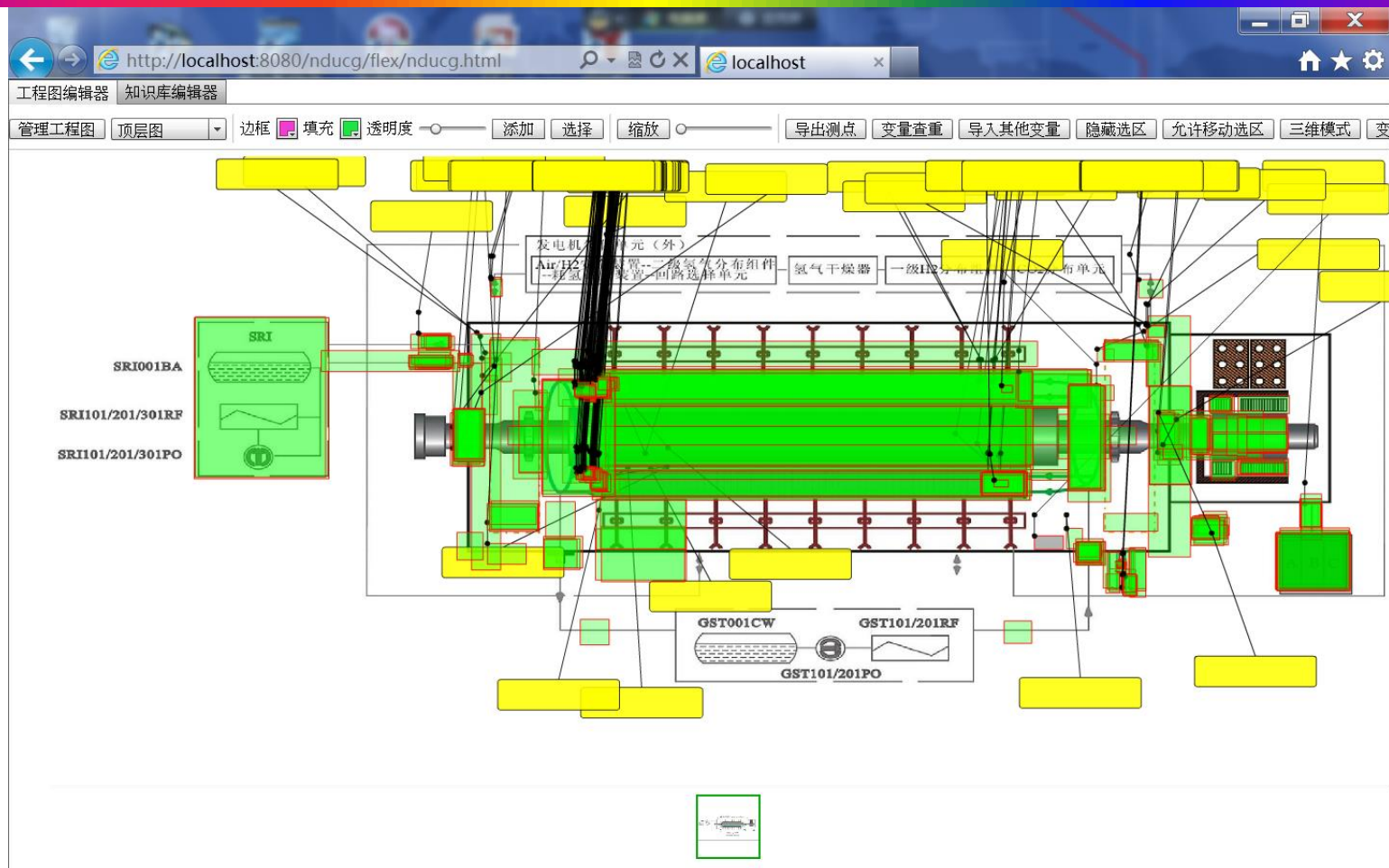
按变量类型 全部 导出变量

变量描述搜索变量:

编号	类型	状态数	变量描述
3	B	2	发电机风扇叶片断裂
3	X	2	101氢冷器冷氢温度GRH411MT_X
4	B	2	发电机转子护环断裂
4	X	2	201氢冷器冷氢温度GRH412MT_X
5	X	2	301氢冷器冷氢温度GRH413MT_X
6	B	2	导电异物进入发电机转子
6	X	2	401氢冷器冷氢温度GRH414MT_X
7	B	2	氢气冷却器的散热片等进入发电机
7	X	2	101氢冷器冷氢温度高报警GRH41

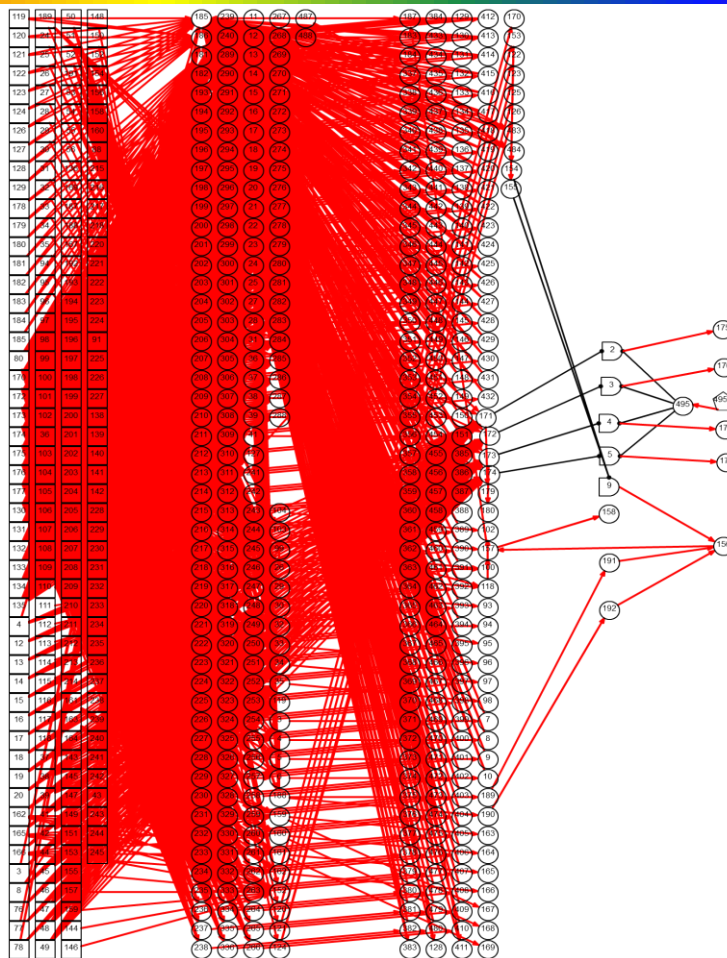
2018/10/14

从图形上定义变量

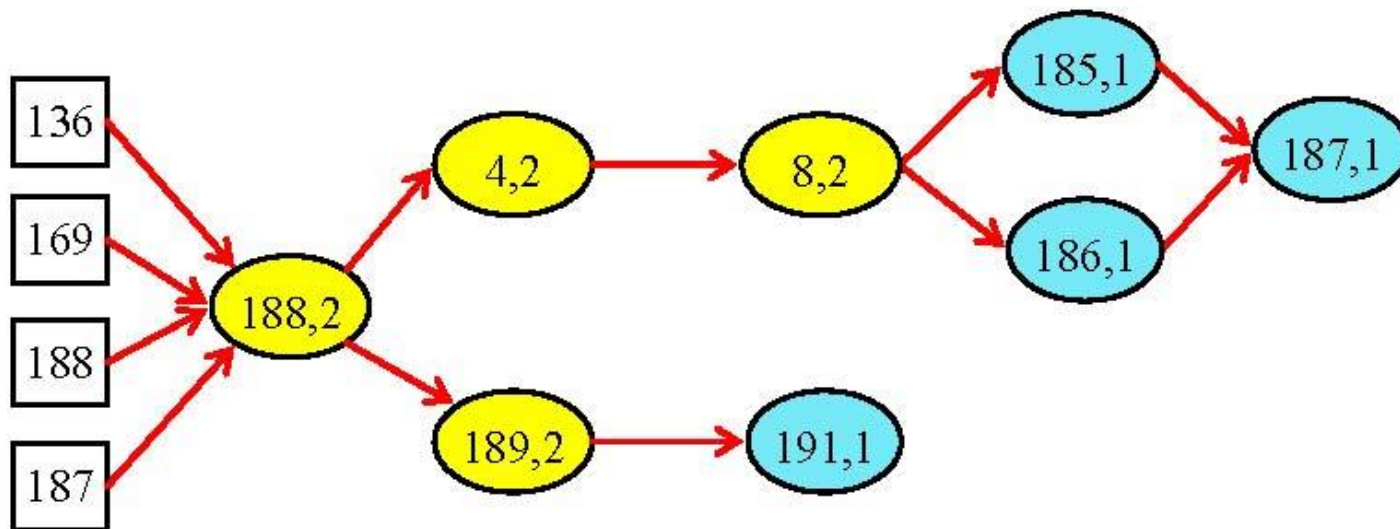


岭东核电站发电机知识库

岭东核电站发
电机知识库，
包含659个变
量，由32个子
DUCG图合成



诊断结果



诊断结果的逻辑关系图显示

内容提要



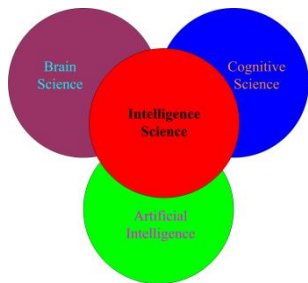
引言

不确定性推理

统计学习

大数据分析

展望



统计学学习方法

近百年来，统计学得到极大的发展。我们可用下面的框架粗略地刻划统计学发展的过程：

- 1900–1920 数据描述
- 1920–1940 统计模型的曙光
- 1940–1960 数理统计时代
 - 随机模型假设的挑战
 - 松弛结构模型假设
- 1990–1999 建模复杂的数据结构

统计学习方法

从1960年至1980年间，统计学领域出现了一场革命，要从观测数据对依赖关系进行估计，只要知道未知依赖关系所属的函数集的某些一般的性质就足够了。引导这一革命的是60年代的四项发现：

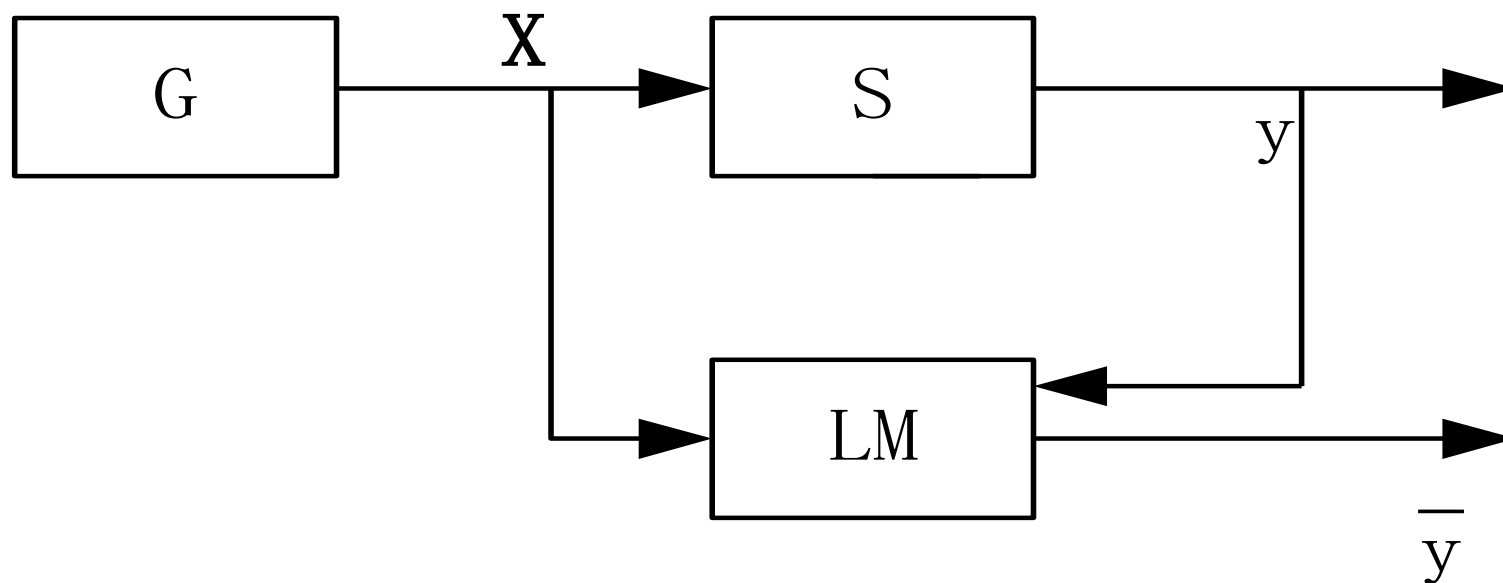
- Tikhonov, Ivanov 和 Philips 发现的关于解决不适定问题的正则化原则；
- Parzen, Rosenblatt 和 Chentsov 发现的非参数统计学；
- Vapnik 和 Chervonenkis 发现的在泛函数空间的大数定律，以及它与学习过程的关系；
- Kolmogorov, Solomonoff 和 Chaitin 发现的算法复杂性及其与归纳推理的关系。

这四项发现也成为人们对学习过程研究的重要基础。

统计机器学习

- 1984: Valiant提出学习框架PAC
- 1989: 数据集知识发现KDD
- 1991: Pawlak 粗糙集专著
- 1992: 波兰Kiekrz召开第一届粗糙集国际讨论会
- 1995: Vapnik提出SVM
- 2000: 知识发现系统MSMiner
- 2005: 第一届IEEE国际粒度计算大会在北京召开
- 2006: Hinton等人提出深度学习的概念
- 2012: 《纽约时报》刊文称“大数据时代”已经降临。

机器学习的问题



机器学习的问题

- 产生器 (G)，产生随机向量 x 属于 R^n ，它们是从固定但未知的概率分布函数 $F(x)$ 中独立抽取的。
- 训练器 (S)，对每个输入向量 x 返回一个输出值 y ，产生输出的根据是同样固定但未知的条件分布函数 $F(y|x)$ 。
- 学习机器 (LM)，它能够实现一定的函数集 $f(x, a)$ ， a 属于 A ，其中 A 是参数集合。

Valiant学习理论



哈佛大学计算机科学家莱斯利·瓦伦特（Leslie Valiant）获2010年图灵奖，他的最杰出贡献是**probably approximately correct (PAC)** 学习理论，被认为是机器学习的重要理论。

Valiant学习理论



BIRTH: 28 March 1949, Budapest, Hungary

EDUCATION: Latymer Upper School, London England; King's College, Cambridge, England (BA, Mathematics, 1970); Imperial College, London, England (DIC in Computing Science); University of Warwick, England (PhD, Computer Science, 1974)

被引用 4215 次

Valiant学习理论

***PAC* learning. Invented by
L.Valiant in 1984.**

***L.G.Valiant*. A theory of the
learnable , Communications. of
the ACM, 1984, vol 27, 11, pp.
1134-1142.**

**This probably approximately
correct (PAC) model has given
rise to a fruitful research area
now known as computational
learning theory.**



Valiant学习理论

Valiant 认为一个学习机必须具备下列性质：

- (1) 机器能够证明地学习所有类的概念。更进一步，这些类可以特征化。
- (2) 对于通用知识概念类是合适的和不平常的。
- (3) 机器演绎所希望的程序的计算过程要求在可行的步数内。

Valiant学习理论

假设 X 是实例空间，一个概念是 X 的一个子集。如果实例在概念中则为正例，否则为反例。概念表示是一种概念的描述，概念类是一组概念表示。学习模型是概念类的有效的可学习性。Valiant 学习理论仅要求对目标概念的很好近似具有极高的概率。允许学习者产生的概念描述与目标概念有一个小的偏差 ϵ ，它是学习算法的一个输入参数。并且，允许学习者失败的概率为 δ ，这也是一个输入参数。两种概念之间的差别采用在实例空间 X 的分布概率 D 来评测：

$$\text{diff}_D(c1, c2) = \sum D(x)$$

Valiant学习理论

根据协议，一个概念类 C 是可学习的当且仅当有一种算法 A ，使用协议，对所有的目标概念表示 $c^* \in C$ 和全部分布 D ，

(1) 执行时间是与 $1/\epsilon$ 、 $1/\delta$ 、 c^* 数目和其它相关参数有关的多项式。

(2) 输出 C 中的概念 c 具有概率 $1-\delta$ ，

$$\text{diff}_D(c, c^*) < \epsilon$$

支持向量机

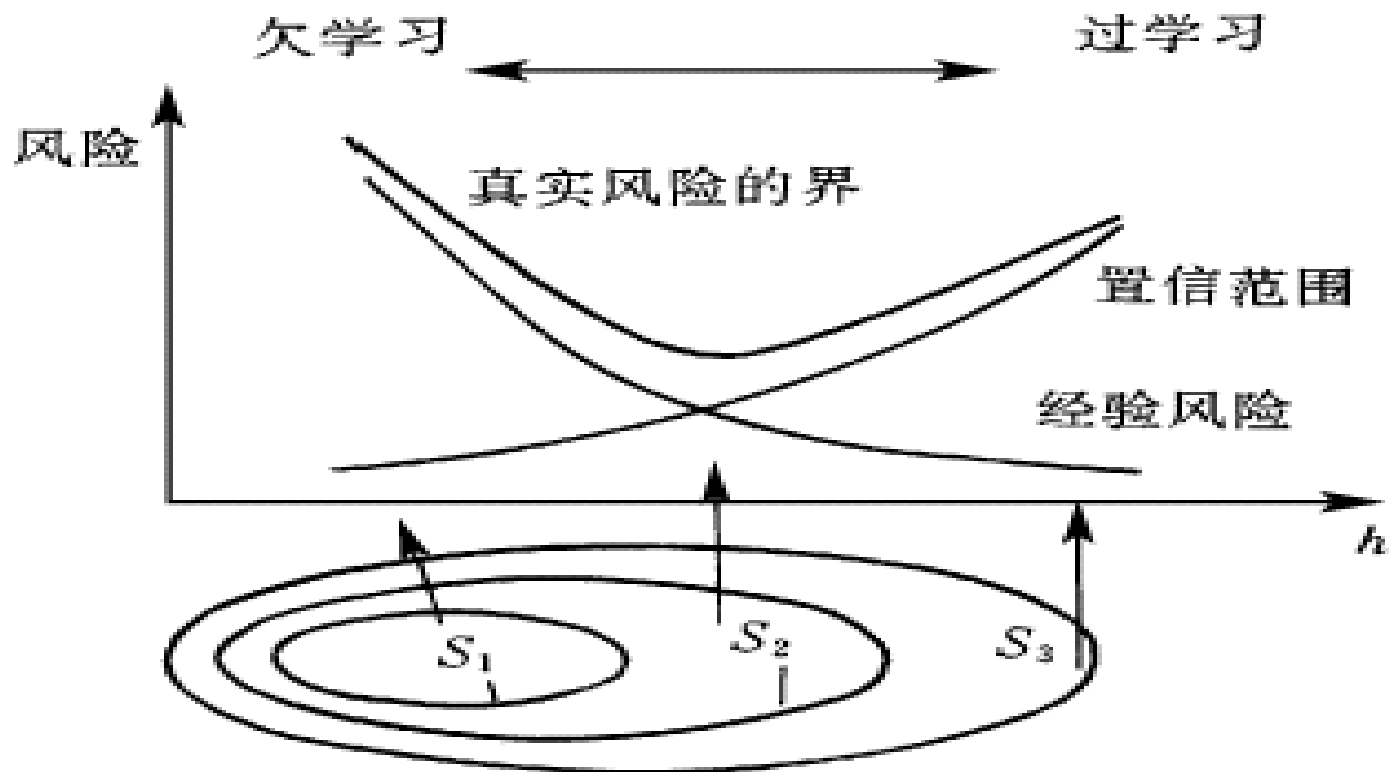
- SVM是一种基于统计学习理论的机器学习方法，它是由Boser, Guyon, Vapnik在COLT-92上首次提出，从此迅速发展起来
- Vapnik V N. 1995. The Nature of Statistical Learning Theory. Springer-Verlag, New York
- Vapnik V N. 1998. Statistical Learning Theory. Wiley-Interscience Publication, John Wiley&Sons, Inc
- 目前已经在许多智能信息获取与处理领域都取得了成功的应用。



