



人工智能基础与进阶

知识与推理

上海交通大学

目录 content



第一节

知识表示

第二节

知识推理

第三节

产生式系统的推理

第四节

知识图谱

第一节 知识表示



知识表示

在人工智能中，自然而然会遇到“**如何表示知识？**”，“**计算机如何理解知识？**”，“**知识能否被程序利用？**”这一类问题，怎样使机器懂得知识，对其进行处理，并能够以人类能理解的方式呈现是最基本的问题之一。

知识表示**就是要把问题求解中所需要的对象、前提条件、算法等知识构造为计算机可处理的数据结构以及解释这种结构的某些过程。**

智能活动主要是一个获得并应用知识的过程，而知识必须有适当的表示方法才便于在计算机中有效地存储、检索、使用和修改。

知识表示-知识

知 识

1. 是信息接收者通过对信息的提炼和推理而获得得的正确结论；
2. 是人对自然世界、人类社会以及思维方式与运动规律的认识与掌握；
3. 是人的大脑通过思维重新组合并和、系统化后的信息集合。

知识表示-知识

在知识表示中的知识一般有分为如下几类：

1. 有关现实世界中所关心对象的**概念**，即用来**描述现实世界所抽象总结出的概念**。
2. 有关现实世界中发生的**事件**、事件相关所关系对象的行为、状态等内容，也就是说**不光有静态的概念，还有动态的信息**。
3. 关于**过程**的知识，即不光有当前状态和行为的描述，还要有**对其发展的变化及其相关条件、因果关系等描述的知识**。
4. **元知识**，即关于知识的知识，是知识库中的高层知识。包括怎样使用规则、解释规则、校验规则、解释程序结构等知识，包括**如何利用已知知识的知识**。

知识表示-知识

在现阶段我们着重关注知识的以下特性：

1. **条件性**。一定条件下/某种环境中
2. **不确定性**。中间状态/为真程度/随机性/模糊性/经验性/不完全性
3. **可表示性**。语言/文字/图像/视频/图形/音频/神经网络/概率图

知识表示-知识表示方法

经过国内外学者的共同努力，目前已经有许多知识表示方法得到了深入的研究，目前使用较多的知识表示方法主要有以下几种知识表示方法。

- **一阶谓词逻辑表示**：用谓词来表示动作的主客体，是一种叙述性的知识表示方法。
- **产生式表示法**：又称规则表示或 IF-THEN 表示。表示一种条件-结果形式，是一种比较简单的知识表示方法
- **知识图谱表示法**：一种语义网络的知识库。
- **框架表示法、脚本表示法、基于XML的表示法等。**

知识表示-知识表示过程

从一般意义上讲，知识表示就是为描述世界所做的一组约定，是**知识的符号化、形式化或模型化**；从计算机科学的角度来看，知识表示是研究计算机表示知识的可行性、有效性的一般方法，是**把人类知识表示成机器能处理的数据结构和系统控制结构的策略**。



知识表示-知识表示过程

一个完整知识表示过程是：

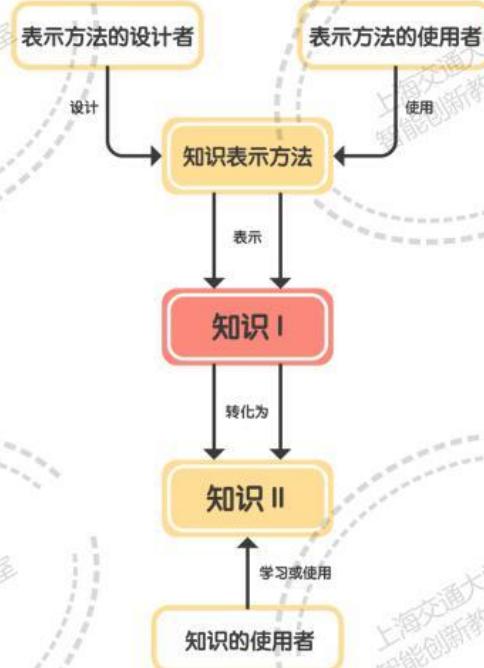
1. 首先**表示方法的设计者**针对各种类型的问题**设计多种知识表示方法**；
2. 然后**表示方法的使用者**选用合适的表示方法**表示某类知识**；
3. 最后**知识的使用者**使用或者学习**经过表示方法处理后的知识**。

知识表示-知识表示过程

知识表示的客体就是知识

知识表示的主体包括3类：

1. 表示方法的设计者
2. 表示方法的使用者
3. 知识的使用者



知识表示-知识表示过程

假设有这样一个知识需要表示：

小潘是计科系的学生，但他不喜欢编程。

用一阶谓词逻辑来表示它就需要采用如下的步骤：

首先，定义谓词

$\text{Computer}(x)$: x 是计科系的学生

$\text{Like}(x, y)$: x 喜欢 y

其次，用谓词公式表示上述知识：

$\text{Computer}(\text{xiaopan}) \wedge \text{not Like}(\text{xiaopan}, \text{programing})$

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第二节 知识推理



上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

知识推理

知识推理是指在计算机或智能系统中，模拟人类的智能推理方式，依据**推理控制策略**，利用形式化的知识进行**机器思维**和求解问题的过程。即**使用知识解决问题**。



知识推理

智能系统的知识推理包括两个基本问题：

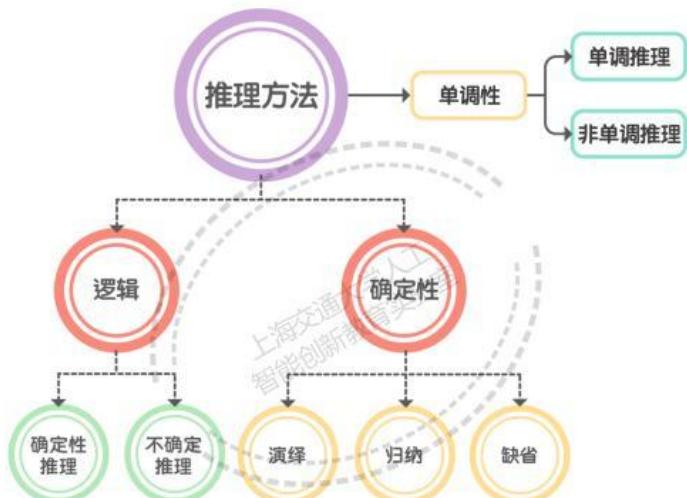
一是**推理方法**，研究的是前提与结论之间的种种逻辑关系及其可信度传递规律等；

