

弘扬 Simon的源头创新精神 开拓“AI”的新理念新路径

钟义信
北京邮电大学
e-mail: zyx@bupt.edu.cn

一，概说

源头创新：Simon学术思想的精髓

Simon学术思想的内涵非常丰富，精髓在于“源头创新”。表现在：

- 提出“Physical Symbol Systems”假设：开创了基于符号逻辑的功能主义人工智能和认知科学两个学科的新时代；
- 提出“Means-Ends Analysis”，开创了“智能求解编程”新思路。如 GPS (General Problem Solver)等。

源头创新：今日AI的现实需求

今日的AI不是昨日的AI。它已有“各领风骚”的三路大军：

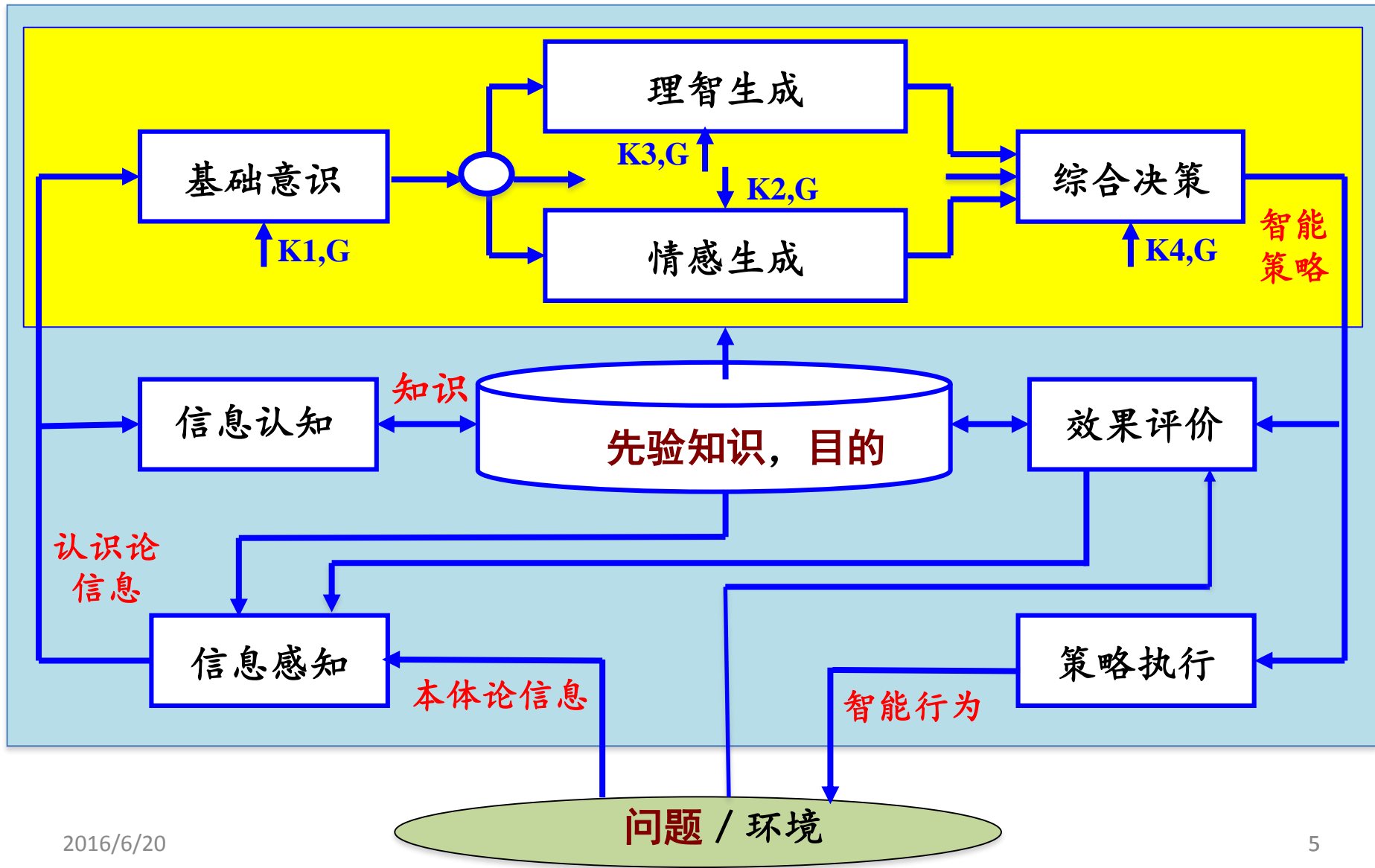
- 结构主义的人工神经网络（系统结构决定系统能力）；
- 功能主义的物理符号系统（系统功能主导系统能力）；
- 行为主义的感知动作系统（系统行为表现系统能力）。