结构风险最小化

- 传统机器学习方法中普遍采用的经验风险最小化原则在样本数目有限时是不合理的, 因此, 需要同时最小化经验风险和置信范围。
- 统计学习理论提出了一种新的策略, 即把函数集构造为一个函数子集序列, 使各个子集按照VC维的大小排列; 在每个子集中寻找最小经验风险, 在子集间折衷考虑经验风险和置信范围, 取得实际风险的最小。这种思想称作**结构风险最小化** (Structural Risk Minimization), 即SRM准则。

结构风险最小化



函数集子集： $S_1 \subset S_2 \subset S_3$

VC 维： $h_1 \leq h_2 \leq h_3$

深度学习

- 深度学习是使用包含复杂结构或由多重非线性变换构成的多个处理层对数据进行高层抽象的算法
- 深度学习是一种基于对数据进行表征学习的方法。观测值（例如一幅图像）可以使用多种方式来表示，如每个像素强度值的向量，或者更抽象地表示成一系列边、特定形状的区域等。而使用某些特定的表示方法更容易从实例中学习任务（例如，人脸识别或面部表情识别）。深度学习的好处是用非监督式或半监督式的特征学习和分层特征提取高效算法来替代手工获取特征

深度学习

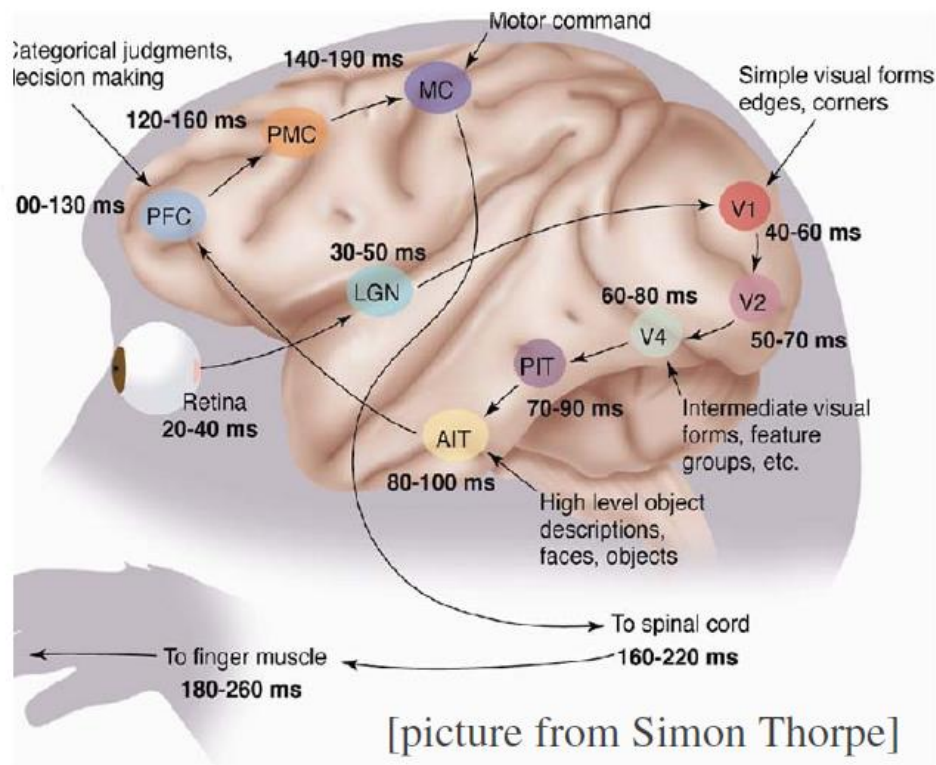
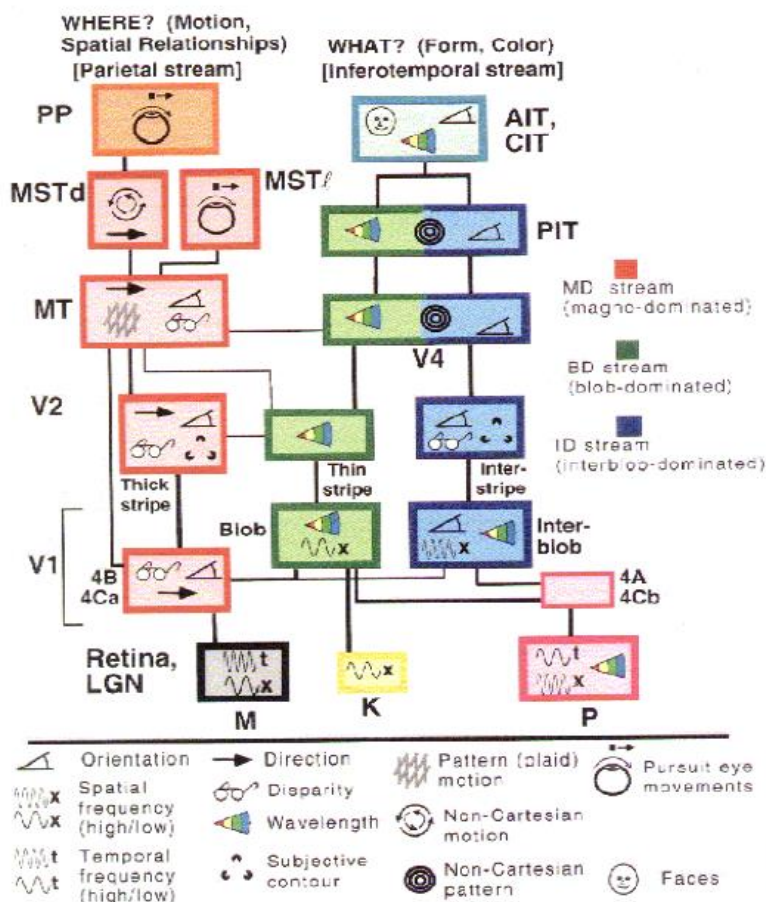
- 2006年的3篇关于深度学习的突破性论文：
- Hinton, G. E., Osindero, S. and Teh, Y., A fast learning algorithm for deep belief nets. *Neural Computation* 18:1527-1554, 2006
- Yoshua Bengio, Pascal Lamblin, Dan Popovici and Hugo Larochelle, Greedy LayerWise Training of Deep Networks, in J. Platt et al. (Eds), *Advances in Neural Information Processing Systems 19 (NIPS 2006)*, pp. 153-160, MIT Press, 2007
- Marc'Aurelio Ranzato, Christopher Poultney, Sumit Chopra and Yann LeCun Efficient Learning of Sparse Representations with an Energy-Based Model, in J. Platt et al. (Eds), *Advances in Neural Information Processing Systems (NIPS 2006)*, MIT Press, 2007

深度学习

- 2012年6月，《纽约时报》披露了Google Brain项目，吸引了公众的广泛关注。这个项目是由著名的斯坦福大学的机器学习教授Andrew Ng和在大规模计算机系统方面的世界顶尖专家Jeff Dean共同主导，用16000个CPU Core的并行计算平台训练一种称为“深度神经网络”（DNN, Deep Neural Networks）的机器学习模型（内部共有10亿个节点。这一网络自然是不能跟人类的神经网络相提并论的。要知道，人脑中可是有150多亿个神经元，互相连接的节点也就是突触数更是如银河沙数。曾经有人估算过，如果将一个人的大脑中所有神经细胞的轴突和树突依次连接起来，并拉成一根直线，可从地球连到月亮，再从月亮返回地球），在语音识别和图像识别等领域获得了巨大的成功。

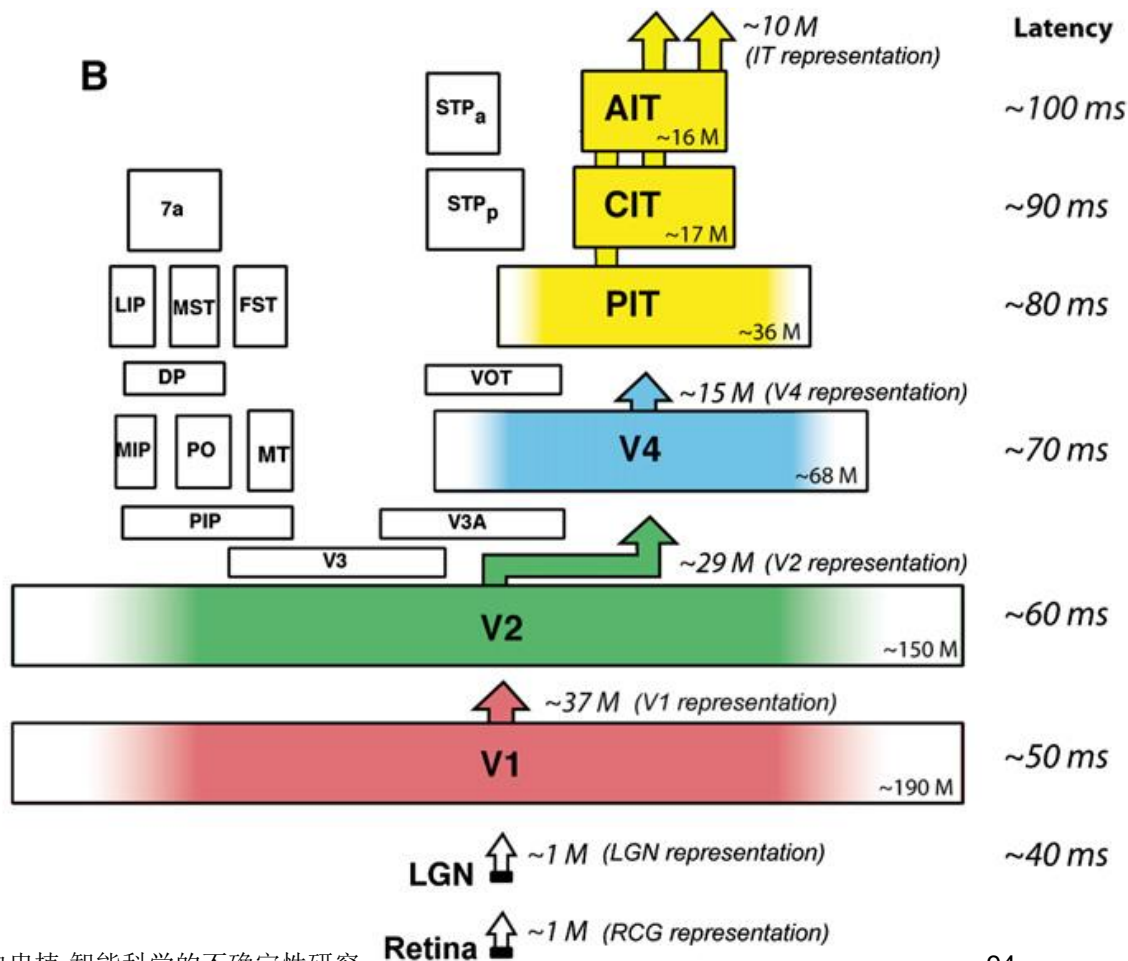
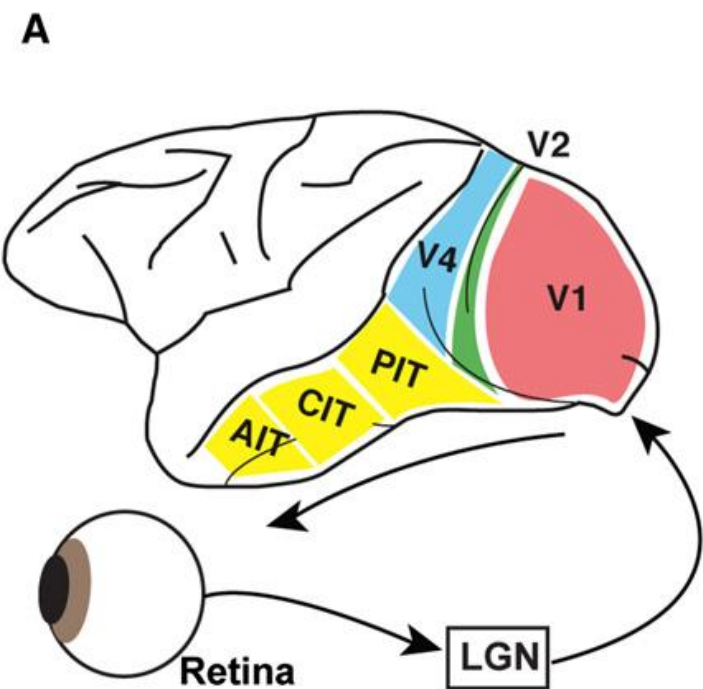
人脑视觉机理

- The ventral (recognition) pathway in the visual cortex has multiple stages
- Retina - LGN - V1 - V2 - V4 - PIT - AIT



[Gallant & Van Essen]

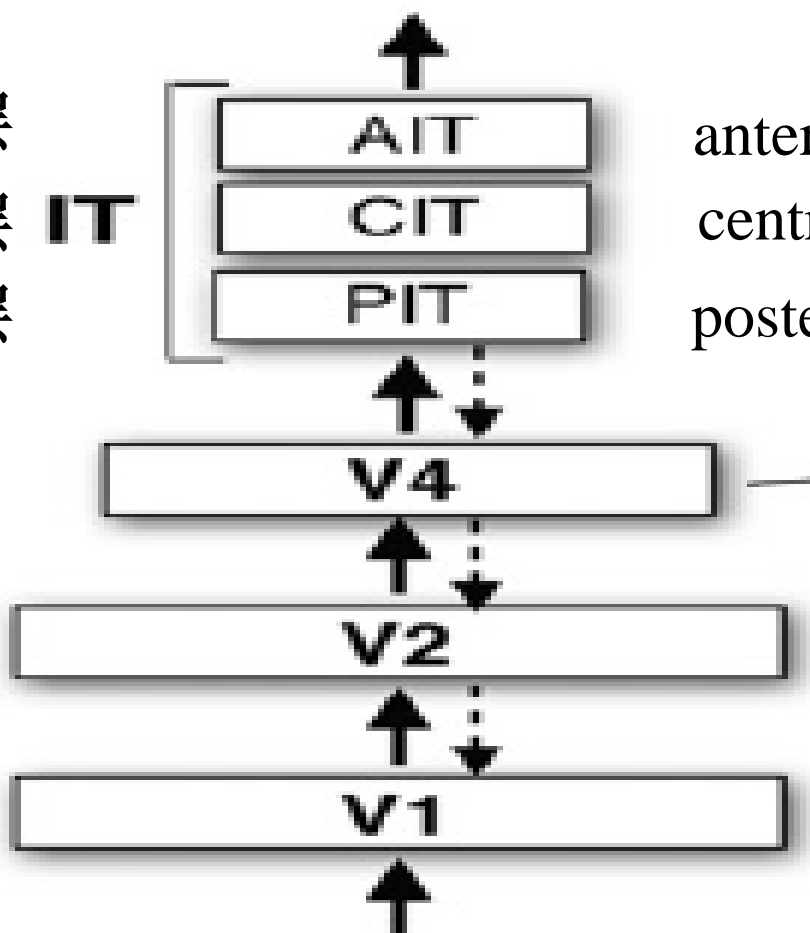
视觉通路



视觉从V1到IT四层

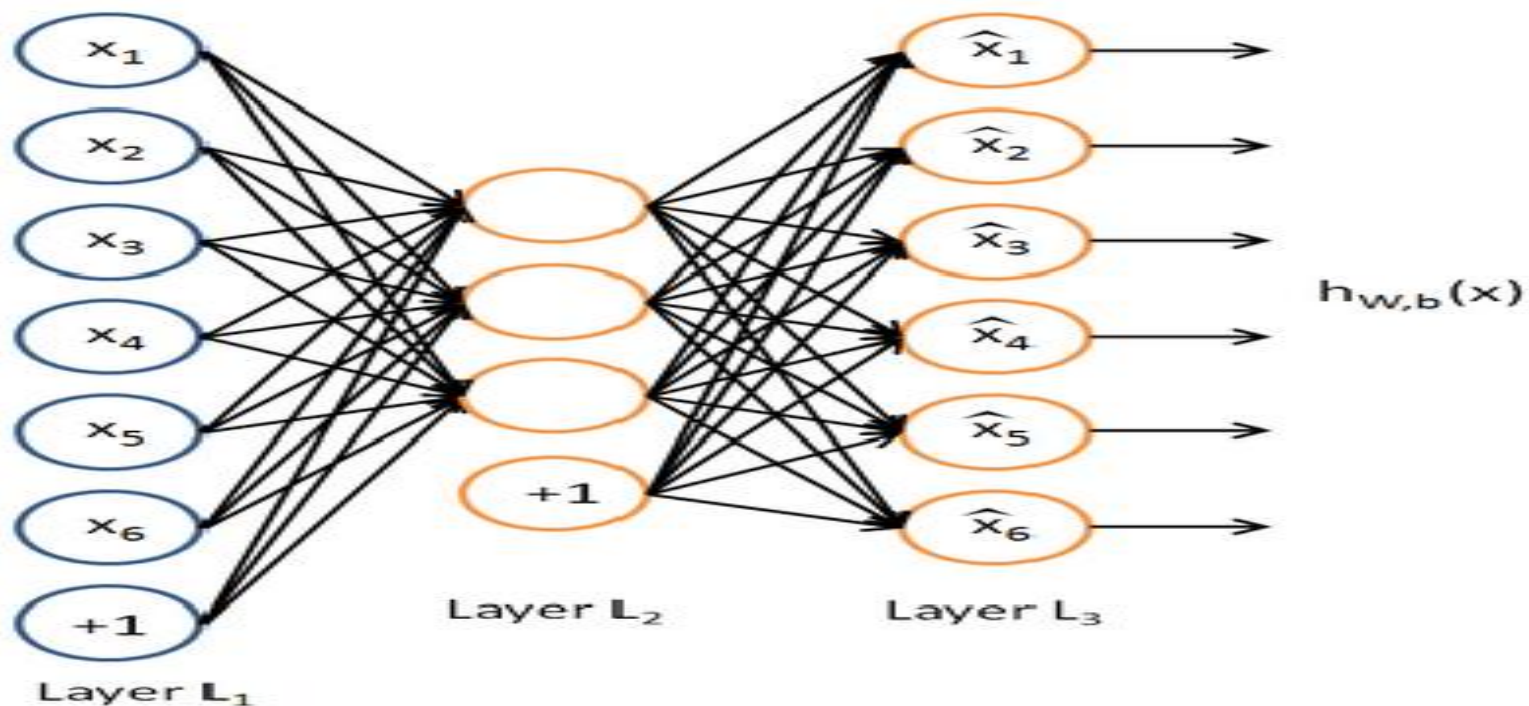
前颞叶皮层
中颞叶皮层
后颞叶皮层

IT



anterior inferotemporal
central inferotemporal
posterior inferotemporal

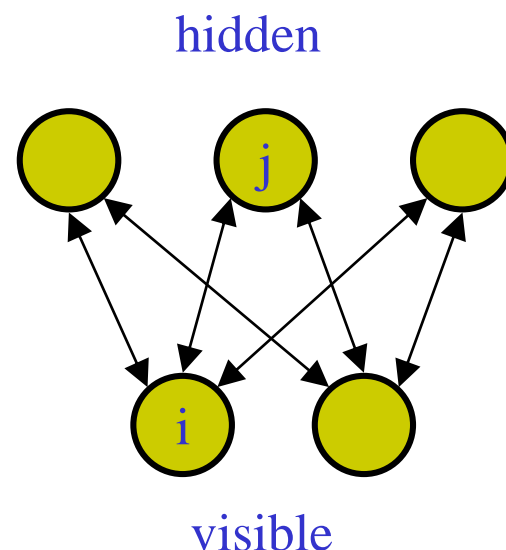
自编码神经网络



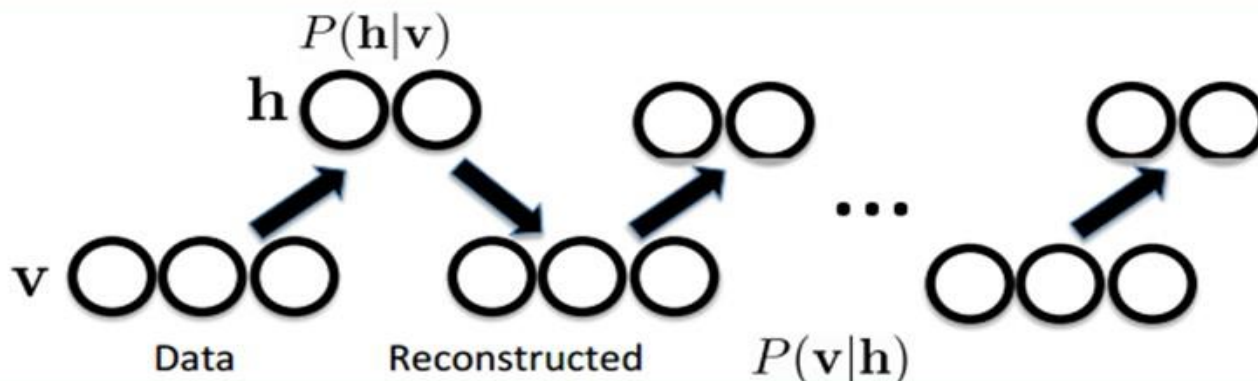
自编码神经网络

受限玻尔兹曼机

- 受限玻尔兹曼机 (restricted Boltzmann machine, RBM) 是一个单层的随机神经网络 (通常我们不把输入层计算在神经网络的层数里), 本质上是一个概率图模型。输入层与隐层之间是全连接, 但层内神经元之间没有相互连接。每个神经元要么激活 (值为1) 要么不激活 (值为0), 激活的概率满足 sigmoid 函数。