二是**推理的控制策略**。而控制策略的采用是为了限制和缩小搜索的空间，使原来的指数型困难问题在多项式时间内求解。

知识推理的方法

推理方法主要解决**在推理过程中前提与结论之间的逻辑关系**，
以及**在非精确性推理中不确定性的传递问题**。

按照分类标准的不同，推理方法主要有以下三种分类方式：



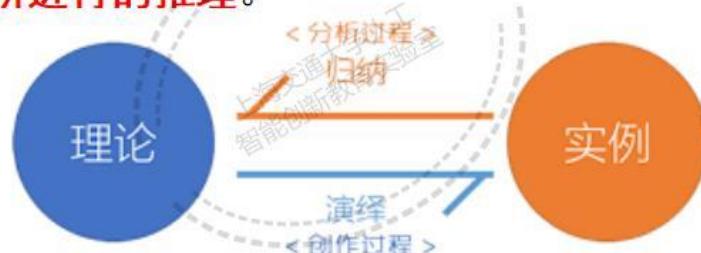
知识推理的方法

1. 从逻辑基础上分，可分为演绎推理、归纳推理和缺省推理

(1) 演绎推理是从已知的一般性知识出发，去推出蕴含在这些已知知识中的适合于某种个别情况的结论。是一种由**一般到个别的推理方法**，其核心是三段论。

(2) 归纳推理是一种由**个别到一般的推理方法**。从足够多的事例中归纳出一般性结论的推理过程。

(3) 缺省推理又称为默认推理，它是在**知识不完全的情况下假设某些条件已经具备所进行的推理**。



知识推理的方法-演绎推理

演绎推理（deduction）：

演绎是从**一般到特殊的过程**。所谓演绎推理，就是从一般性的前提出发，通过演绎，得出具体陈述或个别结论的过程。

最经典的演绎推理就是三段论（syllogism），包括一个一般性原则（**大前提**）、一个附属于大前提的特殊化陈述（**小前提**）、一个由此引申出的特殊化陈述或者符合一般性原则的**结论**。

例如：

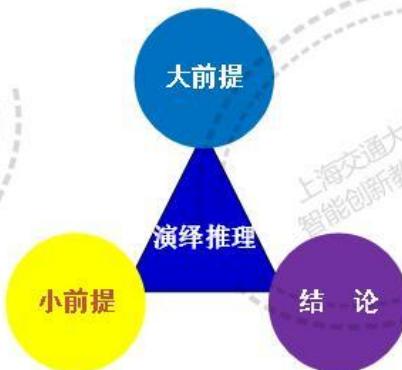
大前提：计算机系的学生都会编程。

小前提：李强是计算机系的学生。

结 论：李强会编程。

知识推理的方法-演绎推理

演绎推理 (deduction)：演绎推理不仅仅局限于三段论，也不是从一般到特殊的过程。它有着强烈的演绎特性，**重在通过利用每一个证据，逐步地推导到目标或以外的结论**，多被用于数学物理证明、思维推导等各类应用。



知识推理的方法-演绎推理

莎士比亚在《威尼斯商人》中，描写富家少女鲍西亚品貌双全，贵族子弟、公子王孙纷纷向她求婚。鲍西亚按照其父遗嘱，由求婚者猜盒订婚。鲍西亚有金、银、铅三个盒子，分别刻有三句话，其中只有一个盒子放有鲍西亚的肖像。求婚者中谁通过这三句话，最先猜中肖像放在哪个盒子里，谁就可以娶到鲍西亚。

金盒子上说：“肖像不在此盒中。”

银盒子上说：“肖像在铅盒中。”

铅盒子. 上说：“肖像不在此盒中。”

鲍西亚告诉求婚者，上述三句话中，最多只有一句话是真的。如果你是一位求婚者，如何尽快猜中鲍西亚的肖像究竟放在哪一个盒子里？

知识推理的方法-演绎推理

大前提：银盒子与铅盒子的话互相矛盾，必一真一假。

小前提：三句话中，最多只有一句话是真的。

结论：金盒子的话为假。

知识推理的方法-归纳推理

归纳推理(induction)：

归纳是从**特殊到一般的过程**。所谓归纳推理，就是从一类事物的大量特殊事例出发，去推出该类事物的一般性结论。我们熟知的数学归纳法就是归纳推理的一个典型例子。

前提：

蓝鲸可以喷射水柱。抹香鲸可以喷射水柱。座头鲸可以喷射水柱

结论：

鲸鱼都可以喷射水柱。

知识推理的方法-归纳推理

演绎推理与归纳推理的区别：

演绎推理是在已知领域内的一般性知识的前提下,通过演绎求解一个具体问题或者证明一个结论的正确性。它所得出的结论实际上早已蕴含在一般性知识的前提中,演绎推理只不过是将已有事实揭示出来,因此它不能增殖新知识。

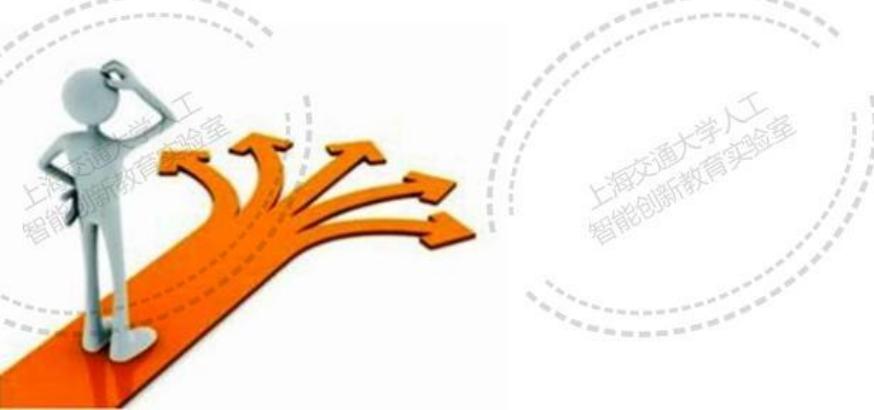
而相反, **归纳推理所推出的结论是没有包含在前提内容中的**。这种由个别事物或现象推出一般性知识的过程,是增殖新知识的过程。

知识推理的方法-确定性推理与不确定性推理

2. 从确定性上分，可分为确定推理和不确定推理

(1) 确定性推理是指推理时所用的知识都是精确的，推出的结论也是确定的，其真值或者为真，或者为假，没有第三种情况出现。

(2) 不确定性推理是指推理时所用的知识不都是精确的，推出的结论也不完全是肯定的，其真值位于真与假之间。



知识推理的方法-确定性推理与不确定性推理

确定性推理: 确定性推理**大多指确定性逻辑推理**, 它具有**完备的推理过程和充分的表达能力**, 可以严格地按照专家预先定义好的规则准确地推导出最终结论。

但是确定性推理很难应对真实世界中, 尤其是存在于网络大规模知识图谱中的不确定甚至不正确的事实和知识。

不确定性推理: 不确定性推理也被称为概率推理, 是统计机器学习中一个重要的议题。**它并不是严格地按照规则进行推理, 而是根据以往的经验和分析**, 结合专家先验知识构建概率模型, 并利用统计计数、最大化后验概率似然估计等统计学习的手段对推理假设进行验证或推测。不确定性推理可以有效建模真实世界中的不确定性。