虽然各有各的辉煌，但是，它们未能实现“殊途同归”。于是，人们在从头寻求新的路径（Approach）：

Russell & Norvig. Artificial Intelligence: A Modern Approach, 1995

Nilsson. Artificial Intelligence: A New Synthesis, 1998

源头创新：需要重审人类智能的系统模型



源头考问：AI系统何以能够生成智能？

系统之所以能够“生成智能”，决定的要素是系统的工作“机制”，而不是系统的结构、功能或行为。

结构和功能都是为实现这种“机制”服务的；行为则是“机制”实现的外在表现。所以，要害是“机制”，而不是“结构 / 功能 / 行为”。

基本范式：一旦给定了合理的“问题—目标—知识”，就可以按照**机制**生成解决问题的智能策略。

模型启迪：生成智能的“机制”是信息转换

模型表明，生成智能的共性核心机制是：

“本体论信息—认识论信息—知识—智能策略”转换。

本体论信息：客体事物呈现的运动状态及其变化方式。

认识论信息：主体从本体论信息感知的“状态及其变化方式”的形式、内容、价值，分别称为语法信息、语义信息、语用信息。

知识：主体从大量认识论信息提取出来的“事物状态及其变化规律”。

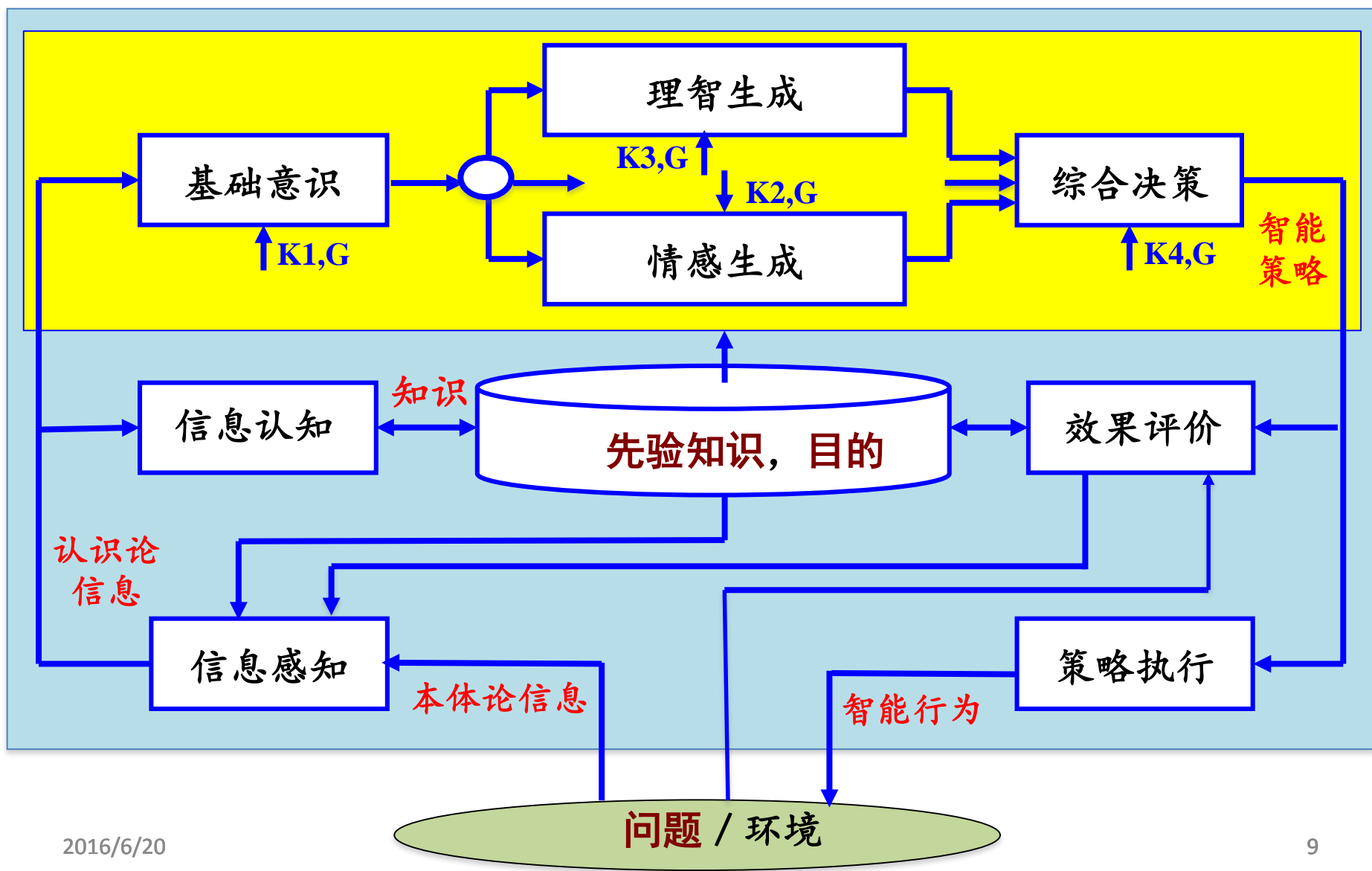
智能策略：在相关的知识支持下、在目标引导下、把认识论信息转换为解决问题的“主客双赢”策略。