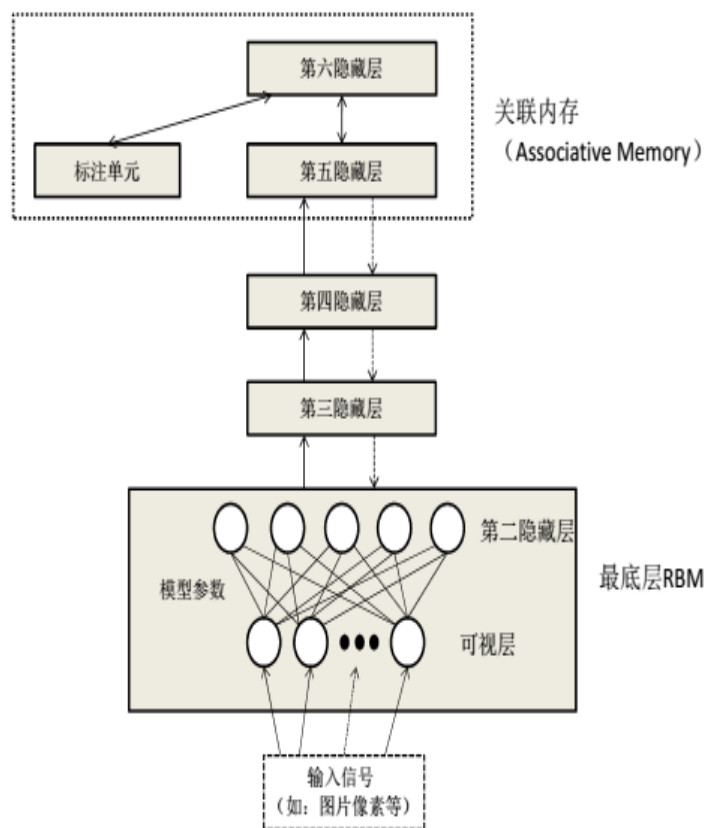
受限玻耳兹曼机 (RBM)



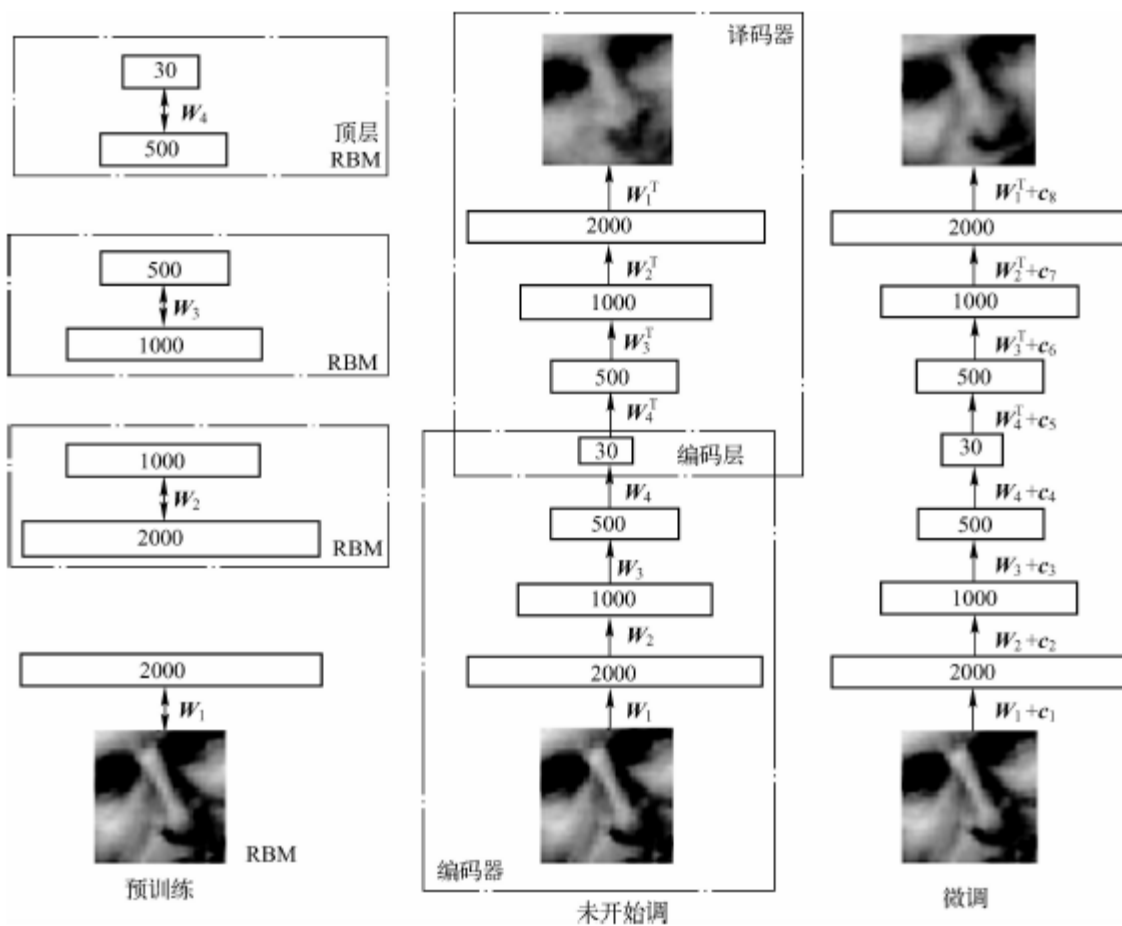
- 因为节点之间不存在连接，**隐藏节点之间是条件独立的**，即 $p(h|v) = p(h_1|v) \cdots p(h_n|v)$ 。
- 在已知隐藏层 h 的情况下，所有的**可视节点都是条件独立的**。
 $p(v|h) = p(v_1|h) \cdots p(v_n|h)$ 。
- 当输入 v 的时候，通过 $p(h|v)$ 可以得到隐藏层 h ，而得到隐藏层 h 之后，通过 $p(v|h)$ 又能得到可视层。
- **隐藏层就是可视层另外一种表达**。
- 训练：最大似然估计。

深度信念网络

- 2006 年，欣顿等提出了一种深度信念网络（deep belief nets, DBN）。一个深度神经网络模型可被视为由若干个 RBM 堆叠在一起，这样一来，在训练的时候，就可以通过由低到高逐层训练这些 RBM 来实现。

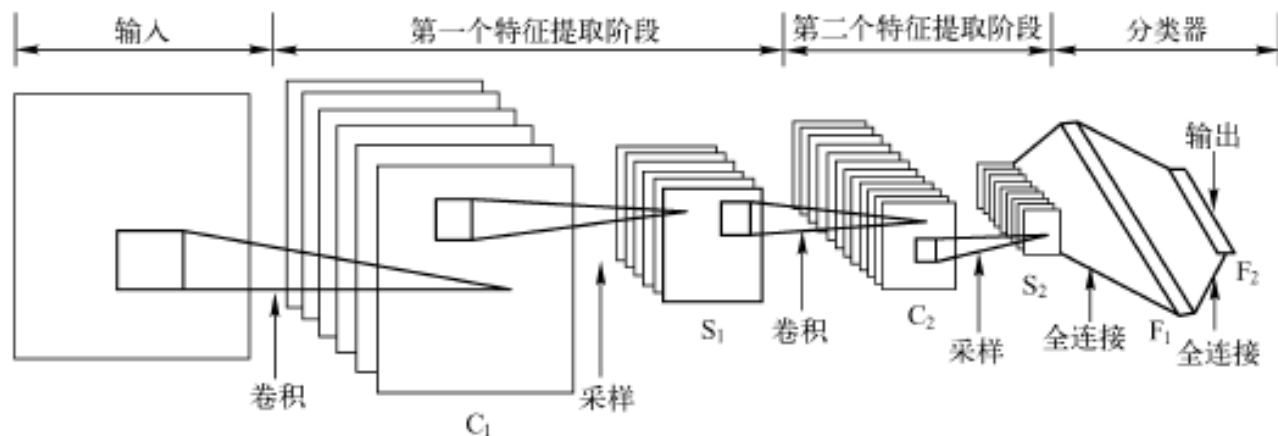


深度信念网络



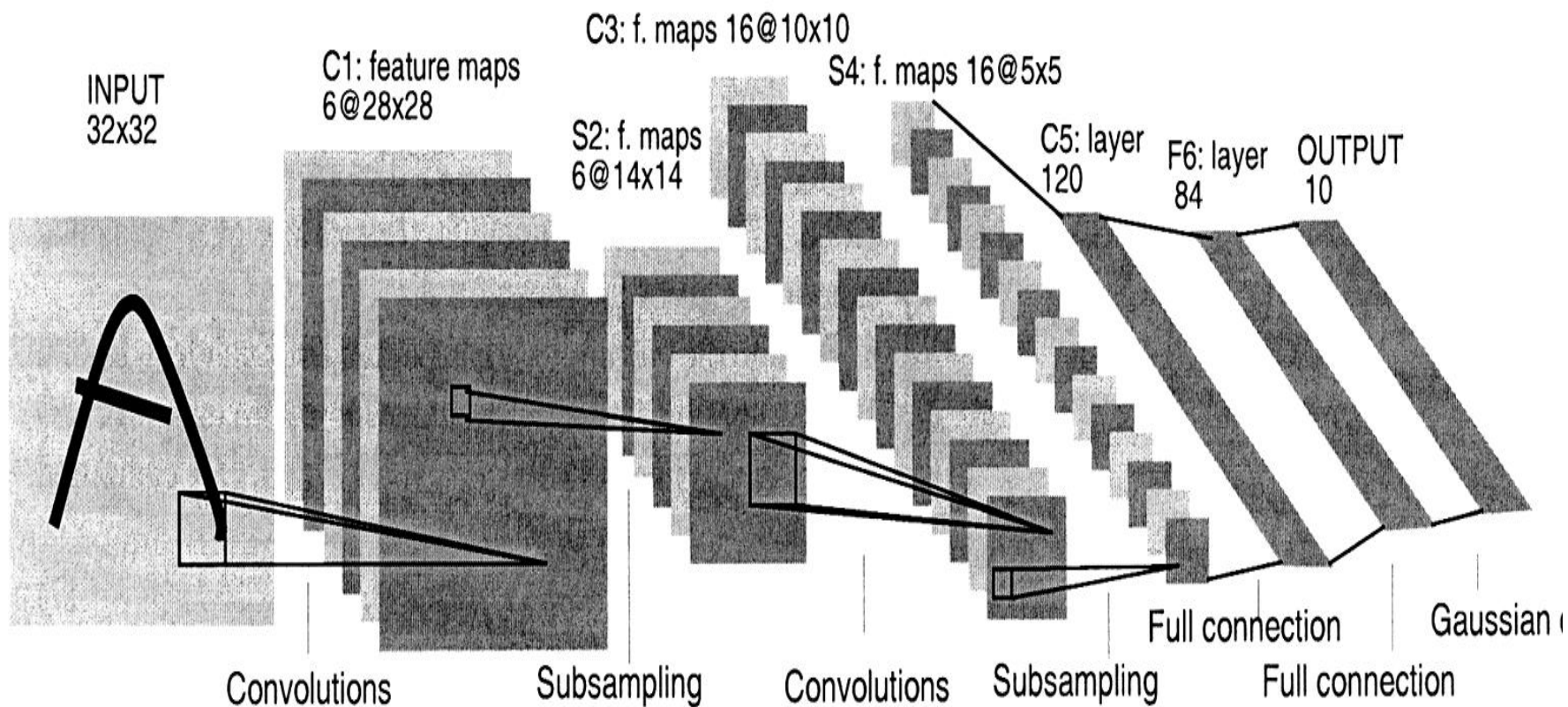
卷积神经网络

卷积神经网络是一种多阶段全局可训练的人工神经网络模型，它可以从经过少量预处理，甚至原始数据中学习到抽象的、本质的和高阶的特征，在车牌检测、人脸检测，手写体识别和目标跟踪等领域得到了广泛的应用。