知识推理的方法-单调推理和非单调推理

3. 从单调性上分，可分为单调推理和非单调推理。

(1) 单调推理推出的结论呈单调增加的趋势，并且越来越接近最终目标。一个演绎推理的逻辑系统有一个无矛盾的公理系统，新加入的结论必须与公理系统兼容，因此新的结论与已有的知识不发生矛盾，结论总是越来越多，所以演绎推理是单调推理。

(2) 非单调推理由新知识的加入，不仅没有加强已推出的结论，反而要否定它。非单调推理的处理过程要比单调推理的过程复杂和困难很多。

知识推理的方法-单调推理和非单调推理

我们需要**非单调推理**的主要原因是：

(1) **由于缺乏完全的知识，只好对部分问题作暂时的假设。**

而这些假设可能是对的，也可能是错的。但错了以后要能够在某时刻得到修正，这就需要非单调推理。

(2) **客观世界变化太快，某一时刻的知识不能持久使用，这也需要非单调推理来维护知识库的正确性。**



知识推理的控制策略

知识推理的控制策略分为三个方面：**推理方向、搜索策略和冲突消解策略。**

知识推理的控制策略

(1) 推理方向

主要包括正向推理、逆向推理和混合推理。

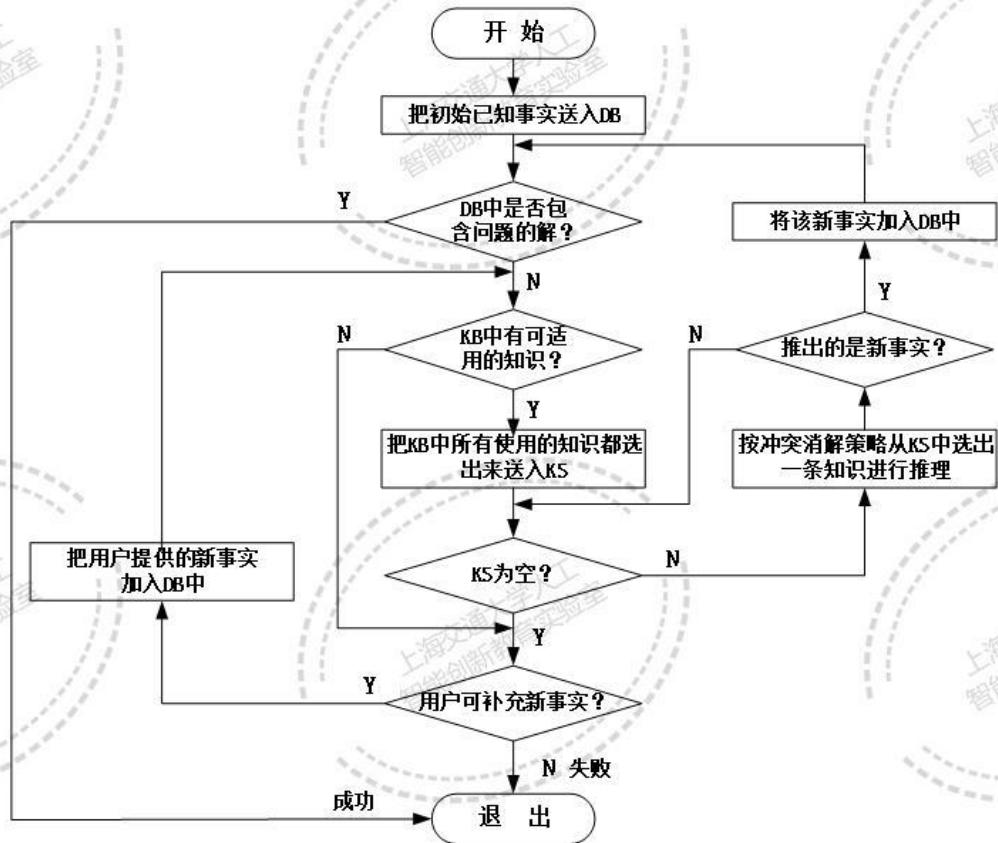
正向推理又称为事实驱动或数据驱动推理，其主要优点是比较直观，允许用户提供有用的事实信息，是产生式专家系统的主要推理方式之一。

知识推理的控制策略-正向推理

基本思想：

- 1) 从用户提供的初始已知事实出发，在**知识库KB**中找出**当前可适用的知识**，构成可适用的**知识集KS**，
- 2) 然后按某种冲突消解策略从知识集中**选出一条知识进行推理**，并将推出的新事实加入到**数据库DB**中，作为**下一步推理的已知事实**。
- 3) 在此之后，再在知识库中**选取可适用的知识进行推理**。如此重复进行这一过程，直到求得所要求的解。