成果之一：“机制主义AI”是AI的统一理论*

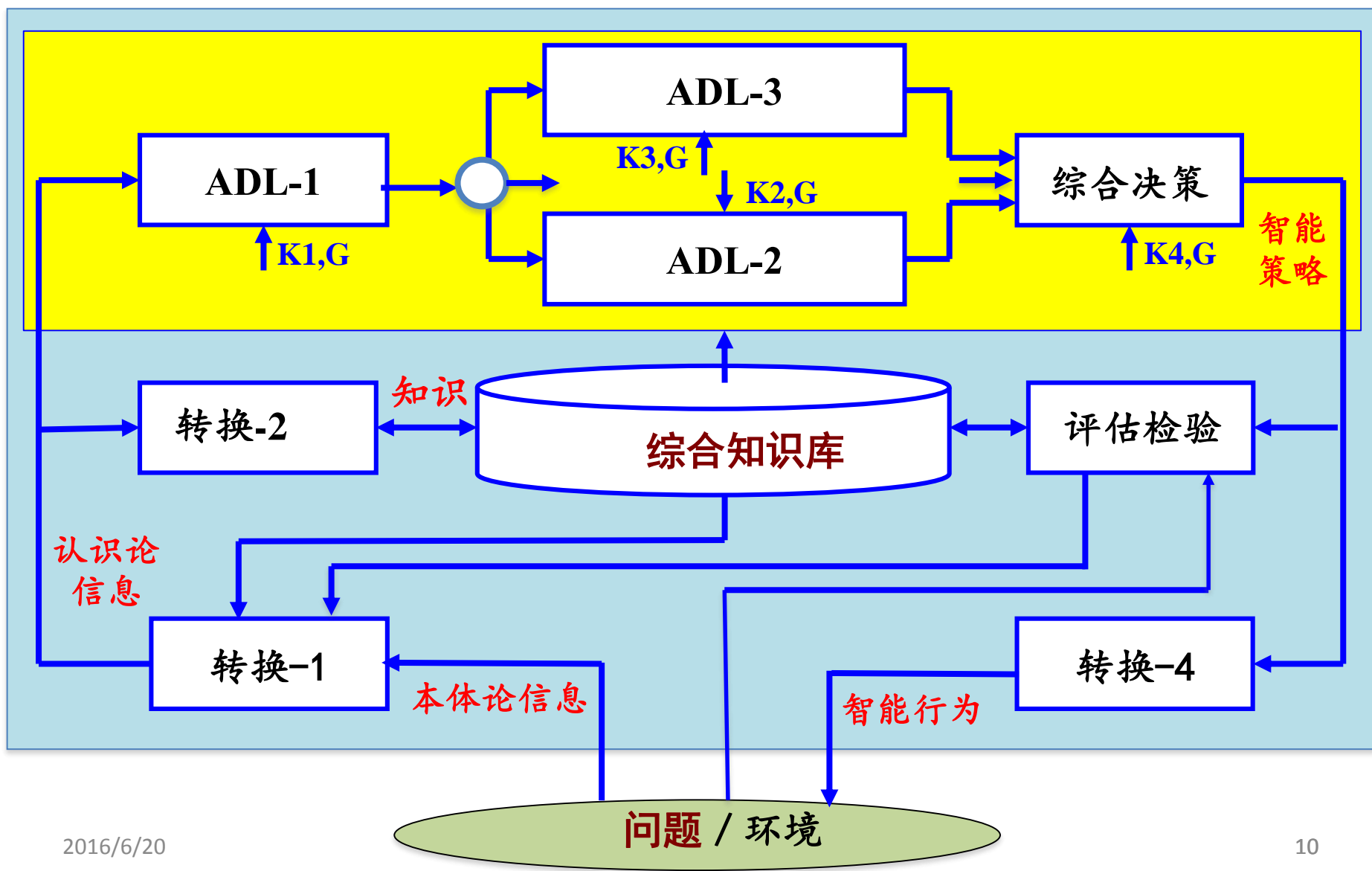
机制主义 AI	信息 → 知识 → 智能策略	实例
A型	信息 → 经验性知识 → 经验型策略	人工神经网络
B型	信息 → 规范性知识 → 规范性策略	物理符号系统
C型	信息 → 常识性知识 → 常识性策略	感知动作系统