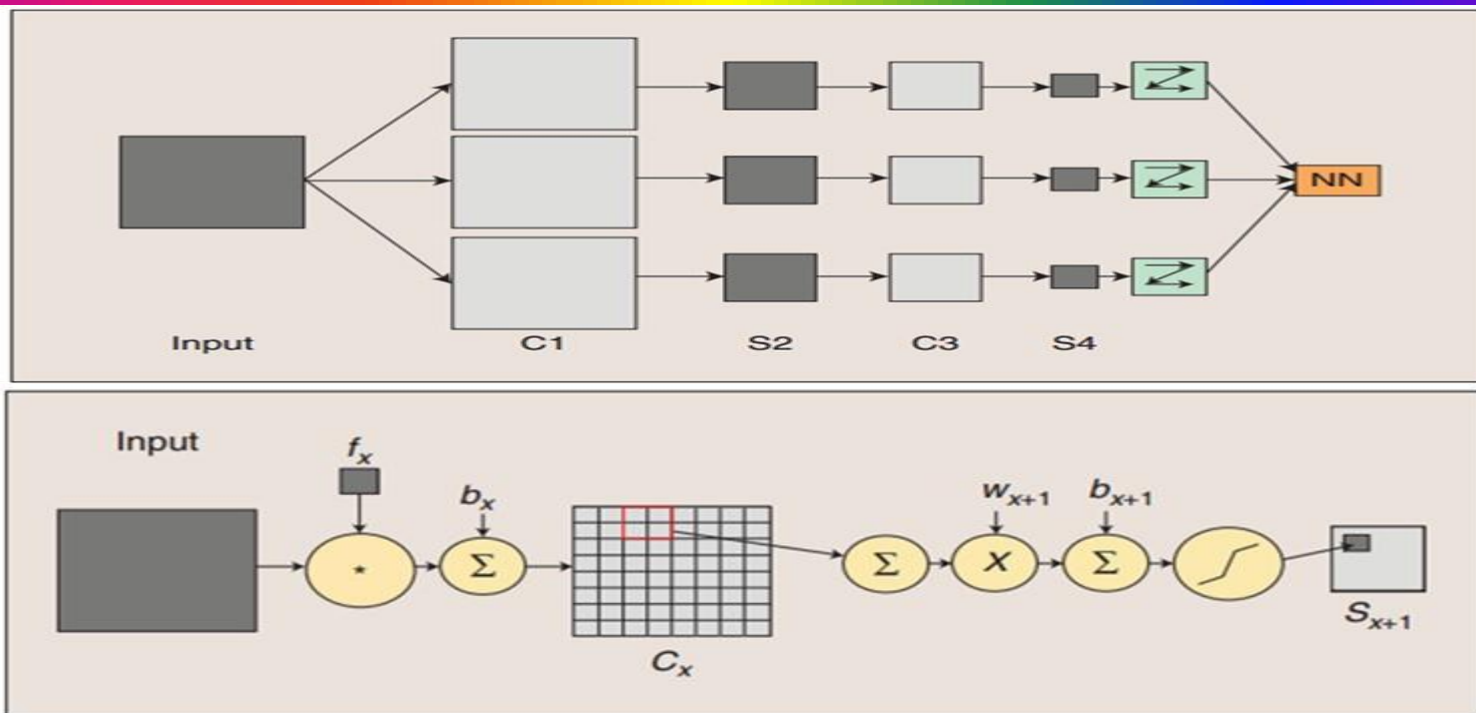


卷积神经网络

The architecture of LeNet5

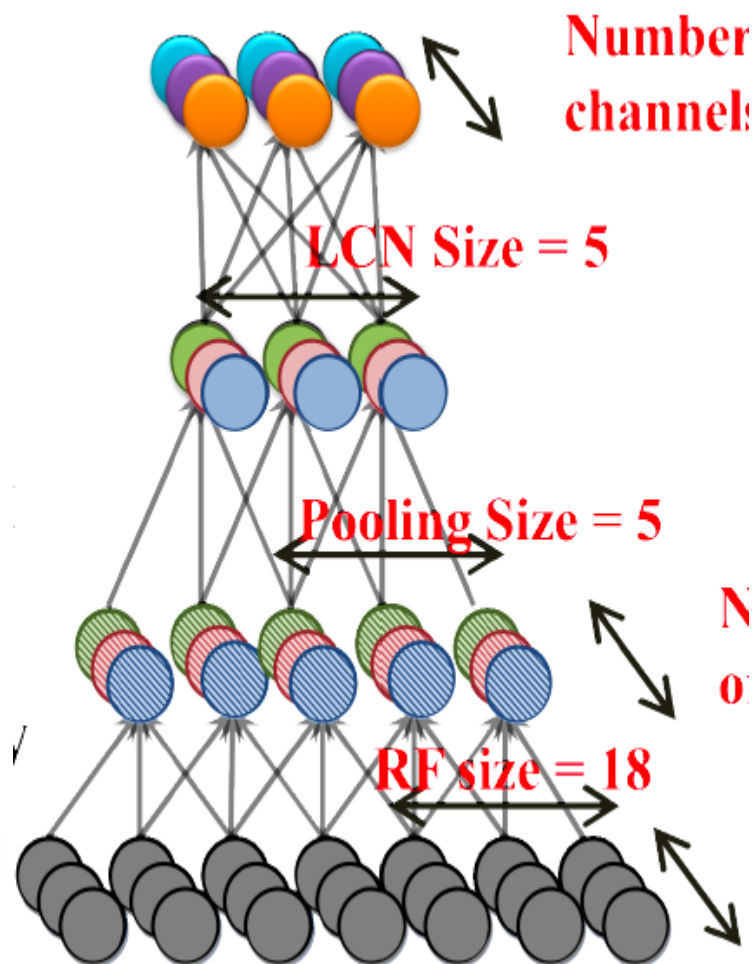
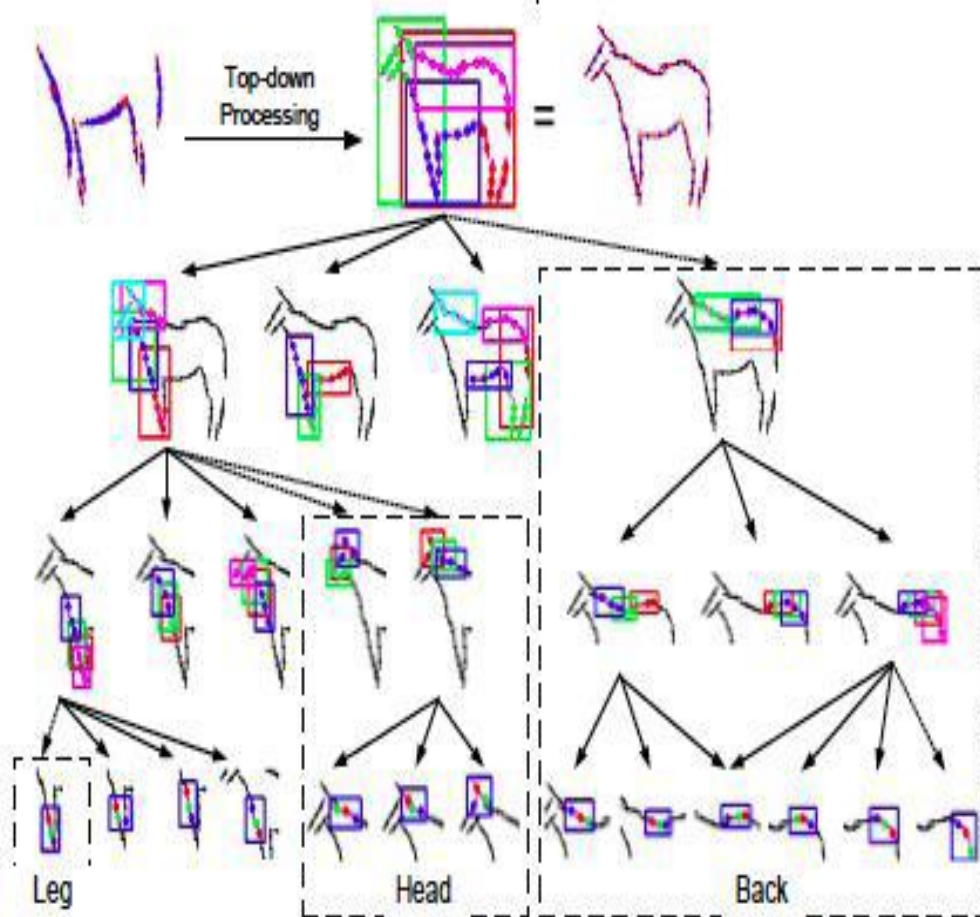


卷积神经网络

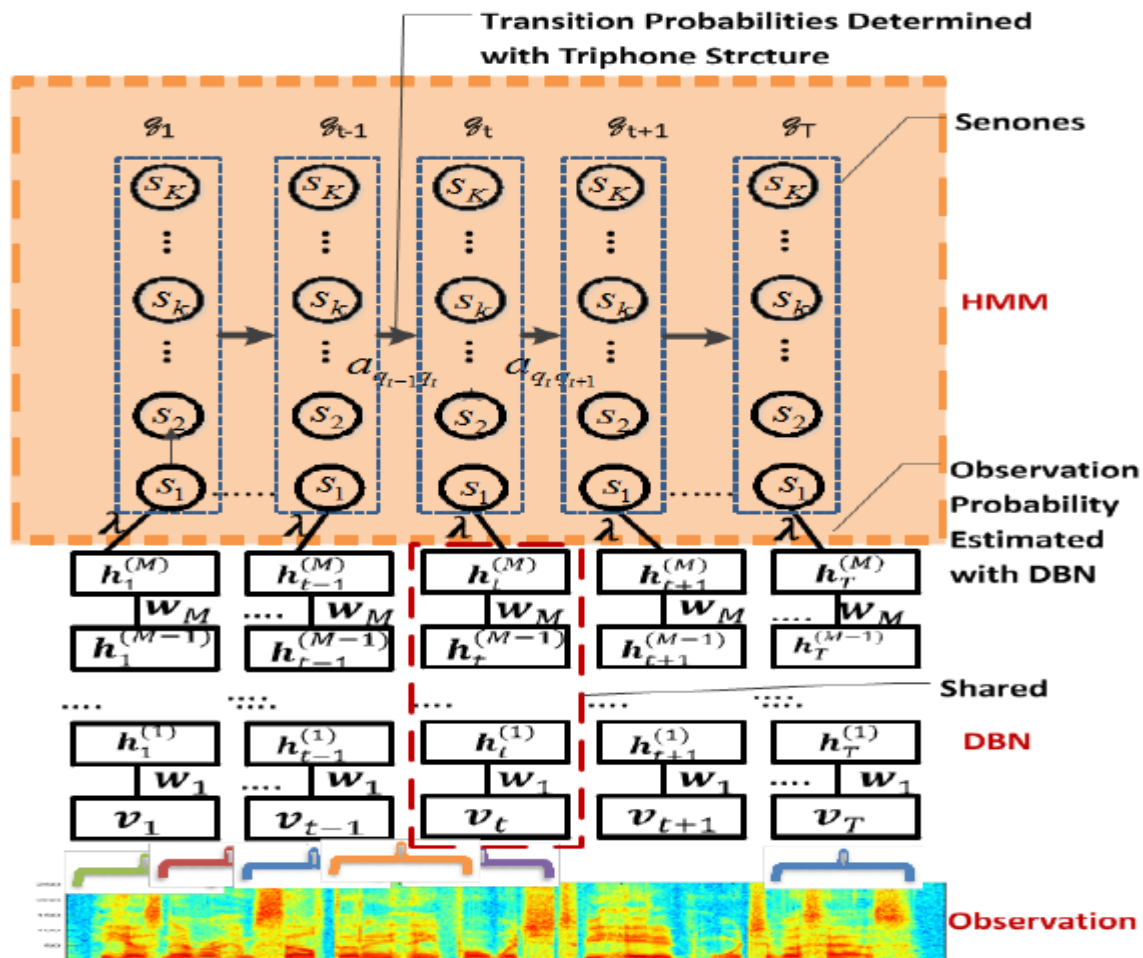


- C层是一个卷积层， S层是一个下采样层，下一C层起到二次特征提取的作用。

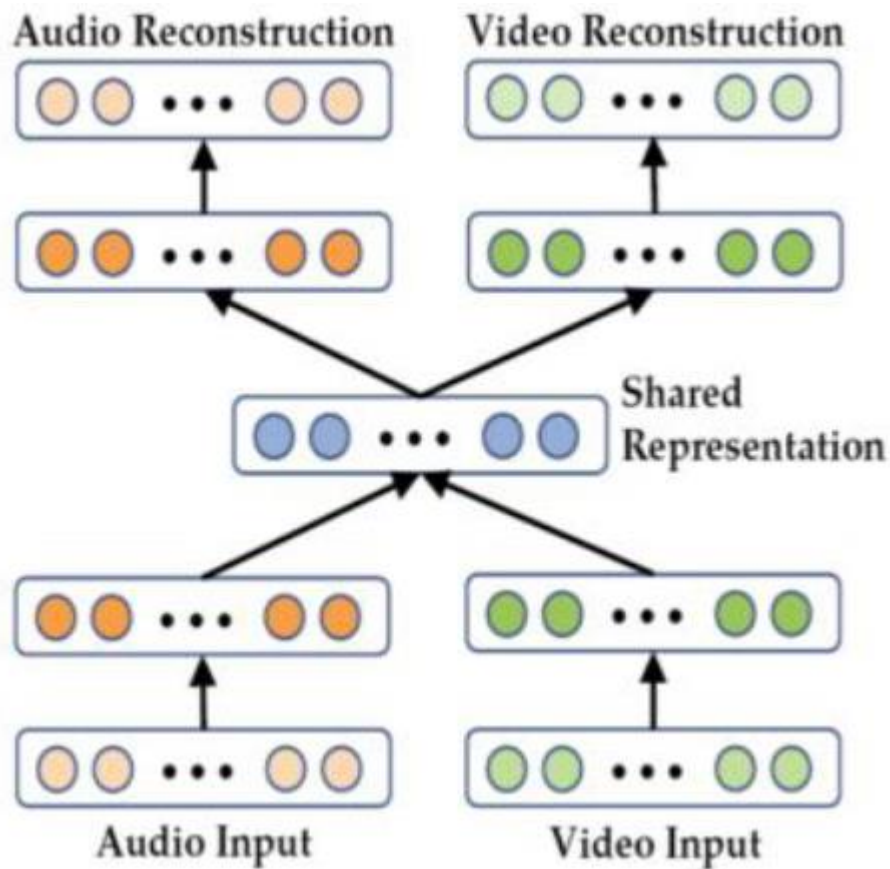
图像识别



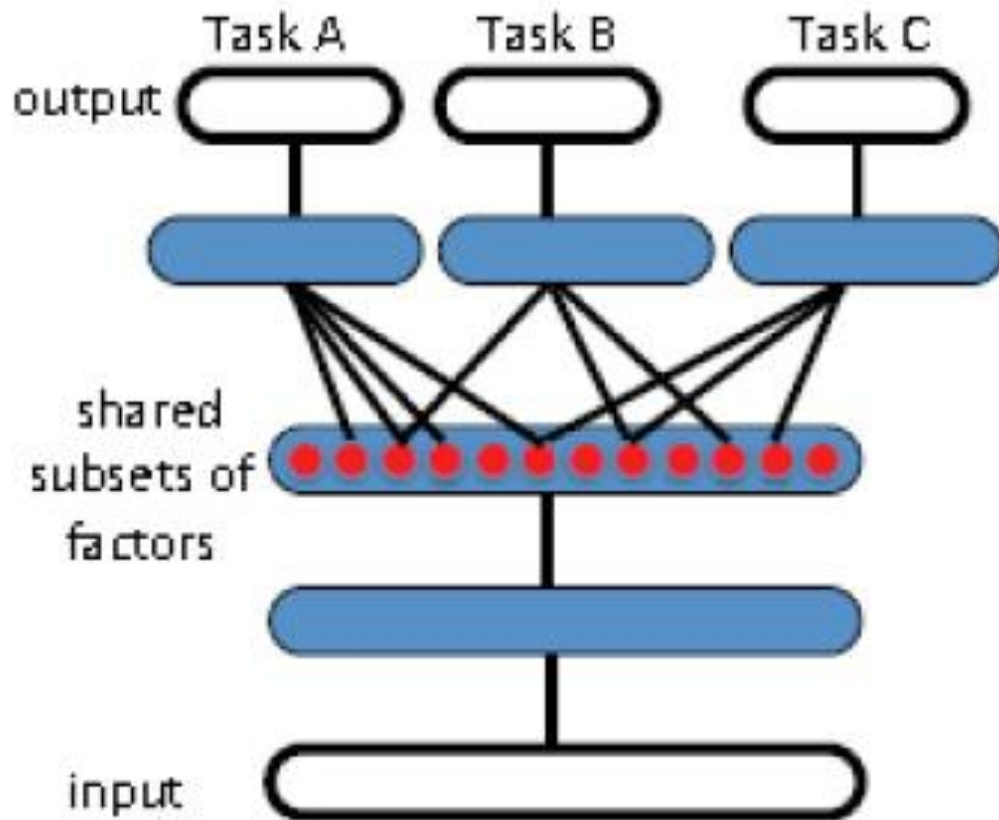
语音识别



声音与视频重构



多任务学习



内容提要



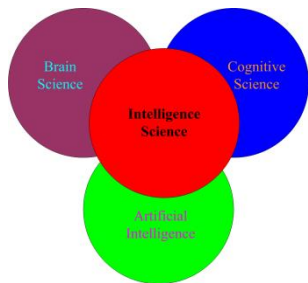
引言

不确定性推理

统计学习

大数据分析

展望



Big Data



UB LIBRARIES

Nature

September, 2008

“Big Data:
Science in the
Petabyte Era”

Source:
<http://www.nature.com/news/specials/bigdata/index.html>

4 September 2008 www.nature.com/nature £70 THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

nature

THE BITER BIT
Viral infections for viruses

TROPICAL CYCLONES
The strong get stronger

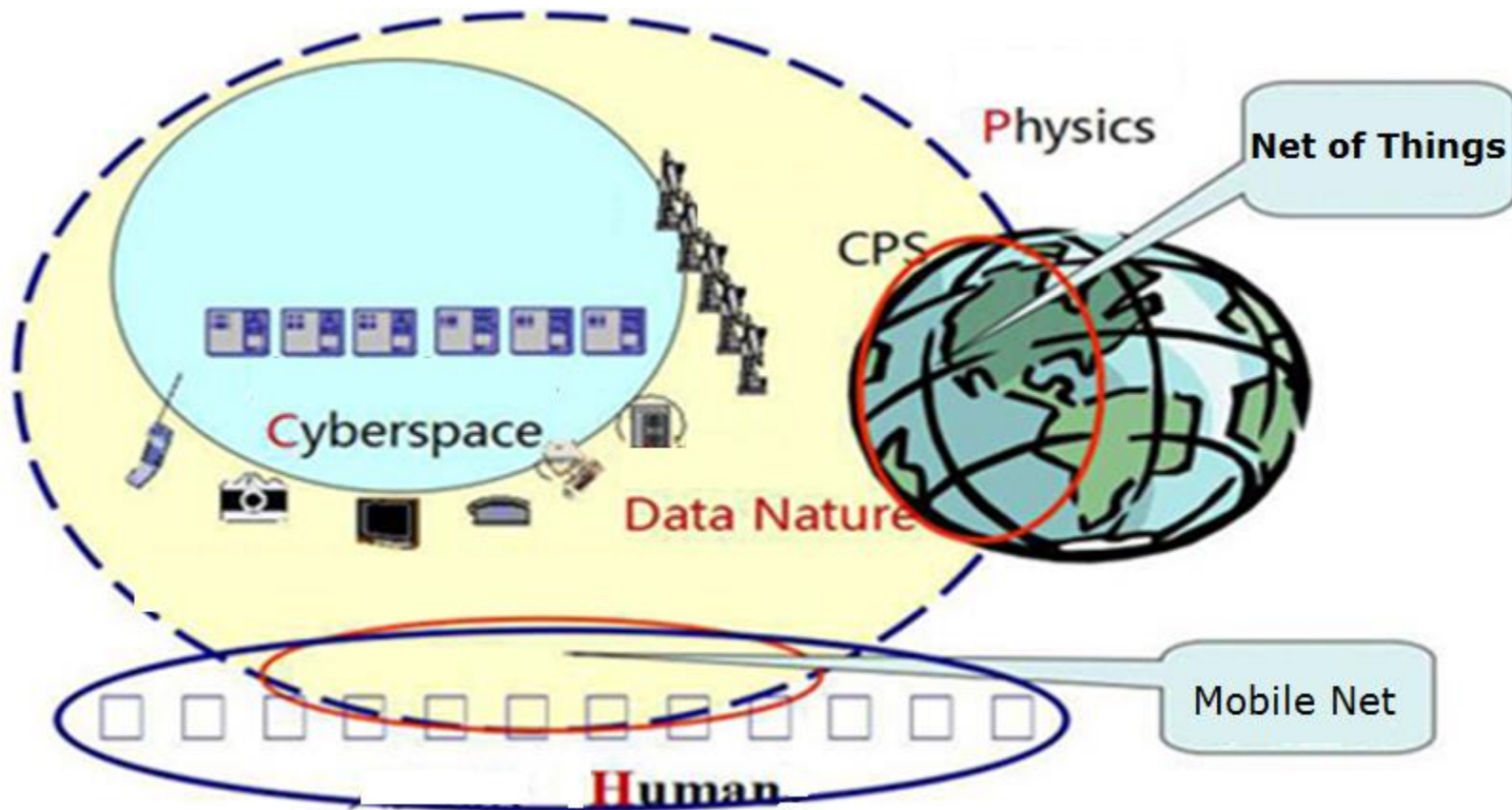
BLACK HOLE PHYSICS
A new window on the
Galactic Centre

BIG DATA

NATUREJOBS
Minnesota musings

**SCIENCE IN THE
PETABYTE ERA**

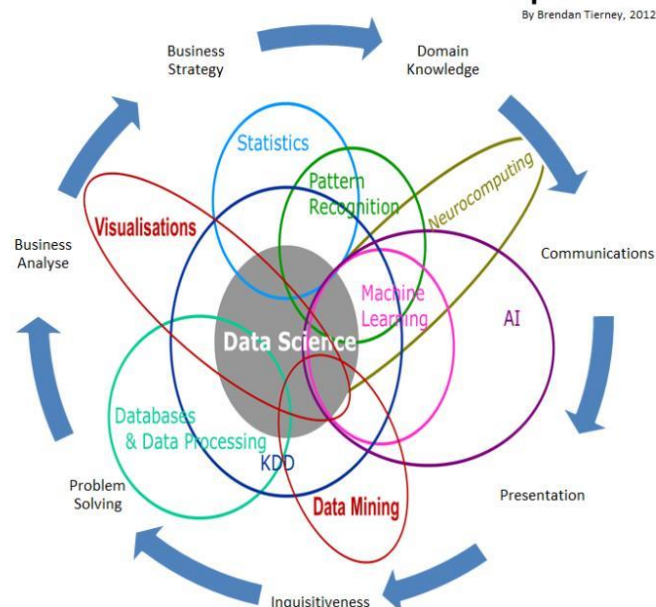
Big Data Nature CPH



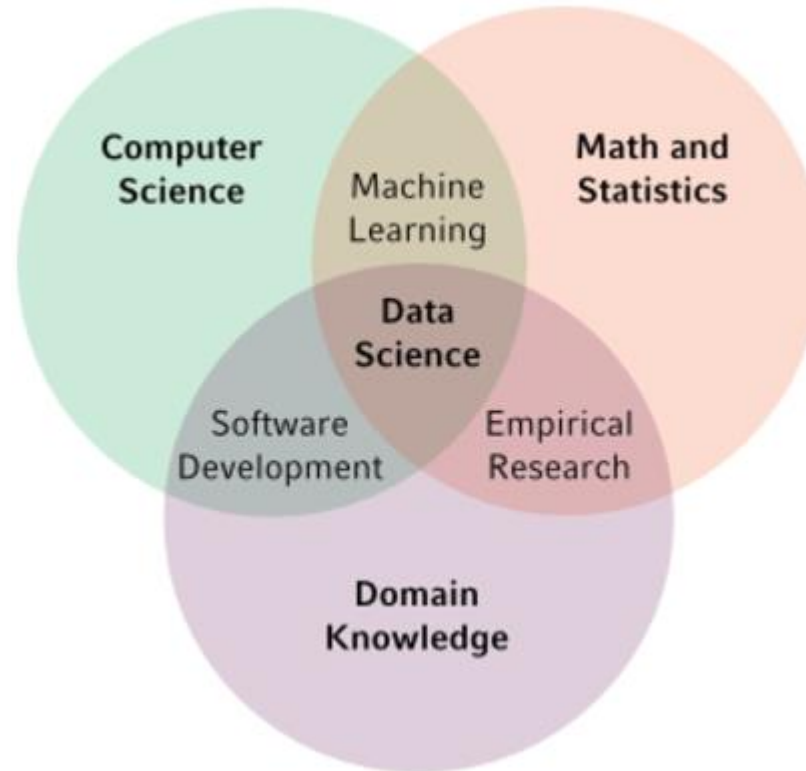
Data Science

Data Science is the art of converting raw data to useful information that can be used to draw conclusions and make decisions. For this goal a data scientist needs a deep understanding of techniques and methodologies from different disciplines in order to capture the full potential that is hidden in the data.