知识推理的控制策略-正向推理



知识推理的控制策略-正向推理

以识别某动物过程为例：

已知事实：一动物 {有毛，吃草，黑条纹}

R1：动物有毛 → 哺乳类

R2：动物产奶 → 哺乳类

R3：哺乳类 ∧ 吃肉 → 食肉类

R4：哺乳类 ∧ 吃草 → 有蹄类

R5：食肉类 ∧ 黄褐色 ∧ 有斑点 → 猎狗

R6：食肉类 ∧ 黄褐色 ∧ 黑条纹 → 虎

R7：有蹄类 ∧ 长脖子 → 长颈鹿

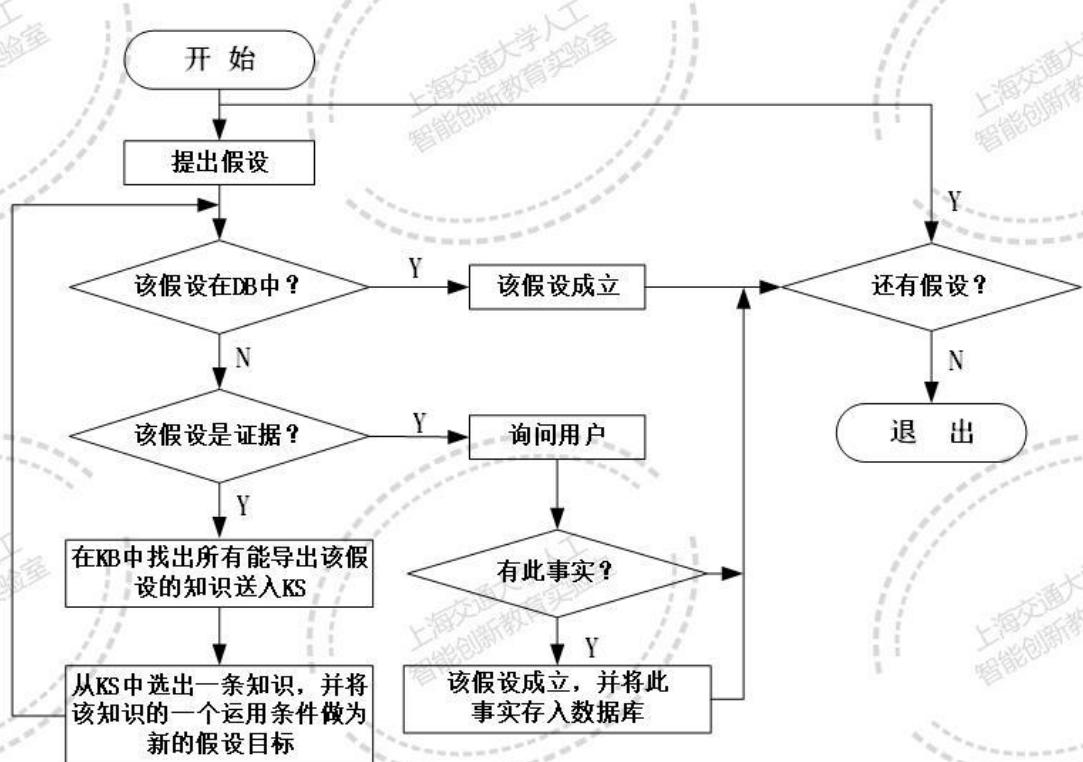
R8：有蹄类 ∧ 黑条纹 → 斑马

在给定的知识库中，根据R1、R4、R8判据，可以认为该{有毛，吃草，黑条纹}动物为斑马。

知识推理的控制策略-逆向推理

逆向推理又称**目标驱动或假设驱动推理**，其主要优点是不必使用与总目标无关的规则，且有利于向用户提供解释。基本思想：
首先选定一个假设目标，然后寻找支持该假设的证据，若所需的证据都能找到，则说明原假设是成立的；若找不到所需要的证据，则说明原假设不成立，此时需要另作新的假设。

知识推理的控制策略-逆向推理



知识推理的控制策略-逆向推理

以识别某动物过程为例：

已知事实：一动物 {有毛，吃草，黑条纹}

R1：动物有毛 → 哺乳类

R2：动物产奶 → 哺乳类

R3：哺乳类 ∧ 吃肉 → 食肉类

R4：哺乳类 ∧ 吃草 → 有蹄类

R5：食肉类 ∧ 黄褐色 ∧ 有斑点 → 猎狗

R6：食肉类 ∧ 黄褐色 ∧ 黑条纹 → 虎

R7：有蹄类 ∧ 长脖子 → 长颈鹿

R8：有蹄类 ∧ 黑条纹 → 斑马

在给定的知识库中，逆向推理依次根据R8、R4、R1可知该动物为斑马。

知识推理的控制策略

(2) 搜索策略

搜索策略主要包括**盲目搜索**和**启发式搜索**，前者包括**深度优先搜索**和**广度优先搜索**等搜索策略；后者包括**A*搜索**和**蚁群搜索**等搜索策略。

知识推理的控制策略

(3) 冲突消解策略

冲突是指同时推理过程中多个知识都匹配成功，相互冲突。

如：正向推理时，多条知识的**前件**（前提）代表的**事实**相同，但**后件**（结论）不同。

逆向推理时，多条知识的**后件**（结论）代表的**假设**相同，但**前件**（前提）不同。

知识推理的控制策略

冲突消解的基本思想都是对知识进行排序，方法有：

- 1) **按针对性排序**: 优先选用针对性强的产生式规则。
- 2) **按已知事实的新鲜性排序**: 优先选用与较多新事实匹配的规则。
- 3) **按匹配度排序**: 在不确定性匹配中，计算两个知识模式的相似度(匹配度)，并对其排序，相似度高的规则先推。
- 4) **按领域问题特点排序**
- 5) **按上下文限制排序**: 把规则按照下上文分组，并只能选取组中的规则。
- 6) **按冗余限制排序**: 冗余知识越少的规则先推。
- 7) **按条件个数排序**: 条件少的规则先推。

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第三节 产生式系统的推理



上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

产生式

产生式，又叫**产生式规则**或**规则**，是一种知识表达形式，其基

本形式为：

<前件> \rightarrow <后件>

产生式规则的语义是：如果前提满足，则可得结论或者执行相应的动作，即后件由前件来触发。

前件表示原因、条件、前提、事实、情况等，而**后件**表示结果、结论、操作、进展、行为等。

产生式

在自然语言表达中，人们广泛使用的各种“原因→结果”，“条件→结论”，“前提→操作”，“事实→进展”，“情况→行为”等结构，都可归结为**产生式**的知识表达形式。

例如，

天下雨，地上湿； “**原因→结果**”

如果把水加热到00以上，冰就会溶化为水； “**条件→结论**”

夜来风雨声，花落知多少； “**事实→进展**”

若能找到一根合适的杠杆，就能撬起那座大山； “**前提→操作**”

产生式系统的推理

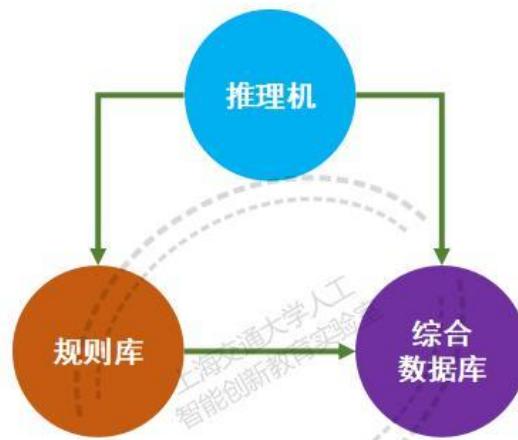
产生式系统的推理是基于产生式知识表示的推理，又称基于规则的推理（Rule-Based Reasoning, RBR），其核心是演绎推理，从一组前提必然推导出某个结论。

特点：

- (1) 具有很强的推理能力和较高的推理效率；
- (2) 知识表示形式简单，通常是IF-THEN结构，易于系统实现；
- (3) 知识获取困难，需要靠人工“移植”方式获取专家知识，知识库维护困难；
- (4) 运行效率随规则库规模的增大而迅速降低；
- (5) 对于非结构化的知识组织形式，求解复杂问题困难。