* 钟义信. 高等人工智能原理. 科学出版社, 2014

成果之二：机制主义也是“高等AI”的理论



成果之三：“信息转换与智能创生定律”涌现



重要启迪：“转型期”内科学理念的挑战

当代是“物质科学主导期”➔“信息科学主导期”的转型期。理念上，科学观：机械唯物论➔辩证唯物论，方法论：分而治之➔生态演化。

在转型期，人们往往会用物质科学的科学观和方法论去研究信息科学的问题，于是，把复杂的“智能”分解为“结构 / 功能 / 行为”，丢失了内在的联系，不能形成合力。理念（科学观—方法论）滞后，是问题的症结。

“机制主义”途径体现了“辩证唯物”和“生态演化”理念，所以能够取得成功。理念是研究的源头，机制主义途径是AI源头创新的成果！

东方思维吻合“辩证唯物”和“生态演化”方法论，中华民族在信息与智能科学的源头创新研究，将面临大好机会，并将大有作为！

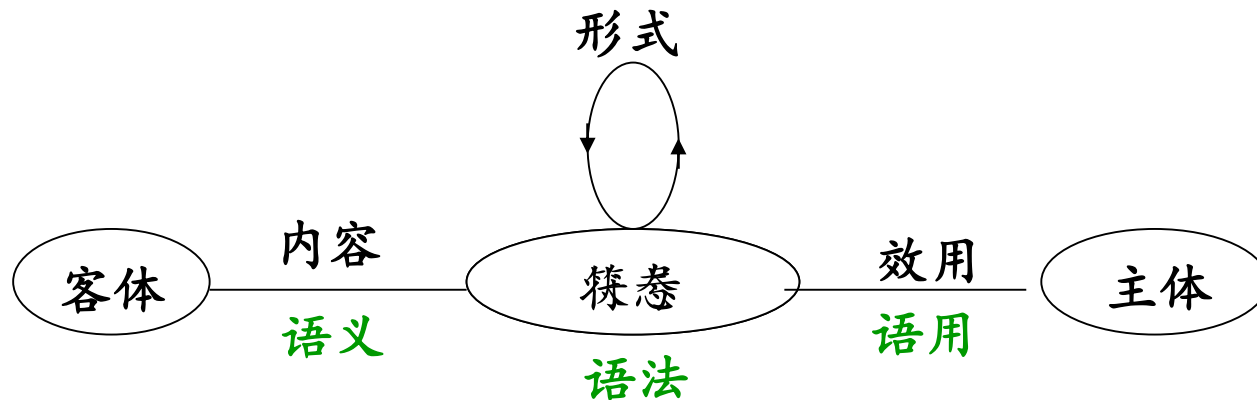
二，详解

“机制主义人工智能”的信息概念

存在两个层次的信息概念，而不是一个（见“生态系统”）！

本体论信息：事物所呈现的运动状态及其变化方式。

认识论信息：主体从本体论信息中所感知的事物的形式（语法信息）、内容（语义信息）和效用（语用信息）。它们的三位一体称为“全信息”，实现了形式、内容、价值的统一。



钟义信，信息科学原理，北邮出版社，1988, 1996, 2002, 2005, 2013

“机制主义人工智能”的知识概念

知识是人类主体在实践中总结的“关于各类事物客体的运动状态及其变化规律的表述”，属于认识论范畴。

所以，知识也有：形态性知识、内容性知识、价值性知识三种分量。

信息是现象，知识才是本质。

钟义信. 知识理论框架. 中国工程科学, Vol.2, No.6, p.50-64, 2000

“机制主义人工智能”的智能概念

智能是智慧概念的子集：智慧是“隐性智慧”和“显性智慧”相互作用相辅相成的统一体，包括“发现问题和定义问题”的能力（称为隐性智慧）以及“解决问题”的能力（显性智慧）。