Data Science Is Multidisciplinary



Data Science



Big Data Values

Perception World



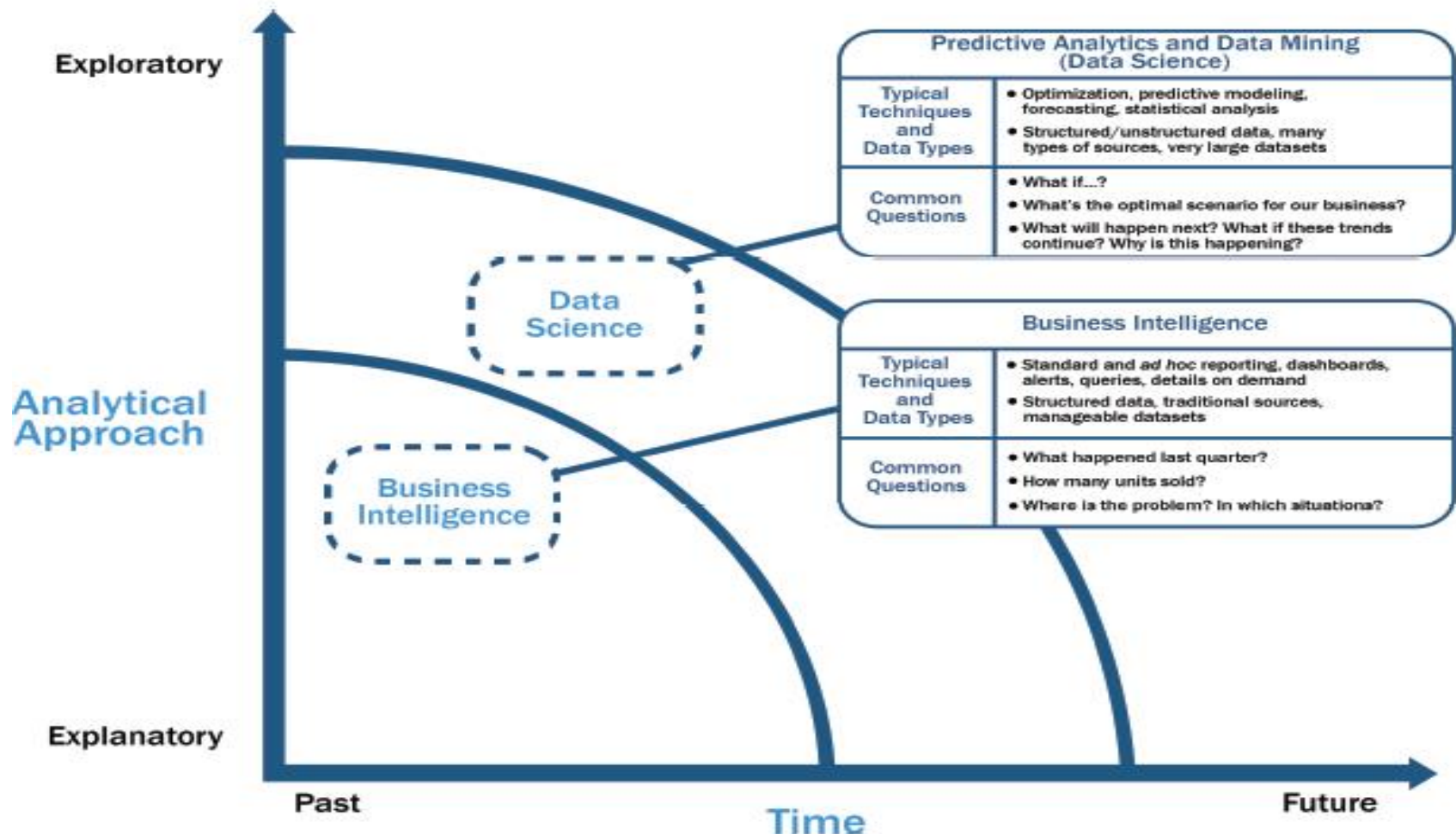
Predict Future



Decision and Adjust

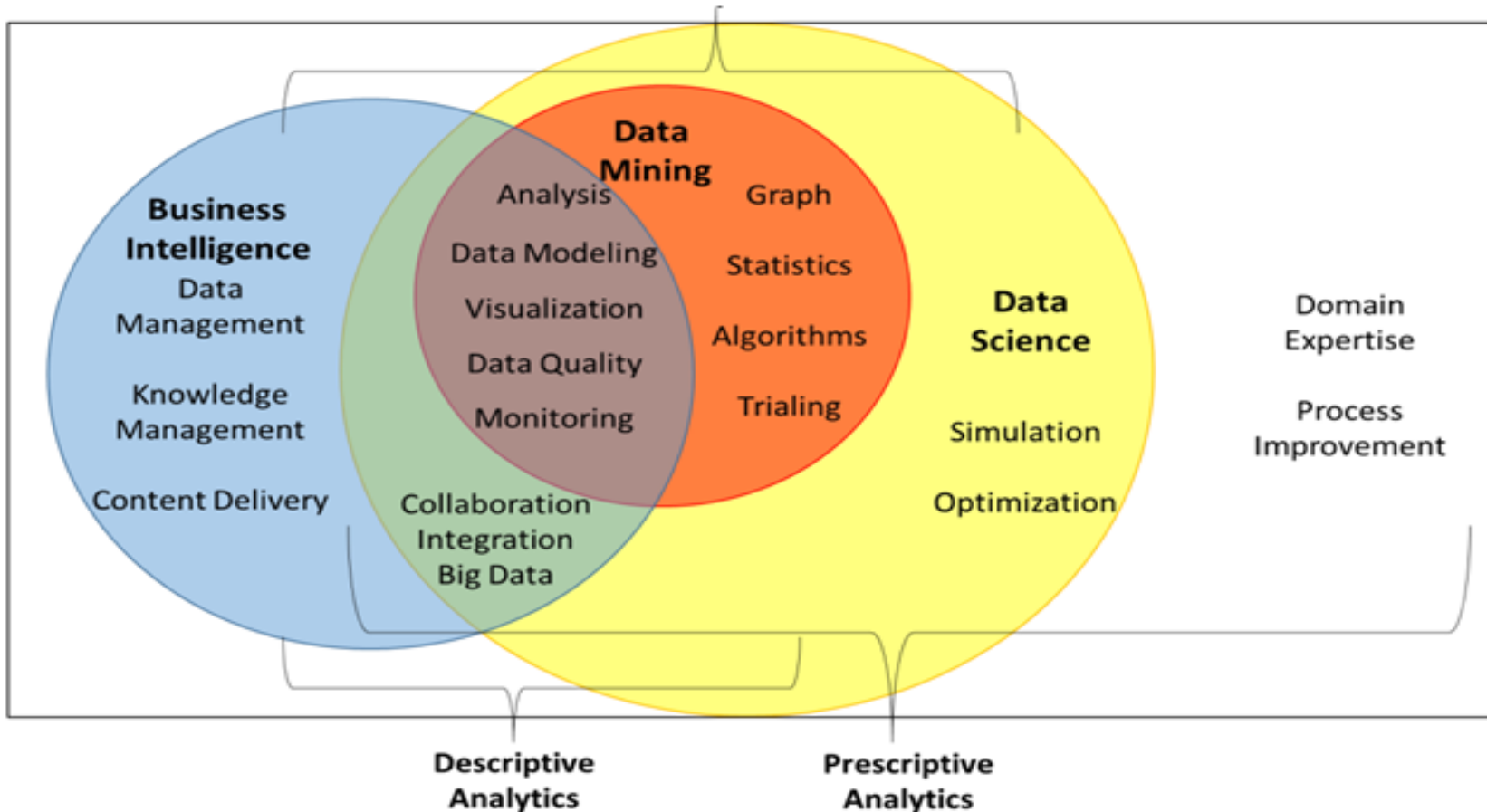


Data Science and Business Intelligence

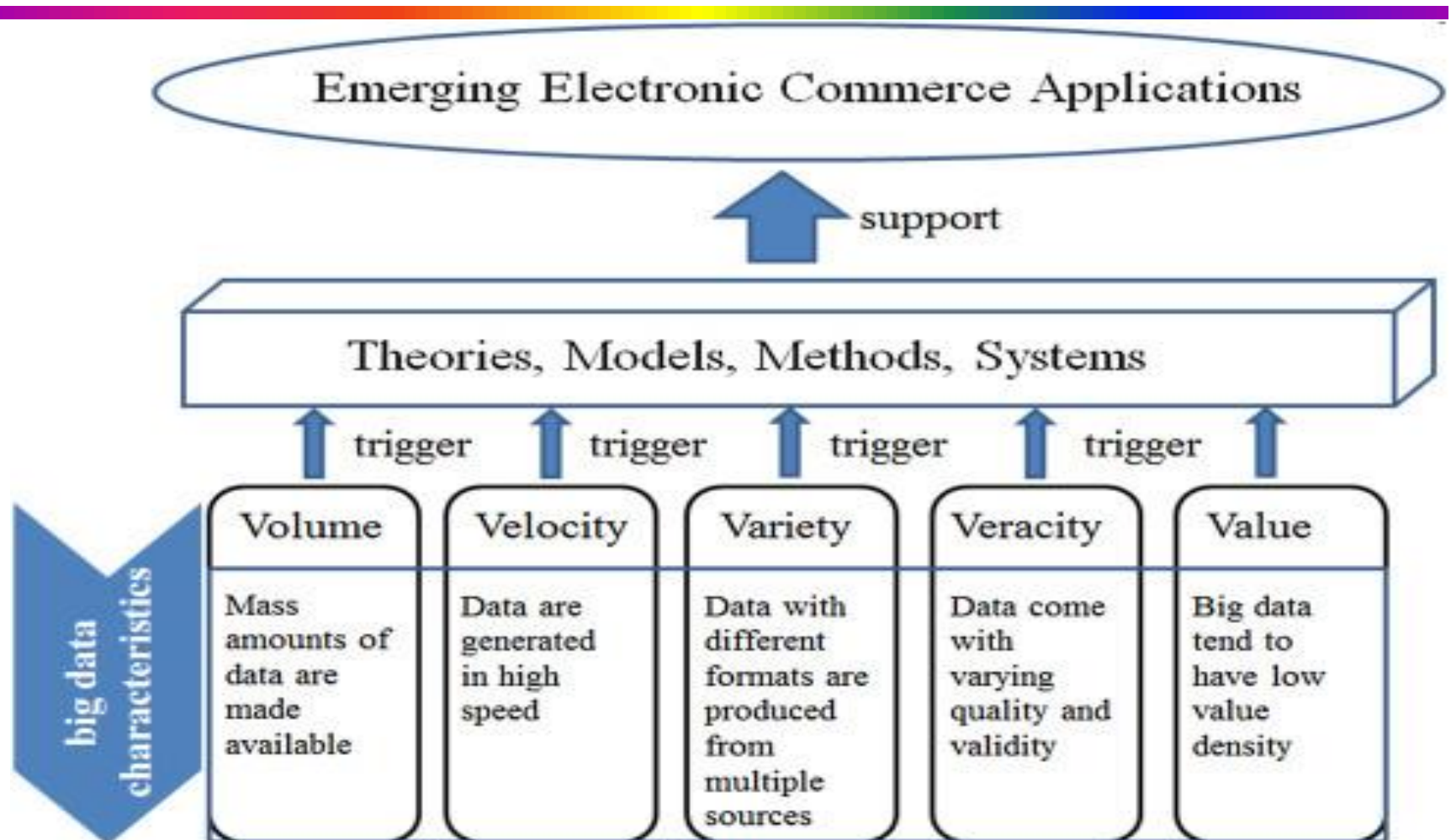


Business Intelligence

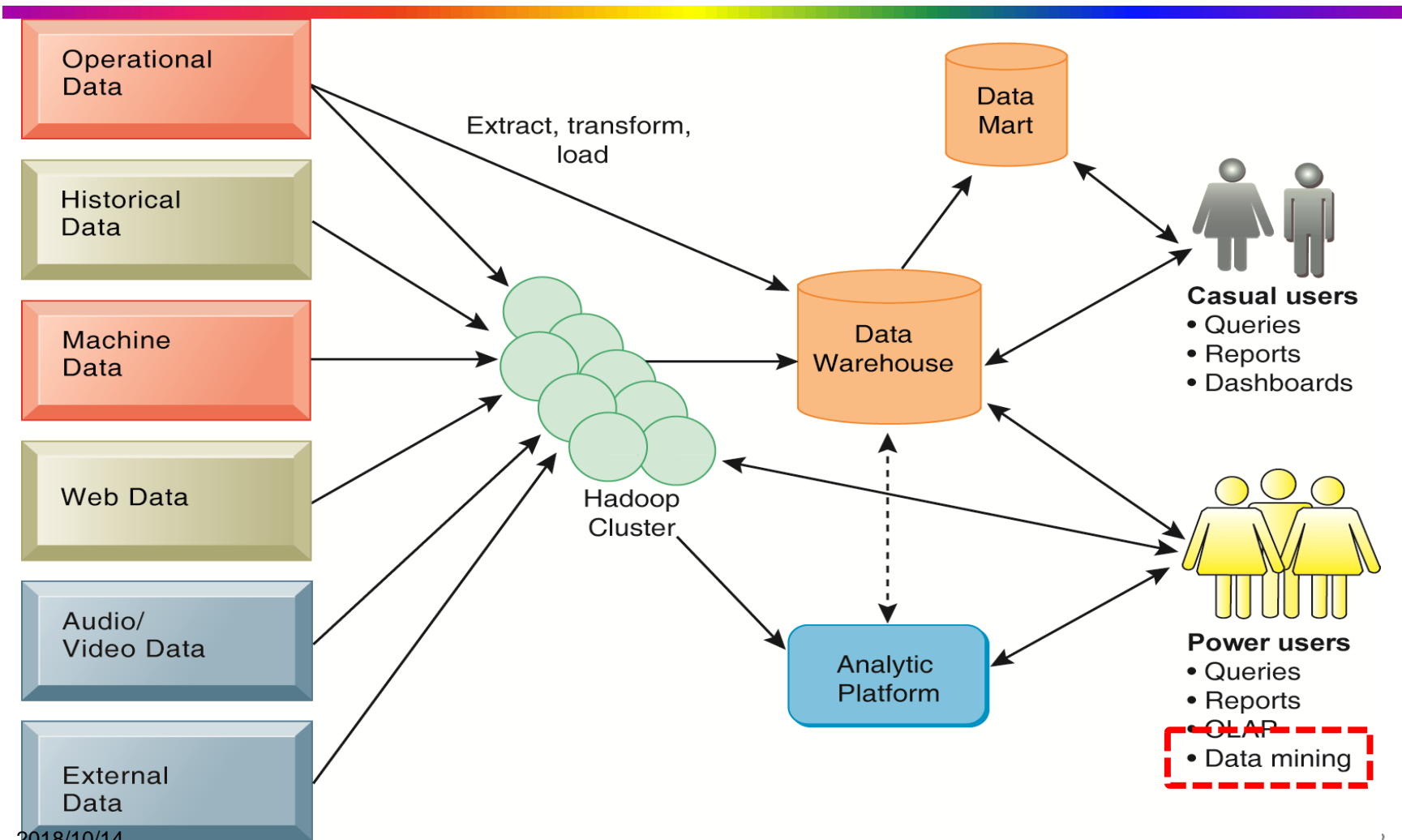
Business Intelligence Predictive Analytics Decision Management



Big Data Commerce



Business Intelligence Infrastructure



2018/10/14

110

CBDME: Cloud-Based Big Data Mining Engine



In the cloud computing environment comprehensive utilization of big data mining, parallel distributed processing, autonomous computing, service computing and other aspects to achieve cross-domain, heterogeneous, dynamic big data integration, analysis, mining, to meet the practical needs of the business and industry.

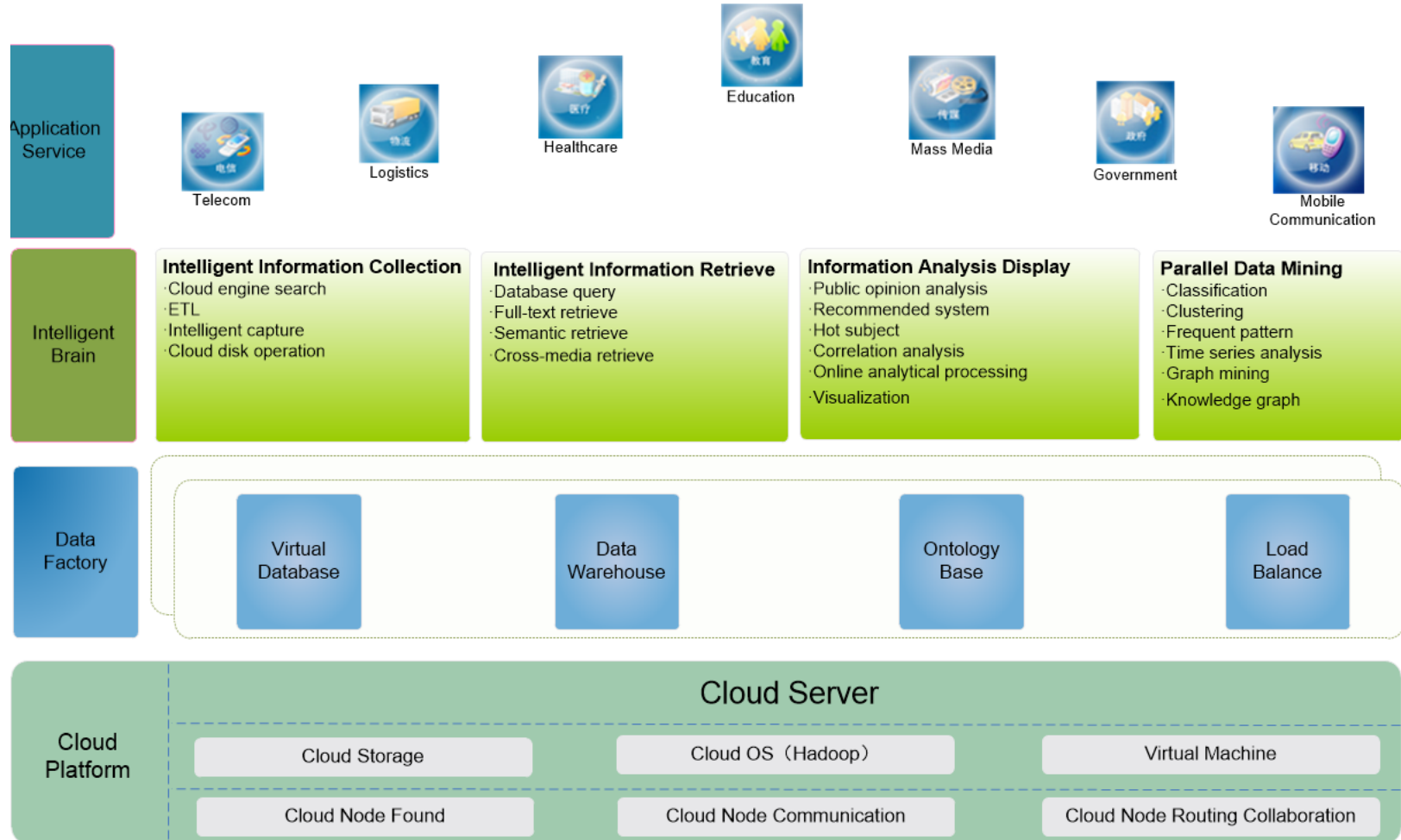
Cloud-Based Big Data Mining Engine-CBDME



CBDME system overall structure is divided into four levels:

- Cloud platform
- Data factory
- Intelligent brain
- Application services.