产生式系统的组成

产生式系统一般由三个基本部分组成：规则库、综合数据库和推理机。



产生式系统的组成

(1) 规则库是用于描述某领域内知识的产生式集合，是某领域知识(规则)的存储器，其中的规则是以产生式形式表示的。规则库中包含着将问题从初始状态转换成目标状态(或解状态)的那些变换规则。

产生式系统的组成

(3) **综合数据库**又称为事实库，用于存放输入的事实、从外部数据库输入的事实以及中间结果(事实)和最后结果。当规则库中的某条产生式的前提可与综合数据库中的某些已知事实匹配时，该产生式就被激活，并把用它推出的结论放入综合数据库中，作为后面推理的已知事实。

显然，综合数据库的内容是在不断变化的，是动态的。

产生式系统的组成

(4) **推理机**是一个或一组程序，用来控制和协调规则库与综合数据库的运行，包含了推理方式和控制策略。控制策略的作用就是确定选用什么规则或如何应用规则。

产生式系统的组成

通常从选择规则到执行操作分3步完成：匹配、冲突解决和操作。

(1) **匹配**：匹配就是将当前综合数据库中的事实与规则中的条件进行比较，如果相匹配，则这一规则称为匹配规则。因为可能同时有几条规则的前提条件与事实相匹配，究竟选哪一条规则去执行呢？这就是规则冲突解决。通过冲突解决策略选中的在操作部分执行的规则称为启用规则。

(2) **冲突解决**：冲突解决的策略有很多种，其中专一性排序、规则排序、规模排序和就近排序是比较常见的冲突解决策略。

(3) **操作**：操作就是执行规则的操作部分。经过操作以后，当前的综合数据库将被修改，其他的规则有可能将成为启用其他规则继续推理。

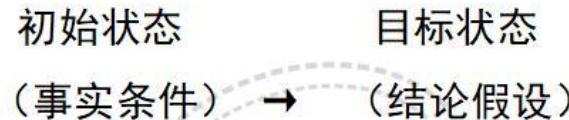
产生式系统的正向推理

(1) 基于规则正向推理：

正向推理是从已知事实出发，通过规则库求得结论。正向推理方式也被称为数据驱动方式或自底向上的方式。

这种推理方式是由数据到结论，所以也叫数据驱动策略。

推理方式如下所示：



正向推理的过程是：

搜索规则（知识）库，逐条检查规则的前提在事实库中是否存在。若各子项不全都存在，则放弃该规则；若全都存在，则执行该规则，并把结论放入综合数据库或对综合数据库进行必要的修改。

产生式系统的正向推理

反复执行以上过程，直至推出目标。举例如下，

问题：患者眼睑局部红肿，顶尖有脓点，观察舌苔薄且黄，自述胃纳差。根据产生式规则，诊断疾病并提出治疗方案

产生式规则

R1：毛囊皮根微红肿→轻型

R2：鼻塞流涕→外感风热

R3：舌苔薄黄 ∧ 胃纳差→胃肠积热

R4：眼睑局部明显红肿→重型

R5：眼睑局部红肿 ∧ 顶尖有脓点→重型

R6：脓点破溃→晚期

R7：轻型 ∧ 外感风热→轻风热型

R8：重型 ∧ 外感风热→重风热型

R9：重型 ∧ 胃肠积热→重积热型

R10：轻型→散瘀法

R11：晚期→外敷药物

R12：轻风热型→散瘀加罐

R13：重风热型→挑瘀加罐

R14：重积热型→放血法

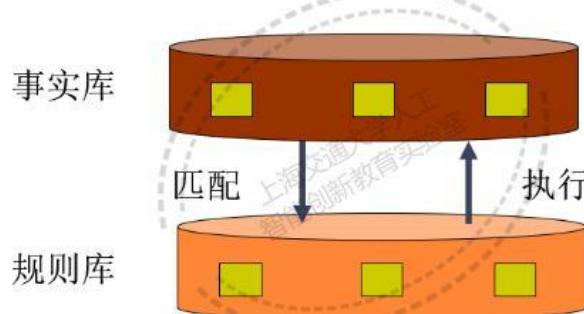
产生式系统的正向推理

输入 “眼睑局部红肿” \wedge “顶尖有脓点” \wedge “舌苔薄黄” \wedge “胃纳差”

循环次数	使用规则	事实变动情况
1	R3	加入 “胃肠积热”
2	R5	加入 “重型”
3	R9	加入 “重积热型”
4	R14	加入 “放血法”

产生式系统的正向推理

计算机利用正向推理论解问题时：先将事实数据存入计算机的事实库中，将领域知识表示为规则，存入规则库中。推理时将问题的事实与规则的前提进行**匹配**。前提可能由条件或子句集合组成，如果规则前提中的所有子句被匹配成功，则**执行**这条规则。将执行后所得的新事实存入事实库中，再次寻找匹配的规则，直至得出结论。



产生式系统的正向推理

正向推理的冲突解决策略：

重要度优先——预先给各规则赋予表示其**重要程度的权值**，在处理冲突规则时，**选择权值最高的规则**。

最近优先法——优先选择与**最近加入**事实库中的事实相匹配的规则。在这种情况下，各数据元素被赋予时间标志。在以**实时控制**为目标的事件驱动型推理中，常使用这种策略。

产生式系统的正向推理

正向推理的一般算法：

1. 扫描规则库，产生可用规则集S，这些规则左边条件均为真，即都被问题的条件事实满足；
2. 调用解决冲突算法，从S中选出规则R；
3. 执行规则R右边的结论部分，将产生的新事实加入事实库；
4. 若目标得证或无新的事实产生，则停止；否则转2。

产生式系统的逆向推理

逆向推理是从**目标（作为假设）**出发，反向使用规则，求得已知事实。这种推理方式也被称为**目标驱动方式或自顶向下的方式**。

这种推理方法由目标到数据，因此也称为**目标驱动策略**。

推理方式如下所示：



反（逆）向推理过程是：

从目标开始，寻找以目标为结论的规则，并对该规则的前提进行判断，若前提中某子项是另一规则的结论，则找此结论的规则重复以上过程，直到对某规则的前提能够进行判断。

产生式系统的逆向推理

交通事故 **逆向推理**示例：

目标假设：李先生不出交通事故。

谓词：

年龄(x) //表示x的年龄

中年人(x) //表示x是中年人

老练(x) //表示x很老练

细心(x) //表示x很细心

有驾驶技术(x) //表示x有驾驶技术

不出交通事故(x) //表示x不出交通事故

事实：

F1: 年龄(李先生) = 43 //李先生的年龄是43

F2: 有驾驶技术(李先生) //李先生有驾驶技术

产生式系统的逆向推理

规则：(逆向规则表示法)