智能是专指人类智慧能力中的“显性智慧”能力，是在“隐性智慧”所设定的框架内（问题—目标—知识）解决问题的能力”。

智能来源于知识和目的，知识来源于信息。这是智能的“生成机制”。

钟义信. 机器知行学原理. 科学出版社, 2007

钟义信. 高等人工智能原理. 科学出版社, 2014

“机制主义人工智能”的信息理论

若给定的事件为离散的有限状态变量 $X: \{x_1, x_2, \dots, x_n\}$, 则

语法信息的描述参量: 肯定度, $C: \{c_1, c_2, \dots, c_n\}$,

语义信息的描述参量: 逻辑真实度, $T: \{t_1, t_2, \dots, t_n\}$,

语用信息的描述参量: 效用度, $U: \{u_1, u_2, \dots, u_n\}$,

全信息的描述矩阵为

$$\begin{array}{c} x_1, x_2, \dots, x_n \\ c_1, c_2, \dots, c_n \\ t_1, t_2, \dots, t_n \\ u_1, u_2, \dots, u_n \end{array} \quad C = \left\{ \begin{array}{l} P, \text{ 若为随机事件} \\ \mu, \text{ 若为模糊事件} \end{array} \right.$$

其中, 概率分布归一; 隶属度分布不归一, 但可化为归一 [2]。

语义信息的描述和处理方法, 是辩证数理逻辑理论。

“机制主义人工智能”的知识理论

概念：知识是人类主体关于各类客体的运动状态及其变化规律的表述，与认识论信息（全信息）相同，是充满生机的生态过程。

理论：知识的外部生态规律

认识论信息 → 知识 → 智能

理论：知识的内部生态规律

经验型知识 → 规范型知识 → 常识型知识

“机制主义人工智能”的智能理论

根据“知识的外生态规律：信息→知识→智能”，提出智能机制模拟。

根据“知识的内生态规律”，三大学派和谐统一于机制模拟方法：

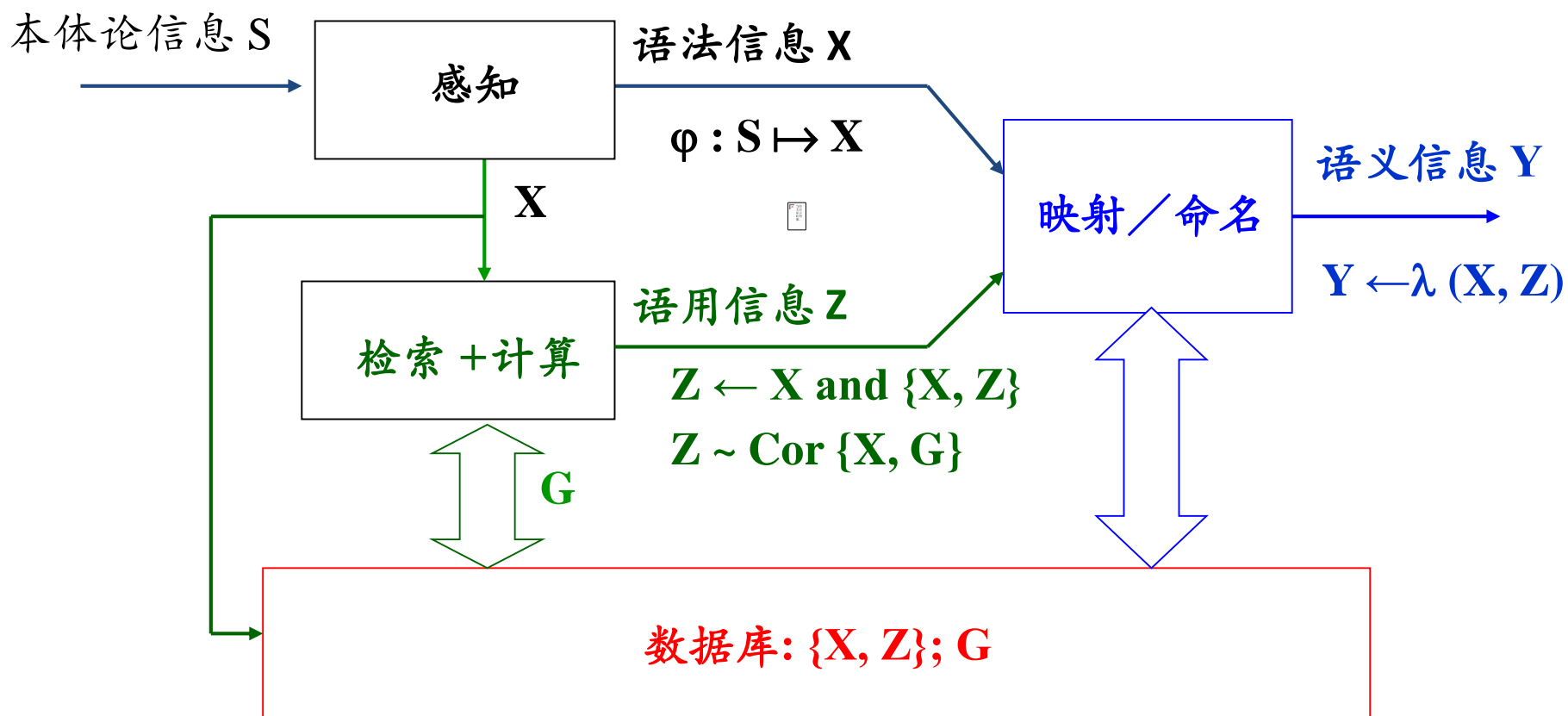
机制模拟 {
A型 (基于经验知识): 人工神经网络 (结构模拟)
B型 (基于规范知识): 经典人工智能 (功能模拟)
C型 (基于常识知识): 智能机器人 (行为模拟)
D型:

机制主义：人工智能的统一理论*

机制模拟方法	信息 → 知识 → 智能策略			实例
A型	信息	经验性知识	经验型策略	人工神经网络
B型	信息	规范性知识	规范性策略	物理符号系统
C型	信息	常识性知识	常识性策略	感知动作系统

* 钟义信. 高等人工智能原理. 科学出版社, 2014

第一类信息转换方法：本体论信息→认识论信息

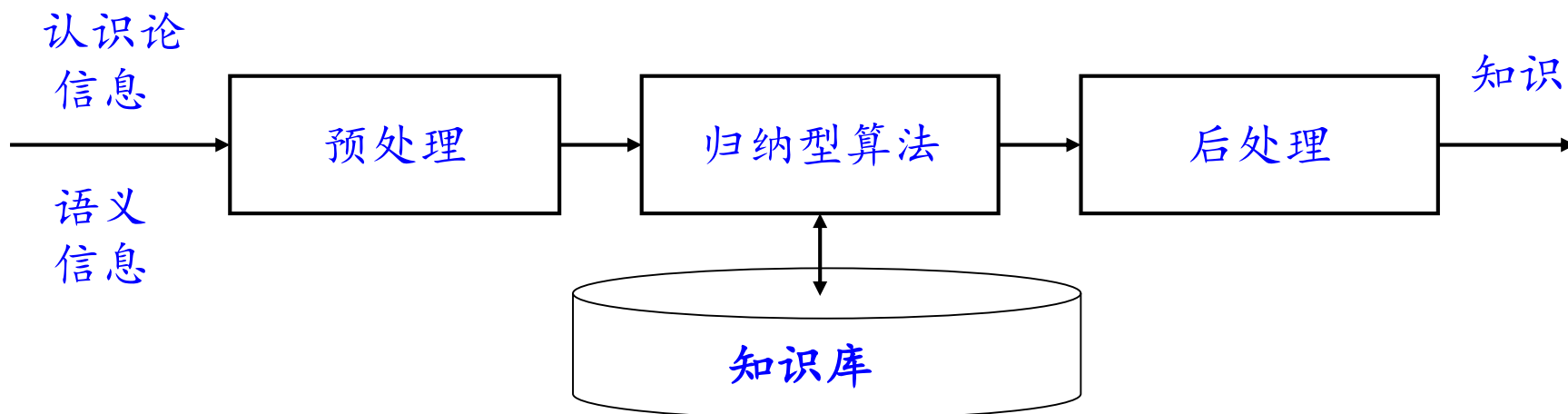


第二类信息转换方法：认识论信息→知识