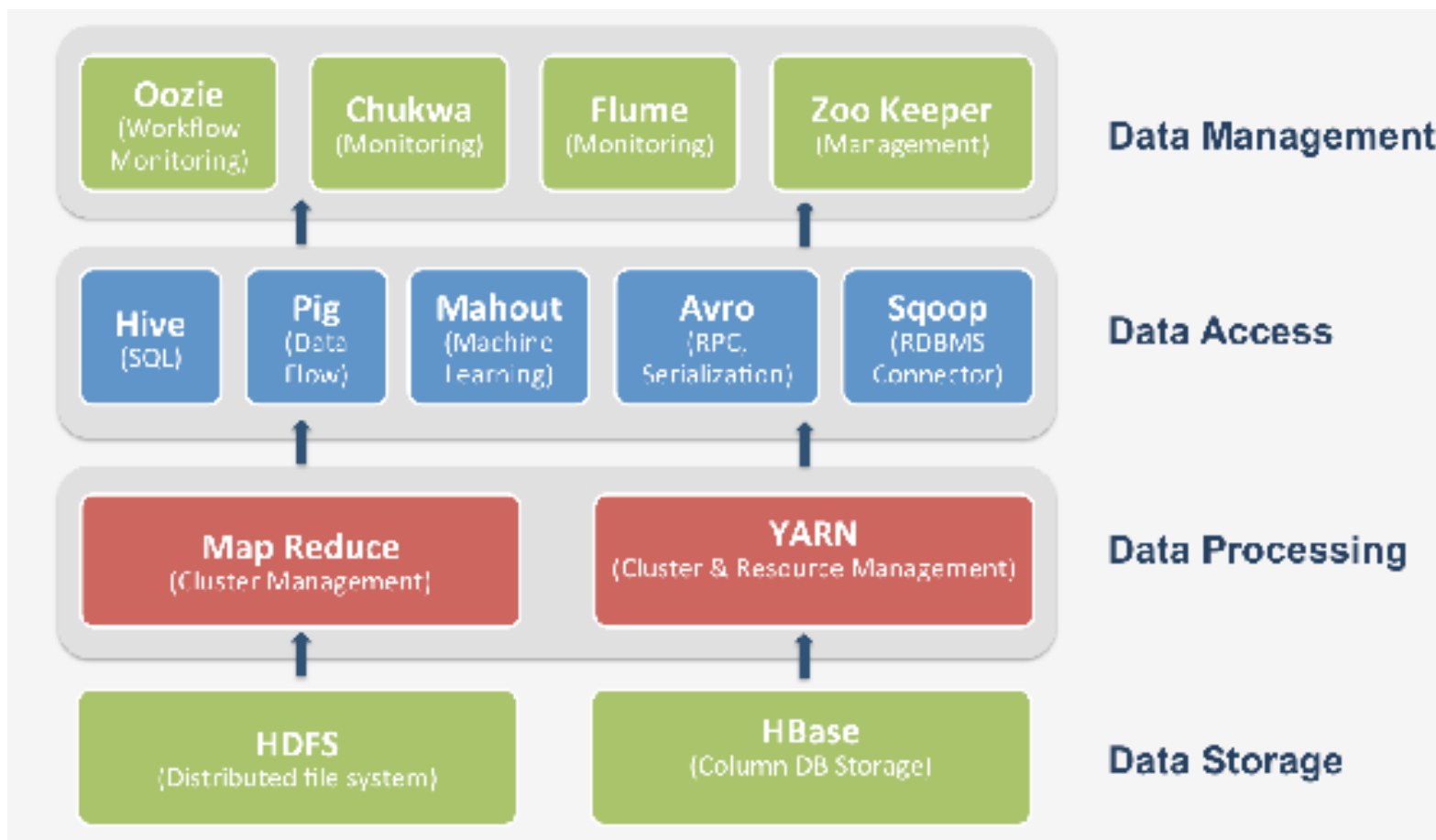
CBDME Architecture



CBDME



Hadoop Ecosystem



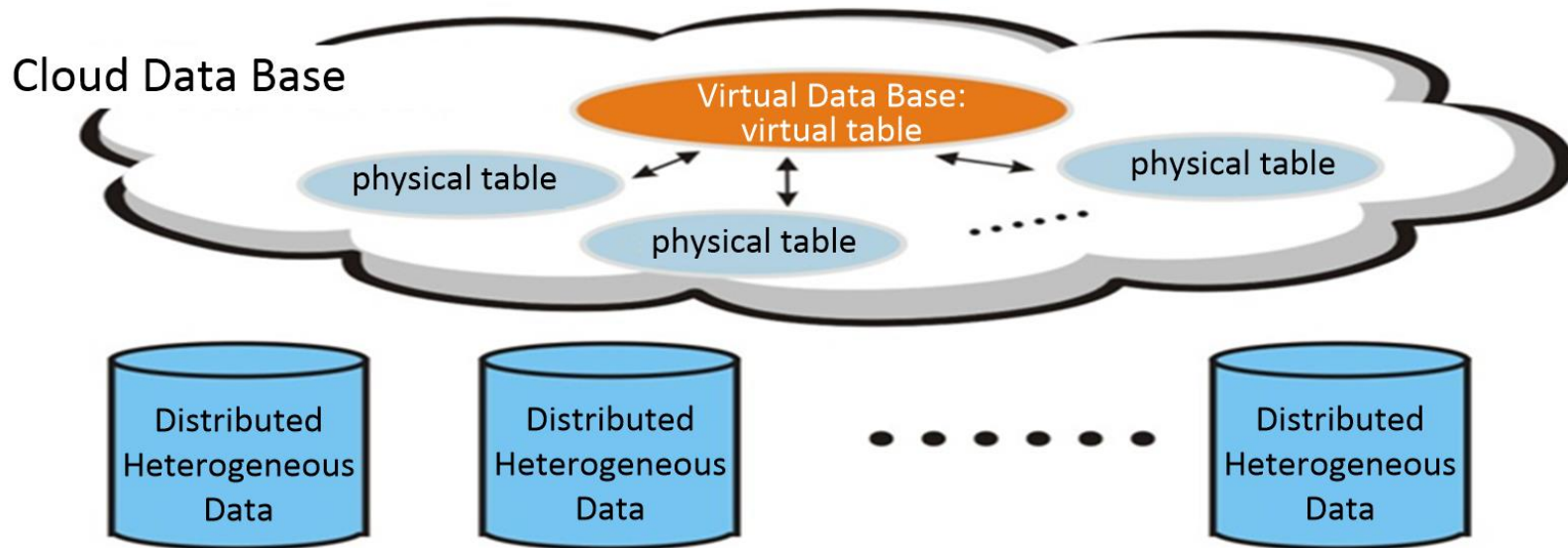
Cloudera Enterprise



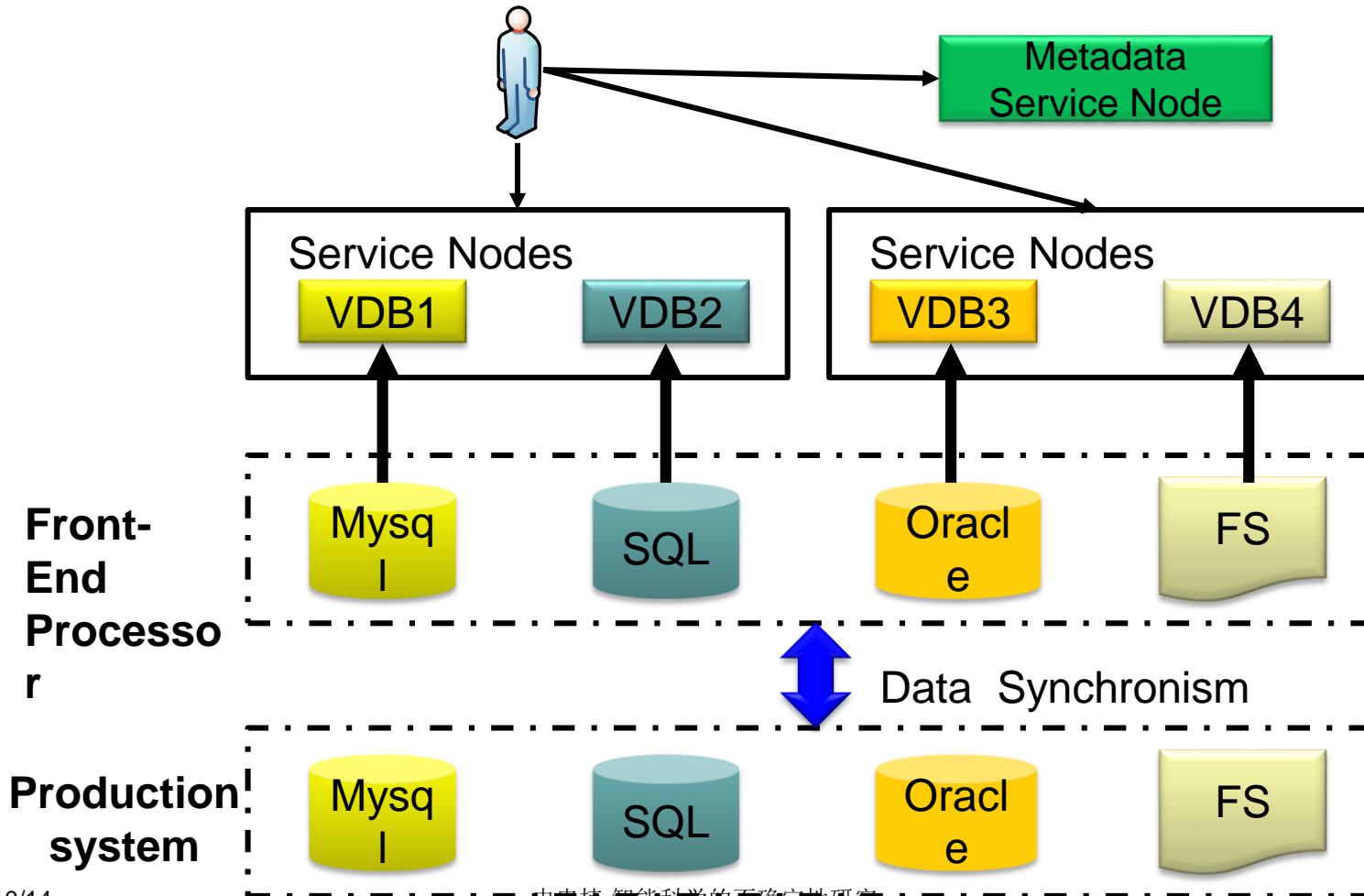
Virtual Database



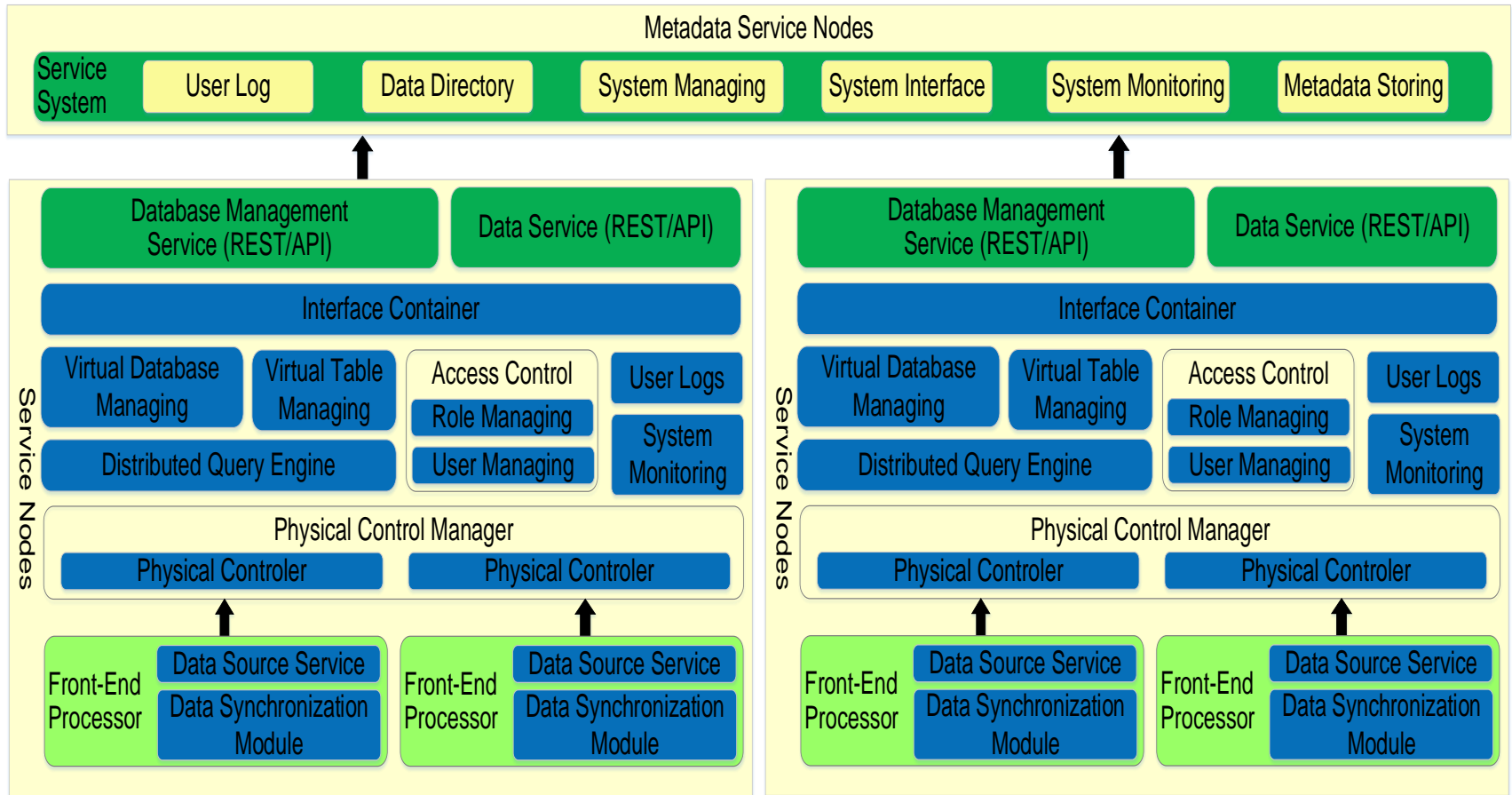
Uniform Interface of Data Standard and Data Access