R1: 中年人 (x)

If 年龄 (x) <= 55 and 年龄 (x) >= 35

R2: 老练 (x)

If 中年人 (x)

R3: 细心 (x)

If 中年人 (x)

R4: 不出交通事故 (x)

If 老练 (x) and 细心 (x) and 有驾驶技术 (x)

产生式系统的逆向推理

推理过程（用目标匹配规则的结论）

不出交通事故(李先生) → (R4)

老练(李先生)

细心(李先生)

有驾驶技术(李先生) ✓ (F2)

老练(李先生) → 中年人(李先生) (R2)

细心(李先生) → 中年人(李先生) (R3)

中年人(李先生) → (R1, F1)

年龄(李先生) 小于等于55 ✓

年龄(李先生) 大于等于35 ✓

产生式系统的逆向推理

逆向推理的算法：

设目标为G，逆向推理算法为：

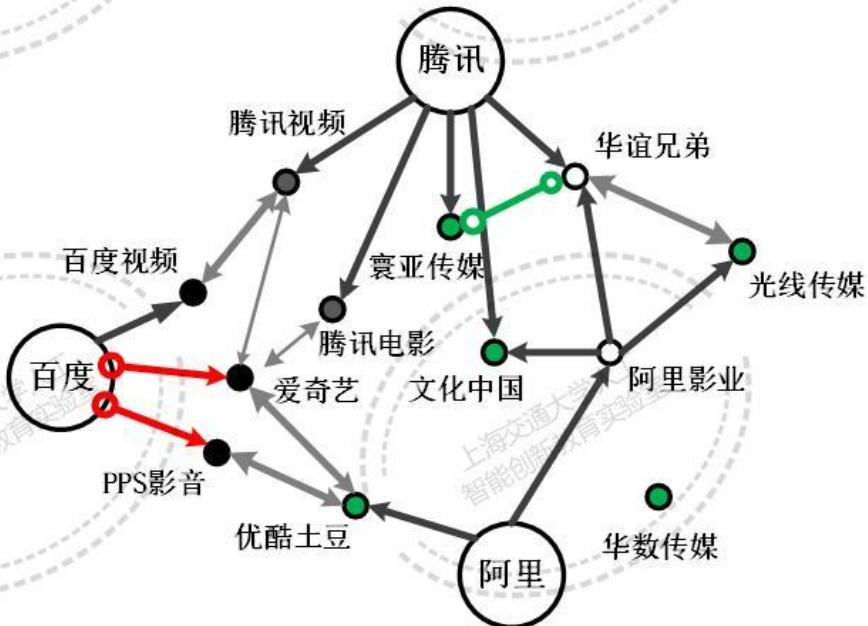
1. 扫描事实库，找出与目标G匹配的事实F，若F存在，则成功返回。
2. 扫描规则库，找出结论与目标G匹配的规则集S
3. 如果S为空：则失败返回。
4. 如果S非空且G未知，重复执行以下操作：
 - ①调用解决冲突算法，从S中选出规则R；
 - ②将规则R的前提部分作为子目标 G' ，
 - ③若 G' 未知，递归调用本算法；
 - ④若 G' 为真，执行R的结论部分，并且从S中删除规则R。

第四节 知识图谱



知识图谱

知识图谱是由Google公司在2012年提出来的一个新的概念。从学术的角度，我们可以对知识图谱给一个这样的定义：“**知识图谱本质上是一种语义网络（Semantic Network）的知识库**”。



语义网络

语义网络（Semantic Web）

语义网络可以看成是一种用于存储知识的数据结构，即基于图的数据结构，这里的图可以是有向图，也可以是无向图。使用语义网络，可以很方便地将自然语言的句子用图来表达和存储，用于机器翻译、问答系统和自然语言理解。

语义网络

下面就举一个例子，比如说，要表示 **John gave a book to Mary** 这样一句话，可以用下面这样的一个语义网络来表示：



知识库

知识库的范围是大于知识图谱的，Google提出的知识图谱也是知识库的一种。构建大规模的知识库一直都是人工智能、自然语言理解等领域的核心任务之一，而知识库的诞生和专家系统密不可分。上世纪八十年代，专家系统获得激增式发展，其核心部分一般由两部分组成：知识库与推理机，人类专家提供知识，再将这种显式的知识映射并存储到知识库中用来推理。

知识图谱的表示-多关系图

知识库和语义网络的概念有点抽象，从实际应用的角度出发其实可以简单地把知识图谱理解成**多关系图**（Multi-relational Graph）。

The screenshot shows a Google search results page for the query "marie curie". The top result is a direct link to the Wikipedia page for Marie Curie. Below it are links to her biography on NobelPrize.org and her profile on the Nobel Foundation website. There are also image search results showing various portraits of Marie Curie.

Marie Curie

Marie Skłodowska-Curie was a French-Polish physicist and chemist famous for her pioneering research on radioactivity. She was the first person honored with two Nobel Prizes—in physics and chemistry. Wikipedia

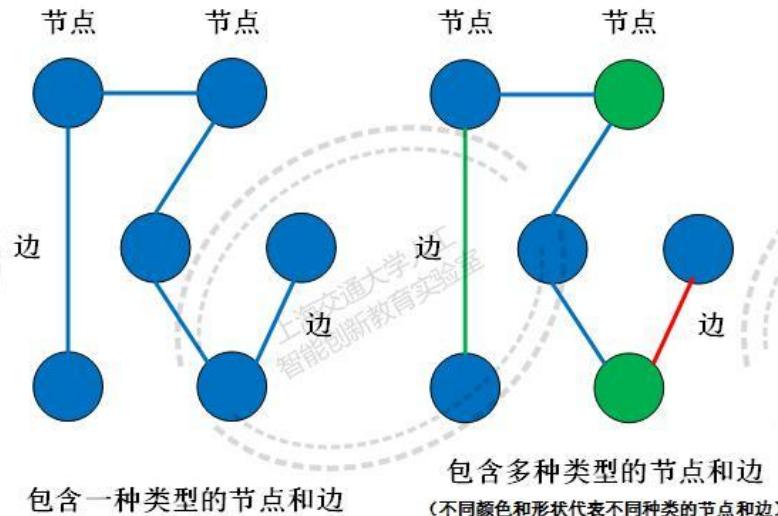
Born: November 7, 1867, Warsaw
Died: July 4, 1934, Sancellemoz
Spouse: Pierre Curie (m. 1895–1906)
Children: Irène Joliot-Curie, Eve Curie
Discovered: Radium, Polonium
Education: Ecole Supérieure de Physique et de Chimie Industrielles de la Ville de Paris, University of Paris