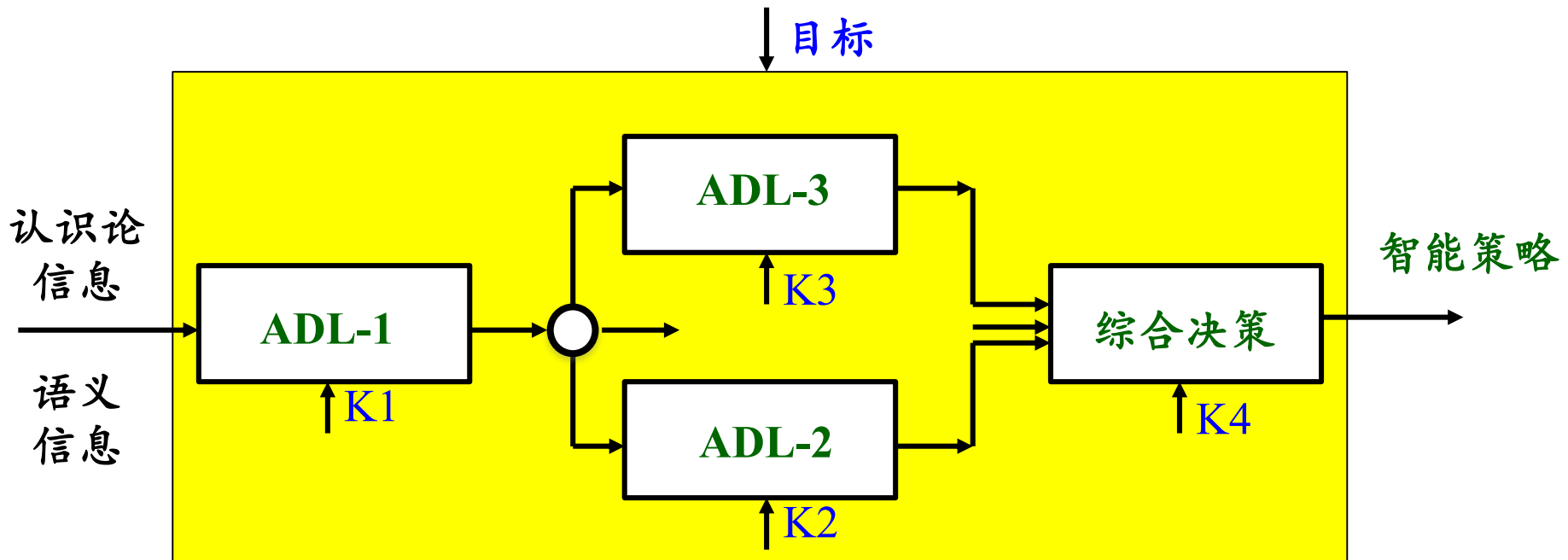
由“认识论信息 / 语义信息”提炼“知识”的主要算法是“归纳型算法”。

如：乌鸦1是黑的，乌鸦2是黑的，...，乌鸦N是黑的 → 乌鸦是黑的。

归纳型算法的特点是不保真；真度与样本的典型性有关。



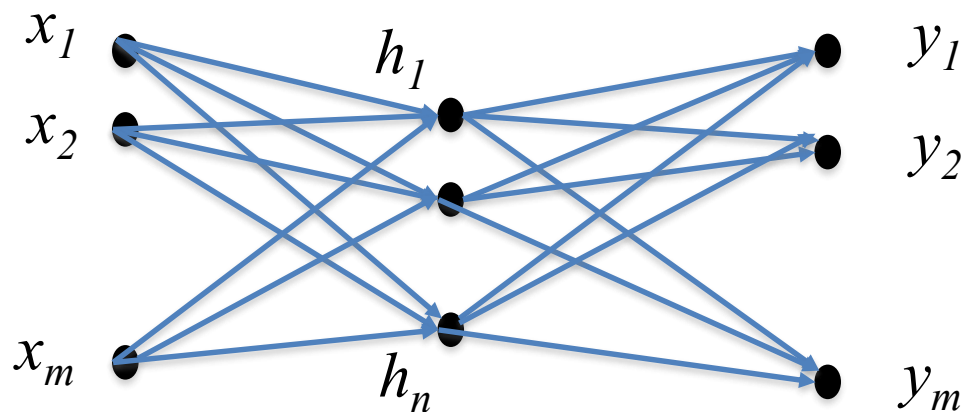
第三类信息转换方法：认识论信息→知识→智能