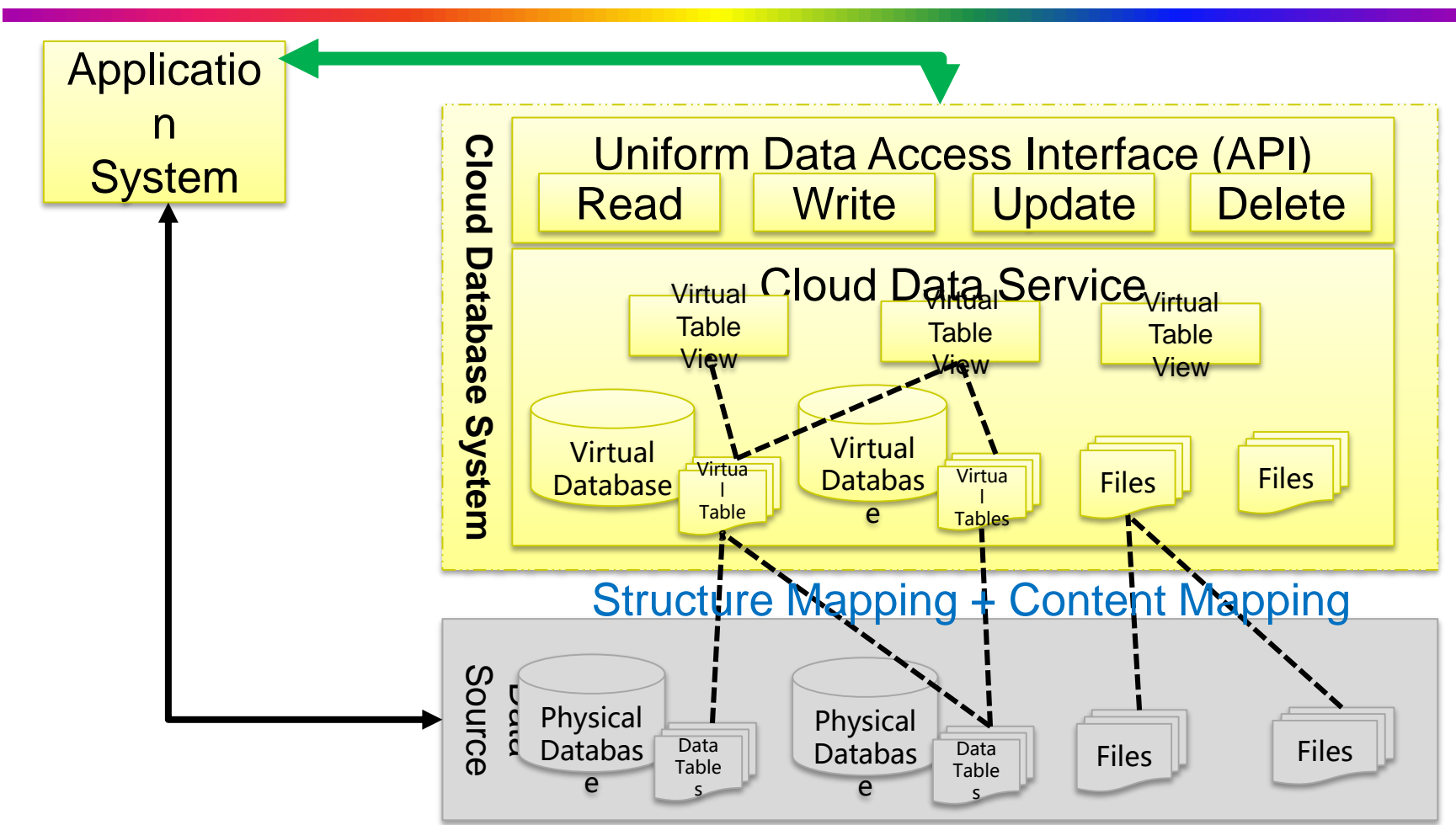
Physical Schema



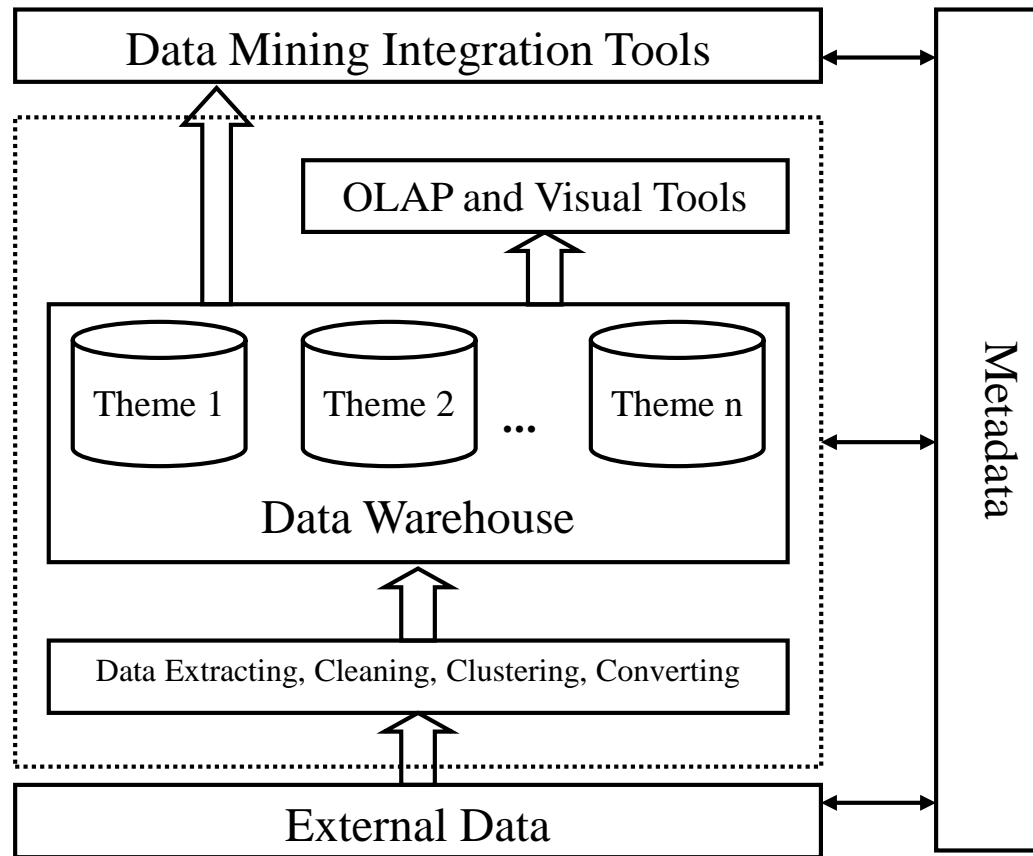
Logic Framework



Workflow



Data Warehouse Framework

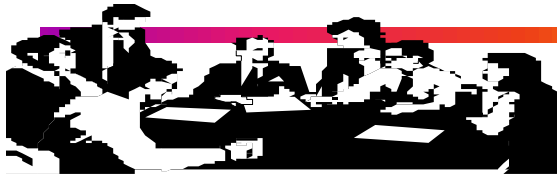


MSMiner Data Warehouse Structure Sketch

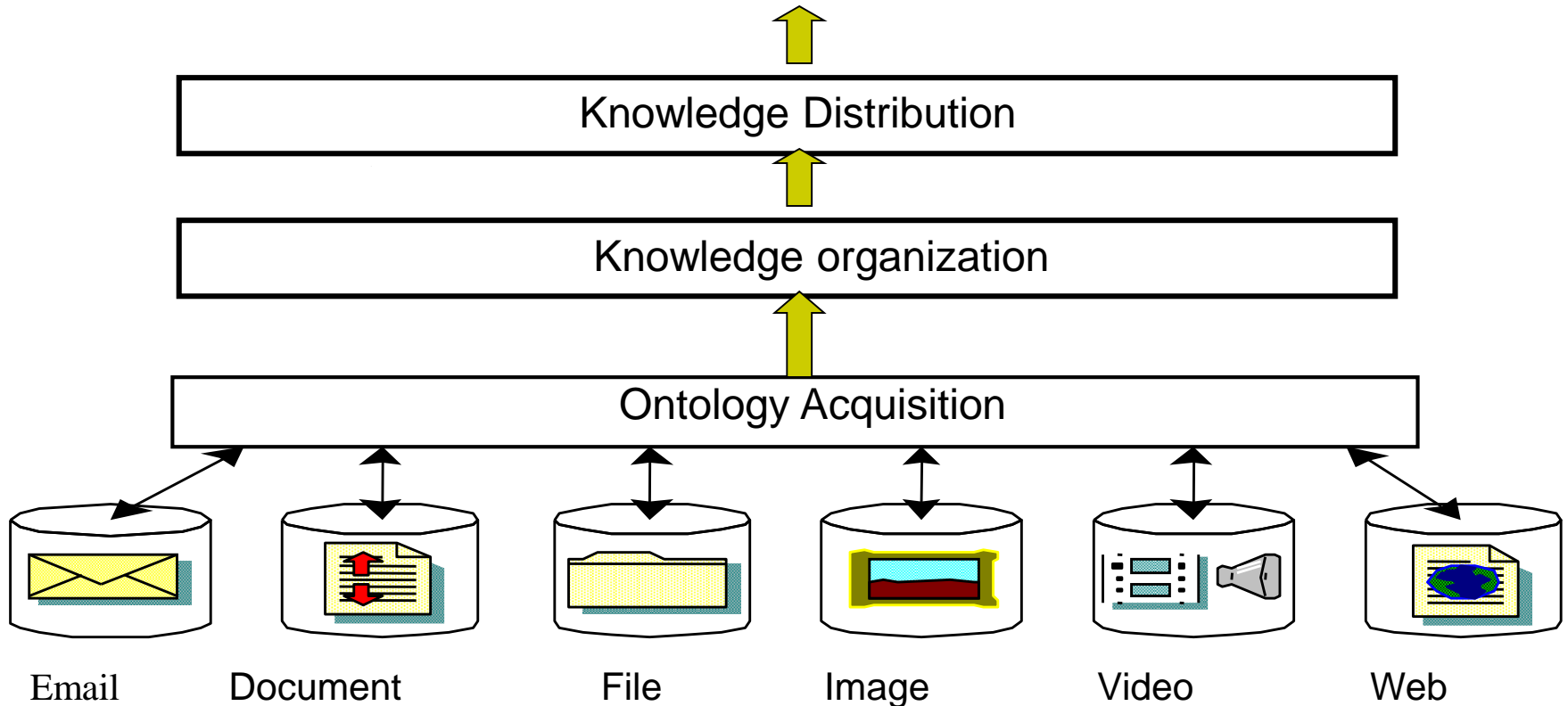
Ontology-Based Knowledge Management



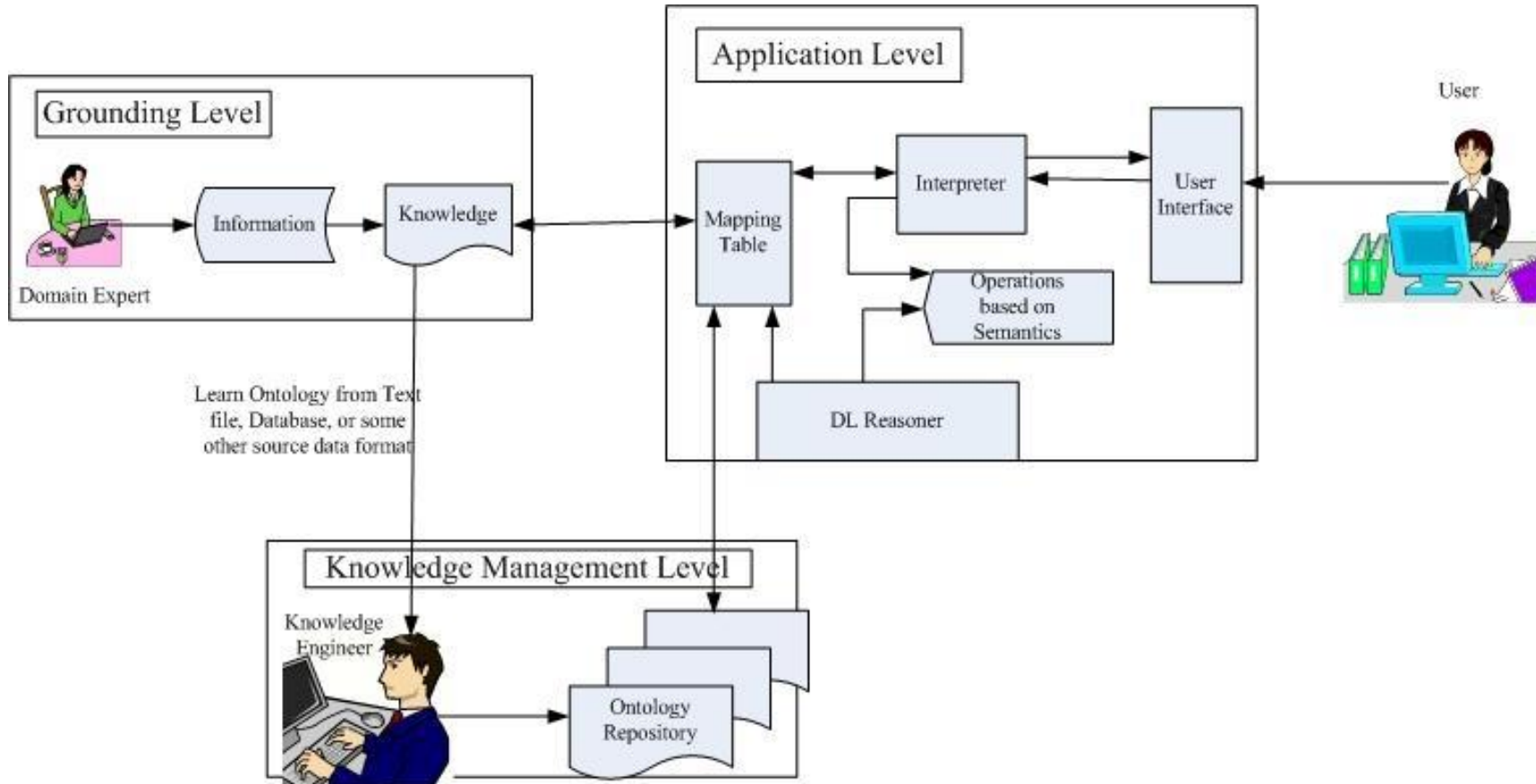
中科院计算所
INSTITUTE OF COMPUTING
TECHNOLOGY



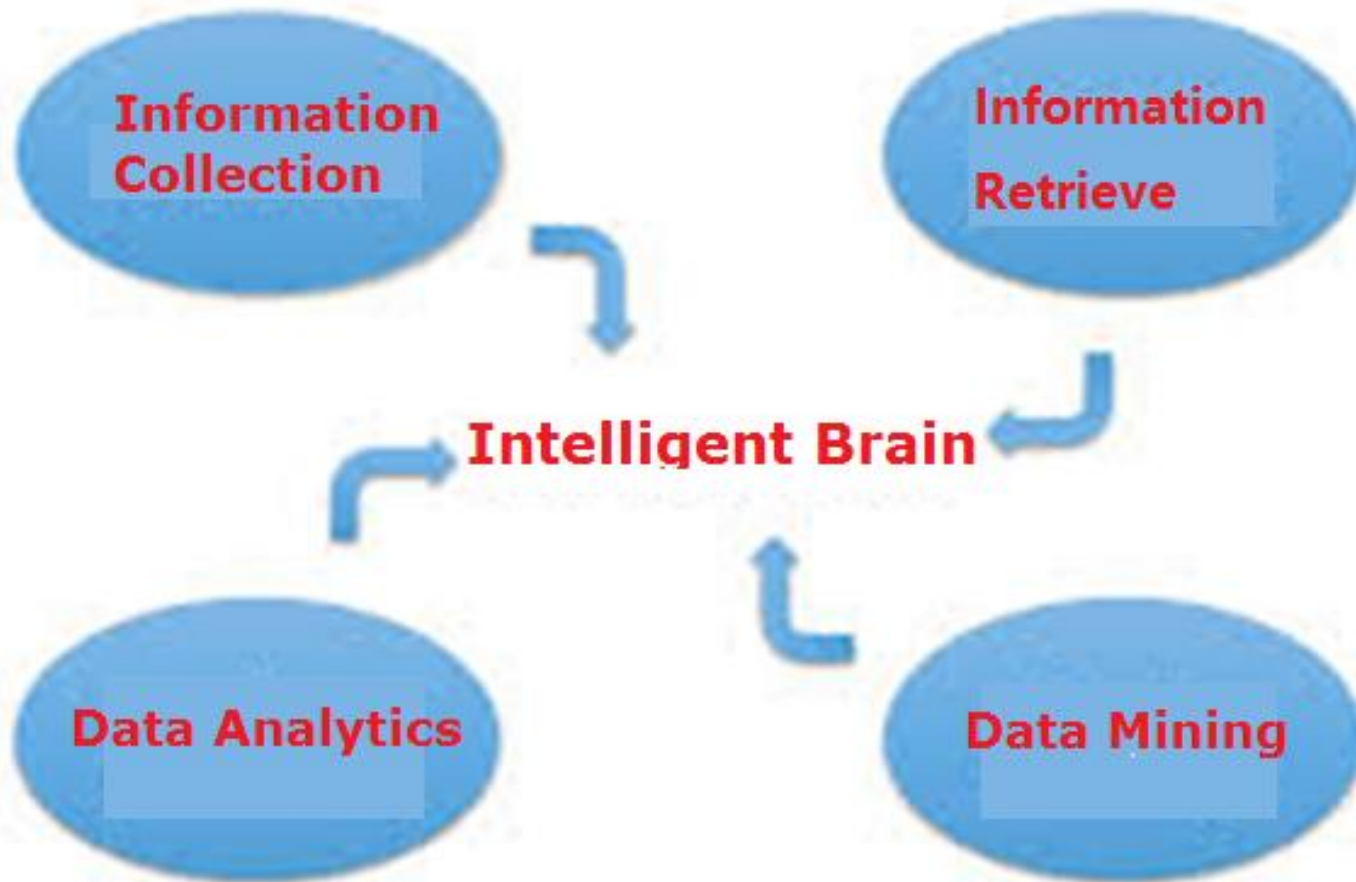
Knowledge
Application



KMSphere System



Intelligent Brain



Intelligent Brain

Intelligent Information Collection

- Cloud engine search
- ETL
- Intelligent capture
- Cloud disk operation

Intelligent Brain

Intelligent Information Retrieve

- Database query
- Full-text retrieve
- Semantic retrieve
- Cross-media retrieve

Intelligent Brain

Data Analytics

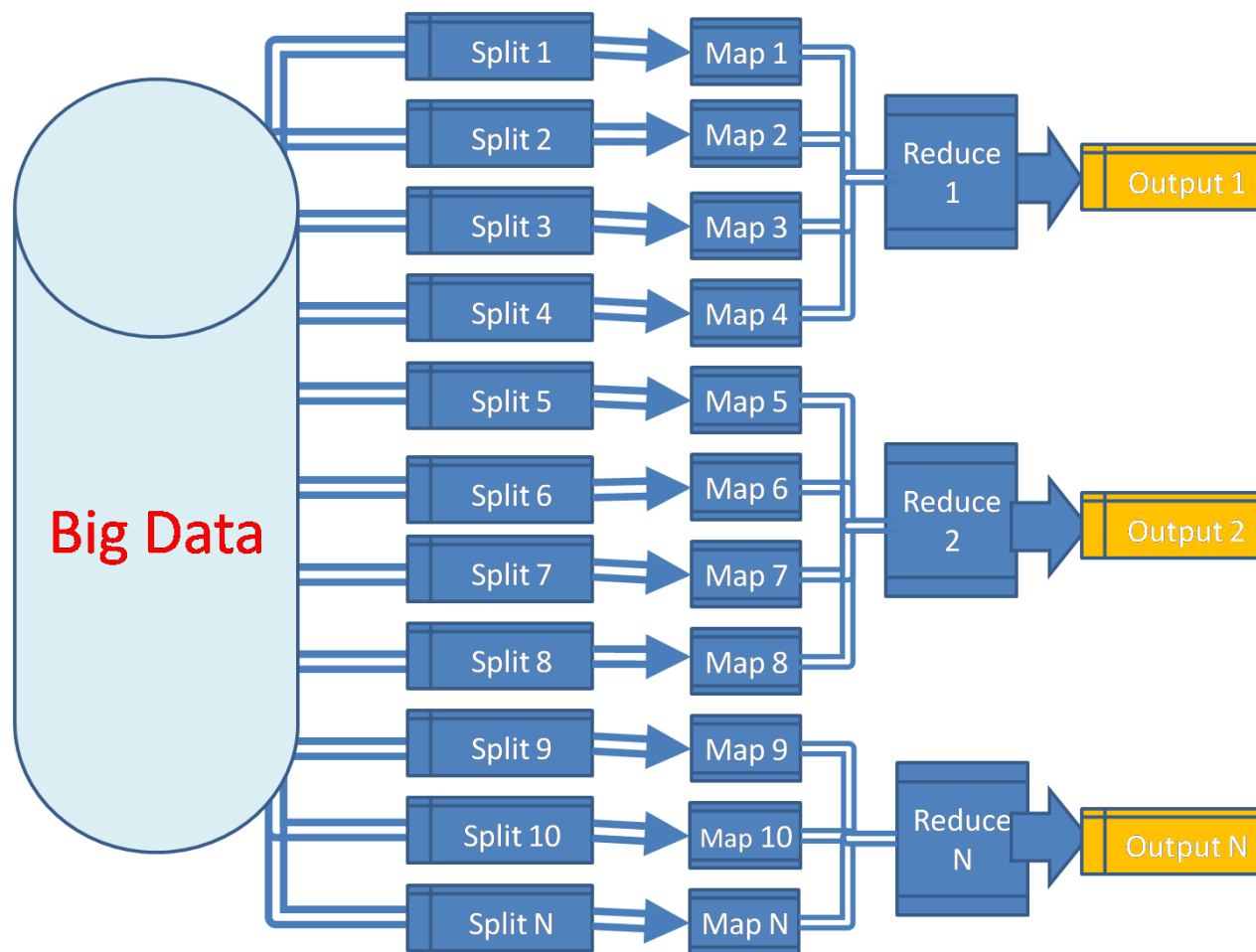
- Public opinion analysis
- Recommended system
- Hot subject
- Correlation analysis
- Online analytical processing
- Visualization

Intelligent Brain

Parallel Data Mining

- Classification
- Clustering
- Frequent pattern
- Time series analysis
- Graph mining
- Knowledge graph

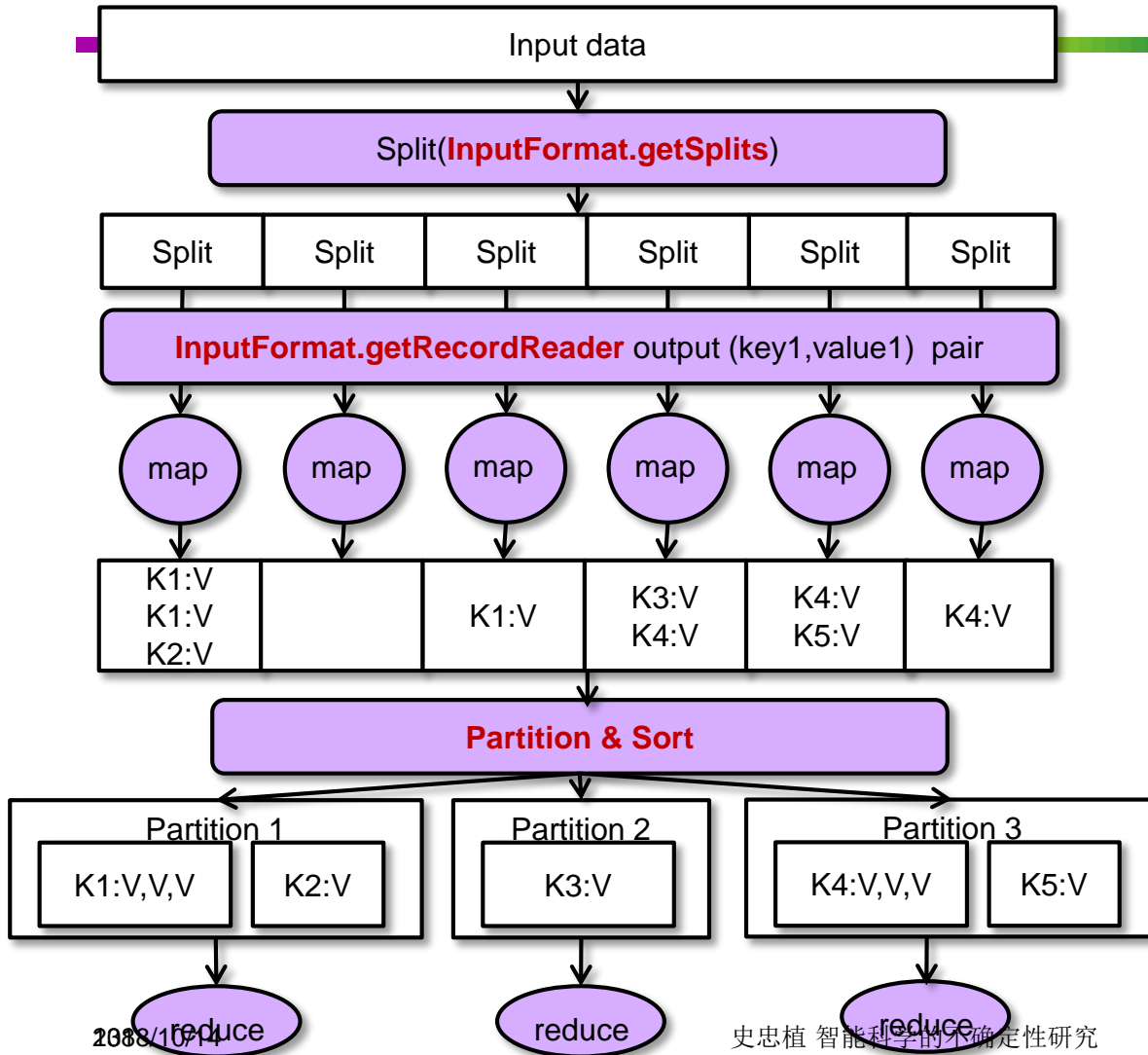
Map/Reduce



Map/Reduce



Data Flow Diagram



(1) `InputFormat.getSplits()` is responsible for creating splits from input data;

(2) `InputFormat.getRecordReader()` divides each split into a set of (key1,value1) pairs;

(3) `Map` function processes each (key1,value1) in turn and produces a set of intermediate (key2,value2) pairs;

(4) `MapReduce` partitions all intermediate (key2,value2) pairs. Each partition is grouped and sorted by intermediate key key2;

(5) `Reduce` function processes each partition and output a smaller set of (key3,value3) pairs.