People also search for:

Albert Einstein, Pierre Curie, Ernest Rutherford, Louis Pasteur, John Dalton

知识图谱的表示-多关系图

图是由**节点（Vertex）**和**边（Edge）**来构成。与一般的图只包含一种节点和边不同，**多关系图**一般包含多种类型的节点和多种类型的边。比如下图表示一个经典的图结构，右边的图则表示多关系图，因为图里包含了多种类型的节点和边。这些类型由不同的颜色来标记。



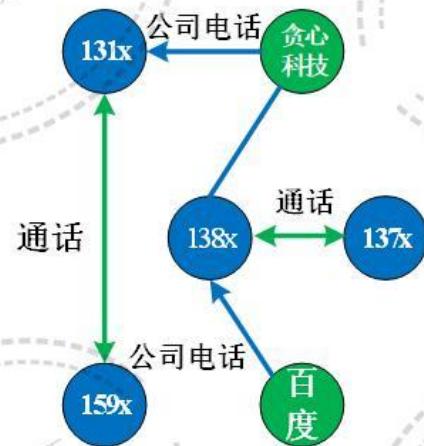
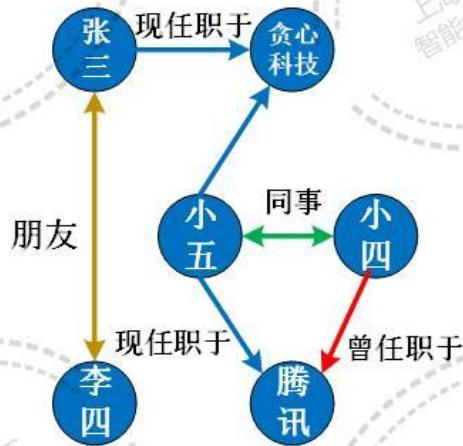
知识图谱的表示-多关系图

在知识图谱里，我们通常用**实体（Entity）**来表达图里的节点、用**关系（Relation）**来表达图里的边。实体指的是现实世界中的事物比如人、地名、概念、药物、公司等，关系则用来表达不同实体之间的某种联系，比如人-“居住在”-北京、张三和李四是“朋友”、逻辑回归是深度学习的“先导知识”等等。

知识图谱的表示-多关系图

现实世界中的很多场景非常适合用知识图谱来表达。比如一个**社交网络图谱**里，我们既可以有“人”的实体，也可以包含“公司”实体。人和人之间的关系可以是“朋友”，也可以是“同事”关系。人和公司之间的关系可以是“现任职”或者“曾任职”的关系。类似的，一个风控知识图谱可以包含“电话”、“公司”的实体，电话和电话之间的关系可以是“通话”关系，而且每个公司它也会有固定的电话。

知识图谱的表示-多关系图



知识图谱的表示-多关系图

知识图谱把大量多关系图构成的知识空间认为是一个知识库，这也是为什么它可以用来回答一些搜索相关问题的原因，比如在 Google 搜索引擎里输入“Who is the wife of Bill Gates?”，我们直接可以得到答案 - “Melinda Gates”。这是因为我们在系统层面上已经创建好了一个包含“Bill Gates”和“Melinda Gates”的实体以及他俩之间关系的知识库。所以，当我们执行搜索的时候，就可以通过关键词提取（"Bill Gates", "Melinda Gates", "wife"）以及知识库上的匹配可以直接获得最终的答案。这种**搜索方式跟传统的搜索引擎是不一样的，一个传统的搜索引擎返回的是网页、而不是最终的答案**，所以就多了一层用户自己筛选并过滤信息的过程。

知识图谱的表示-多关系图

Who is the wife of Bill Gates

All News Images Videos Shopping More Settings Tools

About 24,800,000 results (0.60 seconds)

Bill Gates / Spouse

Melinda Gates

m. 1994



Melinda Ann Gates, DBE is an American philanthropist. She is a former Microsoft employee and co-founder of the Bill & Melinda Gates Foundation. She worked at Microsoft, where she was project manager for Microsoft Bob, Encarta and Expedia. Wikipedia.

People also search for



Bill Gates
Warren Buffet

Paul Allen

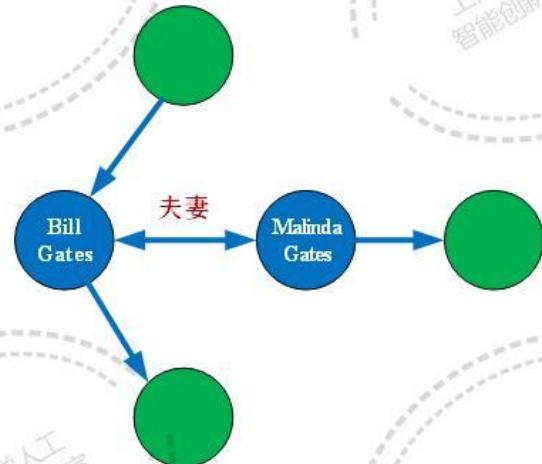
Mark Zuckerberg

Jeff Bezos

William H. Gates Sr.

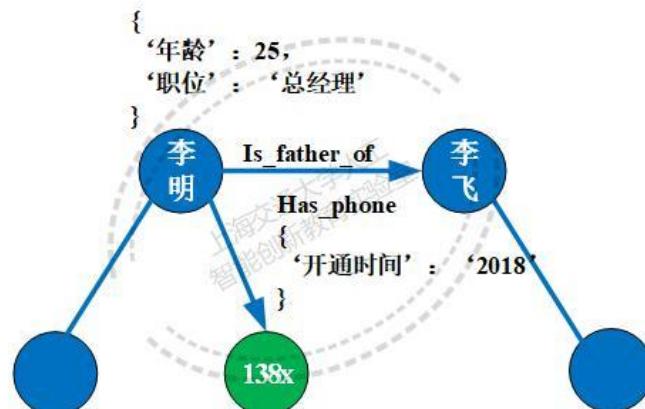
Father-in-law

View 15+ more



知识图谱的表示-属性图

在现实世界中，实体和关系也会拥有各自的属性，比如人可以有“姓名”和“年龄”。当一个知识图谱拥有属性时，我们可以用属性图（Property Graph）来表示。下面的图表示一个简单的属性图。李明和李飞是父子关系，并且李明拥有一个138开头的电话号，这个电话号开通时间是2018年，其中2018年就可以作为关系的属性。类似的，李明本人也带有一些属性值比如年龄为25岁、职位是总经理等。