K1: 本能与常识知识, K2: K1 + 经验型知识, K3: K2 + 规范型知识, K4: K3 + 决策艺术

第三类信息转换的主导算法是演绎型算法。

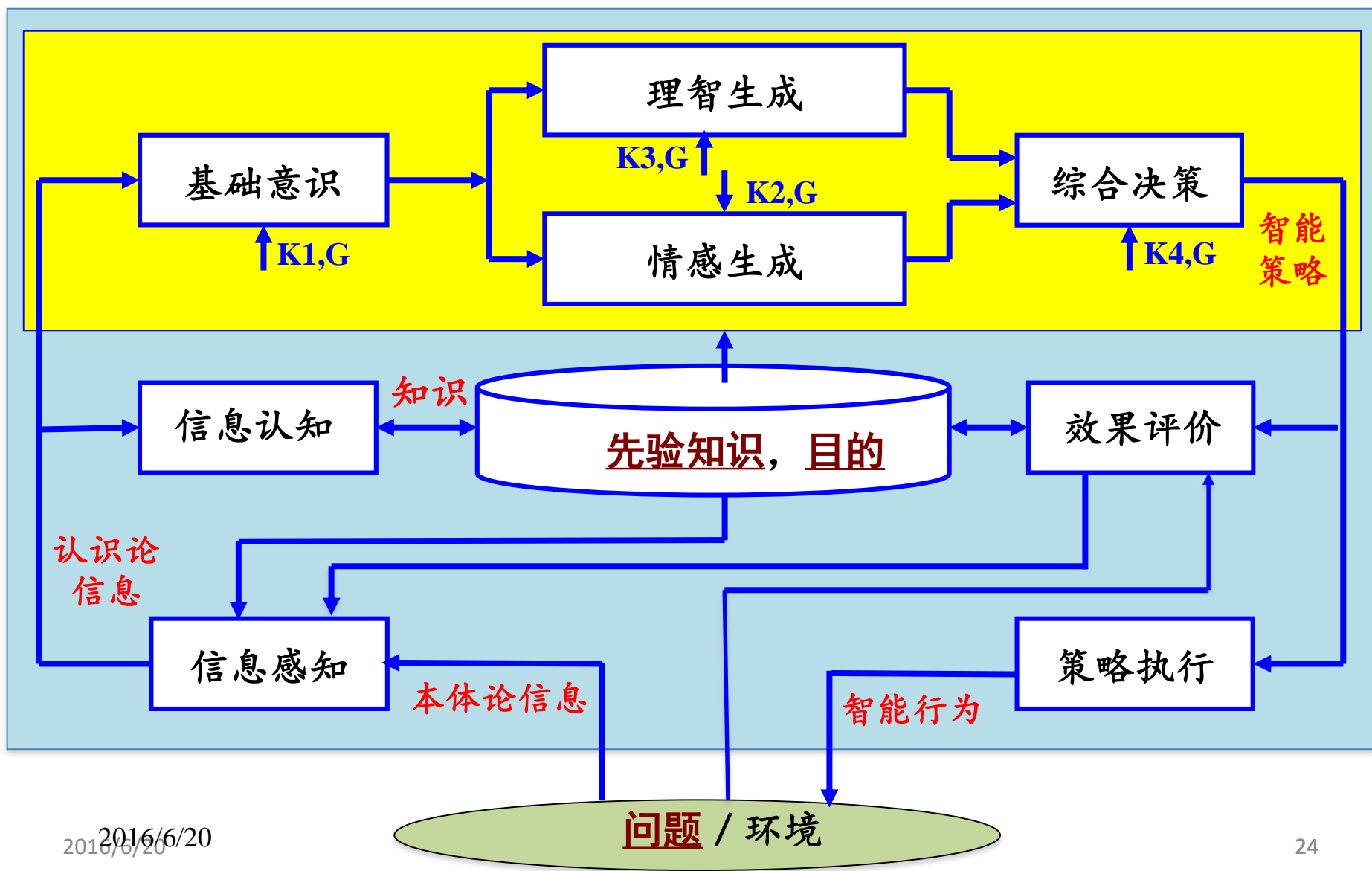
DL: 基本概念



令 $y_k = x_k$, $k = 1, \dots, m$ 且 $n < m$,

把 $\{h_l \mid l = 1, \dots, n\}$, 作为下一级的网络输入, 做类似处理。

信息转换与智能创生定律：信息领域根本定律



定律的学术意义