Cross-Domain CCA



The algorithm preserves the inter-domain correlation and maps the data of each domain from the original high-dimensional feature space to the low-dimensional feature space. In this low-dimensional space, the data between the domains have a similar distribution.

$$\max_{W_A, W_{B_s}, W_{\Gamma_t}} \frac{W_A^T C_{A_s B_s} W_{B_s}}{\sqrt{W_A^T C_{A_s A_s} W_A \cdot W_{B_s}^T C_{B_s B_s} W_{B_s}}} + \frac{W_A^T C_{A_t \Gamma_t} W_{\Gamma_t}}{\sqrt{W_A^T C_{A_t A_t} W_A \cdot W_{\Gamma_t}^T C_{\Gamma_t \Gamma_t} W_{\Gamma_t}}}$$

Where:

$$A_s = [\alpha_1, \alpha_2, \dots, \alpha_{n_s}] \in \mathbb{R}^{|\mathcal{X}_n| \times n_s}$$

$$B_s = [\beta_1, \beta_2, \dots, \beta_{n_s}] \in \mathbb{R}^{|\mathcal{X}_s - \mathcal{X}_n| \times n_s}$$

$$C_{A_s A_s} = A_s A_s^T \in \mathbb{R}^{|\mathcal{X}_n| \times |\mathcal{X}_n|}$$

$$C_{B_s B_s} = B_s B_s^T \in \mathbb{R}^{|\mathcal{X}_s - \mathcal{X}_n| \times |\mathcal{X}_s - \mathcal{X}_n|}$$

$$C_{A_s B_s} = A_s B_s^T \in \mathbb{R}^{|\mathcal{X}_n| \times |\mathcal{X}_s - \mathcal{X}_n|} = C_{B_s A_s}^T$$

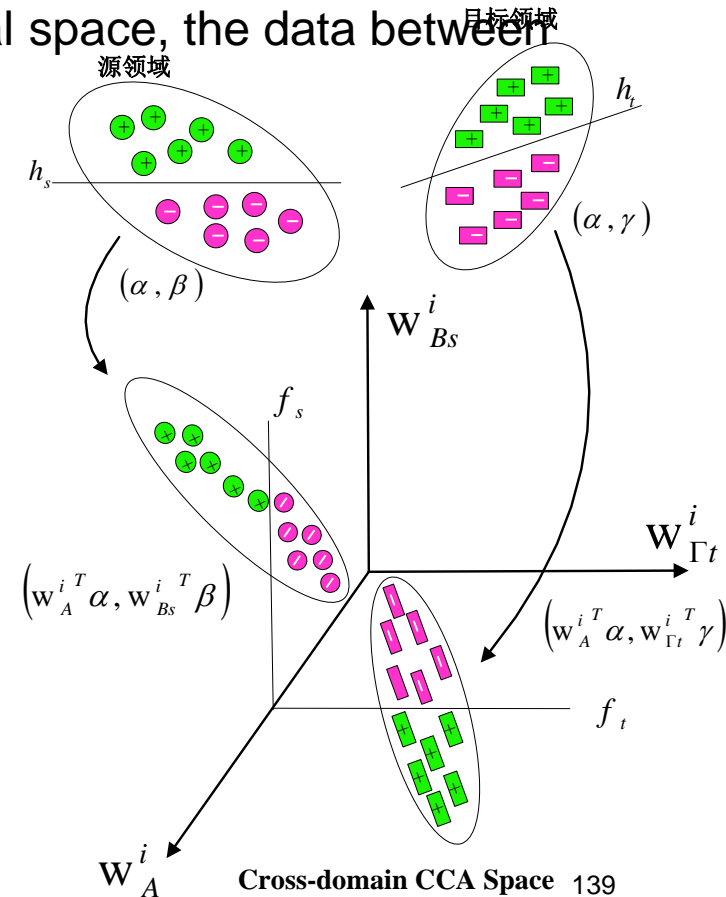
$$A_t = [\alpha_1, \alpha_2, \dots, \alpha_l] \in \mathbb{R}^{|\mathcal{X}_n| \times l}$$

$$\Gamma_t = [\Gamma_1, \Gamma_2, \dots, \Gamma_l] \in \mathbb{R}^{|\mathcal{X}_t| \times l}$$

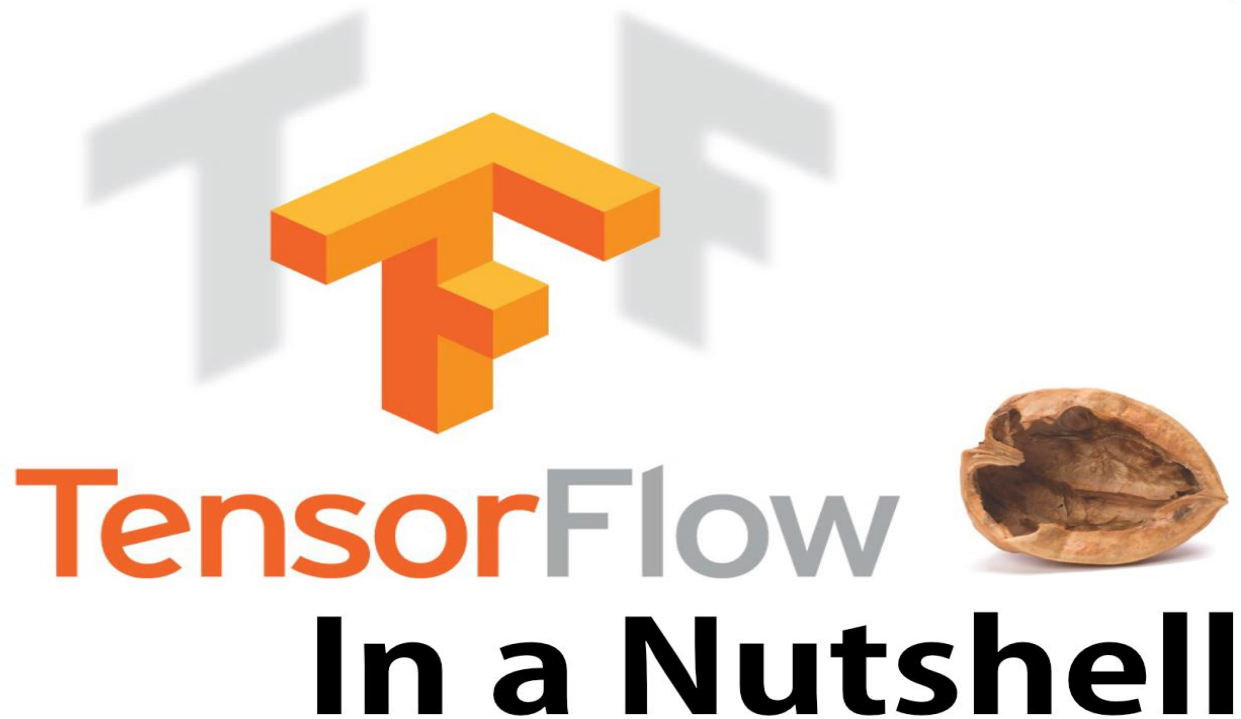
$$C_{A_t A_t} = A_t A_t^T \in \mathbb{R}^{|\mathcal{X}_n| \times |\mathcal{X}_n|}$$

$$C_{\Gamma_t \Gamma_t} = \Gamma_t \Gamma_t^T \in \mathbb{R}^{|\mathcal{X}_t| \times |\mathcal{X}_t|}$$

$$C_{A_t \Gamma_t} = A_t \Gamma_t^T \in \mathbb{R}^{|\mathcal{X}_n| \times |\mathcal{X}_t|} = C_{\Gamma_t A_t}^T$$



TensorFlow



Keras

Keras: Deep Learning library for Theano and TensorFlow

You have just found Keras.

Keras is a minimalist, highly modular neural networks library, written in Python and capable of running either on top of either [TensorFlow](#) or [Theano](#). It was developed with a focus on enabling fast experimentation. Being able to go from idea to result with the least possible delay is key to doing good research.

Use Keras if you need a deep learning library that:

- allows for easy and fast prototyping (through total modularity, minimalism, and extensibility).
- supports both convolutional networks and recurrent networks, as well as combinations of the two.
- supports arbitrary connectivity schemes (including multi-input and multi-output training).
- runs seamlessly on CPU and GPU.

Read the documentation at Keras.io.

Keras is compatible with: - **Python 2.7-3.5** with the Theano backend - **Python 2.7** with the TensorFlow backend

内容提要



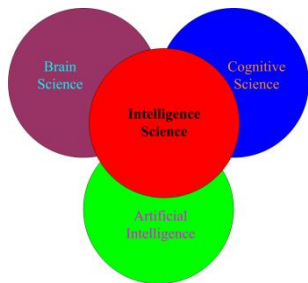
引言

不确定性推理

统计学习

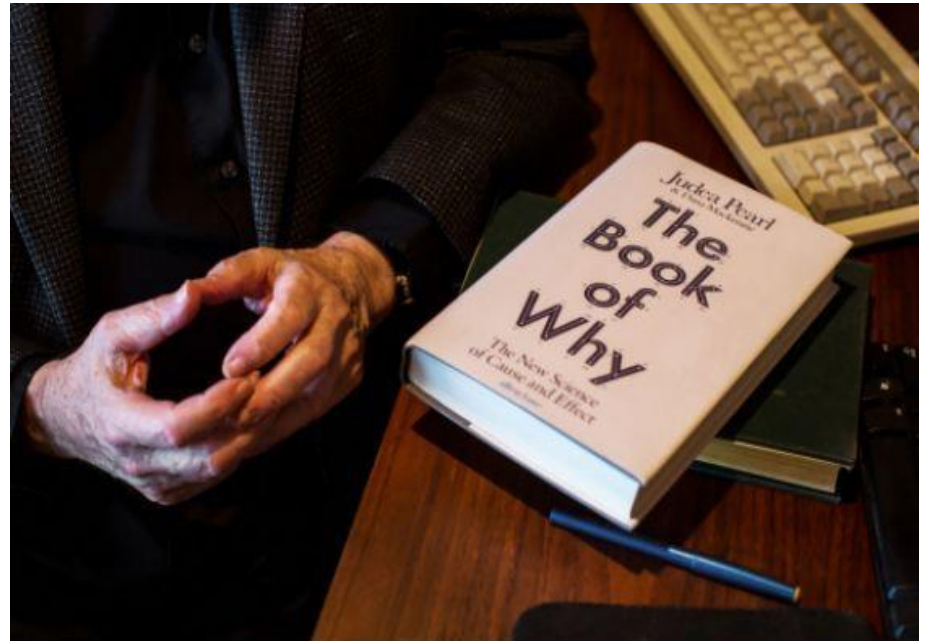
大数据分析

展望



因果推理

Pearl在新书中，详细阐述了智能机器如何真正思考的愿景。他认为，关键在于用因果推理来取代简单推理。机器不仅需要把发热和疟疾联系起来，而且需要能够推断疟疾为什么能引起发烧。一旦有了因果推理能力，机器就有可能查询某种干预而引起的因果关系如何改变 — Pearl将其视为科学思想的基石

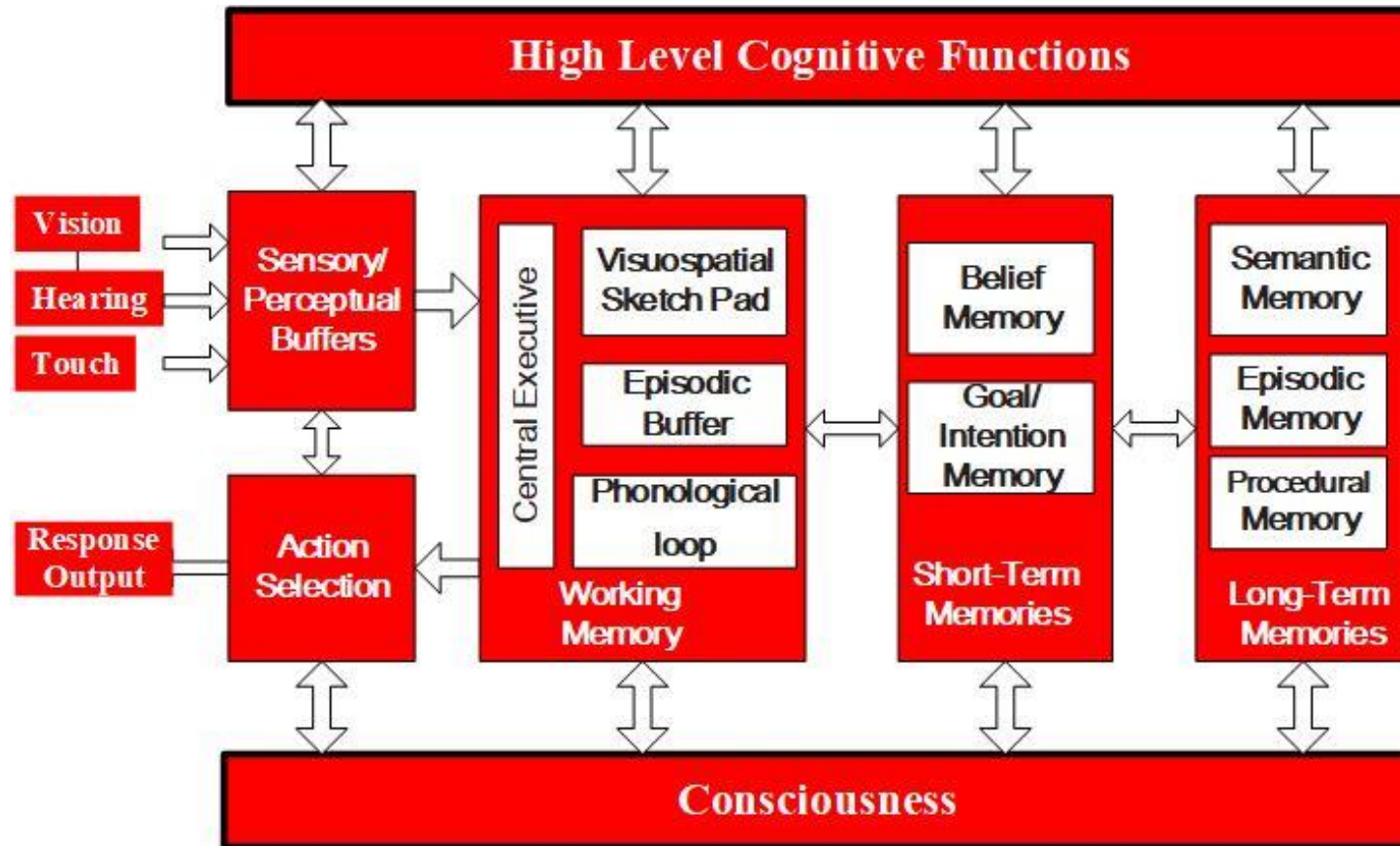


认知机器学习

认知机器学习是指把机器学习与脑认知机理结合起来。

- 学习涌现
- 程序性知识学习
- 学习进化

心智模型 CAM



第三届智能科学国际会议 ICIS2018



The Third International Conference on Intelligence Science (ICIS2018)
November 2-5, 2018, Beijing, China

ICIS2018 Site

- Home
- Call for paper
- Committees
- Keynote Speakers
- Programme
- Dates
- Paper submission
- Review
- Camera-ready and Copyright
- Venue
- Registration
- Accommodation
- Workshop
- Contacts
- Link IFIP AICT 510

The Third International Conference on Intelligence Science (ICIS2018)

ICIS2018, November 2-5, 2018, Beijing, China

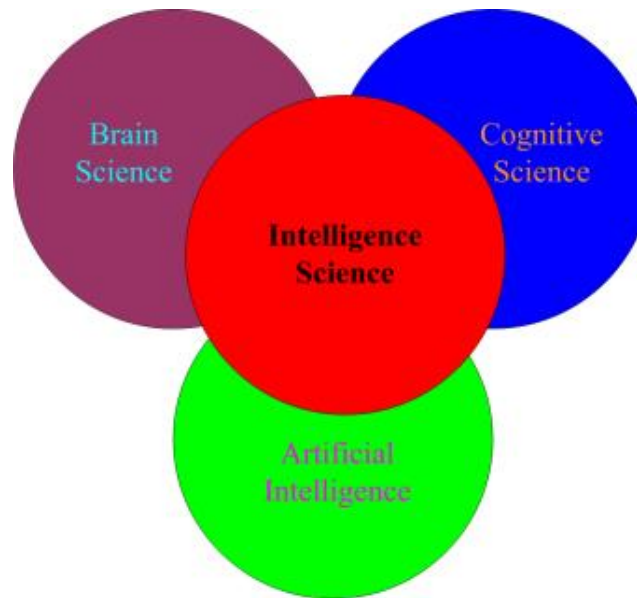
Artificial Intelligence research has made certain substantial progress in some special areas so far. However, the deeper understandings on the essence of intelligence are far from sufficient and, therefore, many state-of-the-art intelligent systems are still not able to compete with human intelligence. To advance the research in artificial intelligence, it is necessary to investigate intelligence, both artificial and natural, in an interdisciplinary context. The objective of this conference is to bring together researchers from brain science, cognitive science, and artificial intelligence to explore the essence of intelligence and the related technologies. The conference provides a platform for discussing some of the key issues related to intelligence science.

The Third International Conference on Intelligence Science (ICIS2018) will be held in Beijing, China, on November 2-5, 2018, focusing on Intelligence Science, Information Science. It is sponsored by Chinese Association for Artificial Intelligence (CAAI), China Chapter of International Society for Information Studies; Organizer is Peking University; and Co-supported by Beijing Association for Science and Technology (BAST), Beijing Association for Artificial Intelligence (BAAI).

Sponsors

Chinese Association for Artificial Intelligence (CAAI)
China Chapter under International Society for Information Studies

Thank You



智能科学 <http://www.intsci.ac.cn/>