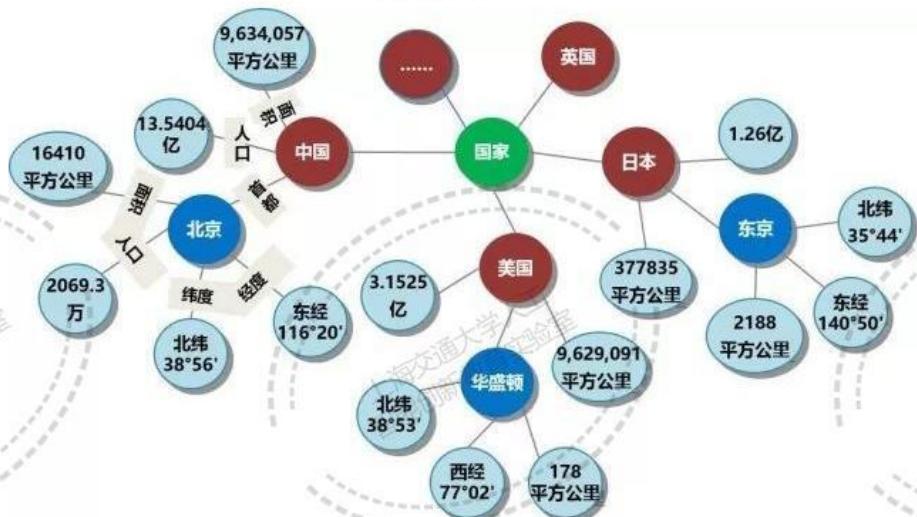


知识图谱的表示-RDF

除了属性图，知识图谱也可以用RDF（Resource Description Framework）来表示，它是由很多的三元组（Triples）来组成。三元组的基本形式主要包括(实体1-关系-实体2)和(实体-属性-属性值)等。每个实体可用一个全局唯一确定的ID来标识，每个属性-属性值对（attribute-value pair，AVP）可用来刻画实体的内在特性，而关系可用来连接两个实体，刻画它们之间的关联。

知识图谱的表示-RDF

如下图的知识图谱例子所示，中国是一个实体，北京是一个实体，**中国-首都-北京** 是一个（实体-关系-实体）的三元组样例。北京是一个实体，人口是一种属性，2069.3万是属性值。北京-人口-2069.3万构成一个（实体-属性-属性值）的三元组样例。



知识图谱的结构-逻辑结构

知识图谱在逻辑上可分为**模式层**与**数据层**两个层次：

数据层由一系列的**事实**组成，也就是知识以事实为单位进行存储。

模式层构建在数据层之上，是知识图谱的核心，通常采用**本体库**来管理知识图谱的模式层。**本体**是**结构化知识库**的概念模板，类似于面向对象思想中的“类”的思想，可以认为本体描述了知识图谱的“类”。通过本体库而形成的知识库不仅层次结构较强，并且冗余程度较小。

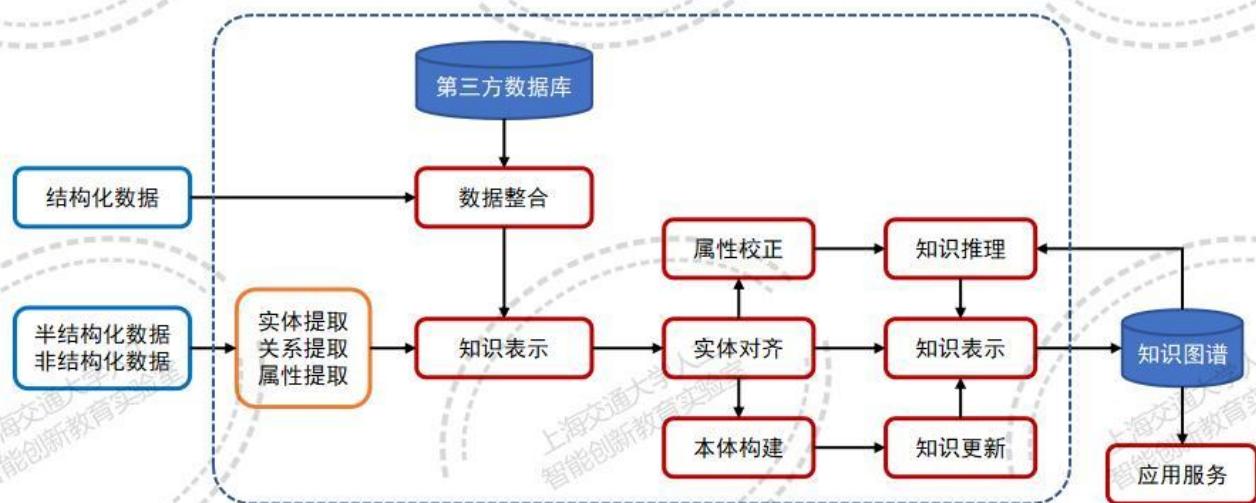
举例：

模式层：(实体-关系-实体), (实体-属性-性值)

数据层：(比尔盖茨-妻子-梅琳达·盖茨), (比尔盖茨-总裁-微软)

知识图谱的结构-技术架构

下面是知识图谱的技术架构。其中虚线框内的部分为知识图谱的构建和更新过程。



知识图谱的结构-技术架构

构建知识图谱是一个迭代更新的过程，根据知识获取的逻辑，每一轮迭代包含四个阶段：

知识提取：从各种类型的数据源中提取出实体、属性以及实体间的相互关系，在此基础上形成本体化的知识表达；

知识表示：采取用某种形式来表示知识；

知识融合：在获得新知识之后，需要对其进行整合，以消除矛盾和歧义，比如某些实体可能有多种表达，某个特定称谓也许对应于多个不同的实体等；

知识加工：对于经过融合的新知识，需要经过质量评估之后（部分需要人工参与甄别），才能将合格的部分加入到知识库中，以确保知识库的质量。

知识图谱的意义

(1) 对Google类的搜索引擎而言

- **找到最想要的信息。**语言可能是模棱两可的——一个搜索请求可能代表多重含义，知识图谱会将信息全面展现出来，让用户找到自己最想要的那种含义。
- **提供最全面的摘要。**有了知识图谱，Google可以更好的理解用户搜索的信息，并总结出与搜索话题相关的内容。
- **让搜索更有深度和广度。**由于知识图谱构建了一个与搜索结果相关的完整的知识体系，所以用户往往会获得意想不到的发现。在搜索中，用户可能会了解到某个新的事实或新的联系，促使其进行一系列的全新搜索查询。

知识图谱的意义

(2) 对广义知识图谱而言

知识图谱是知识库技术的一大进步，虽然它在搜索引擎中率先使用，但是随着大数据时代的到来，知识图谱很快会在日常生活中得到应用。

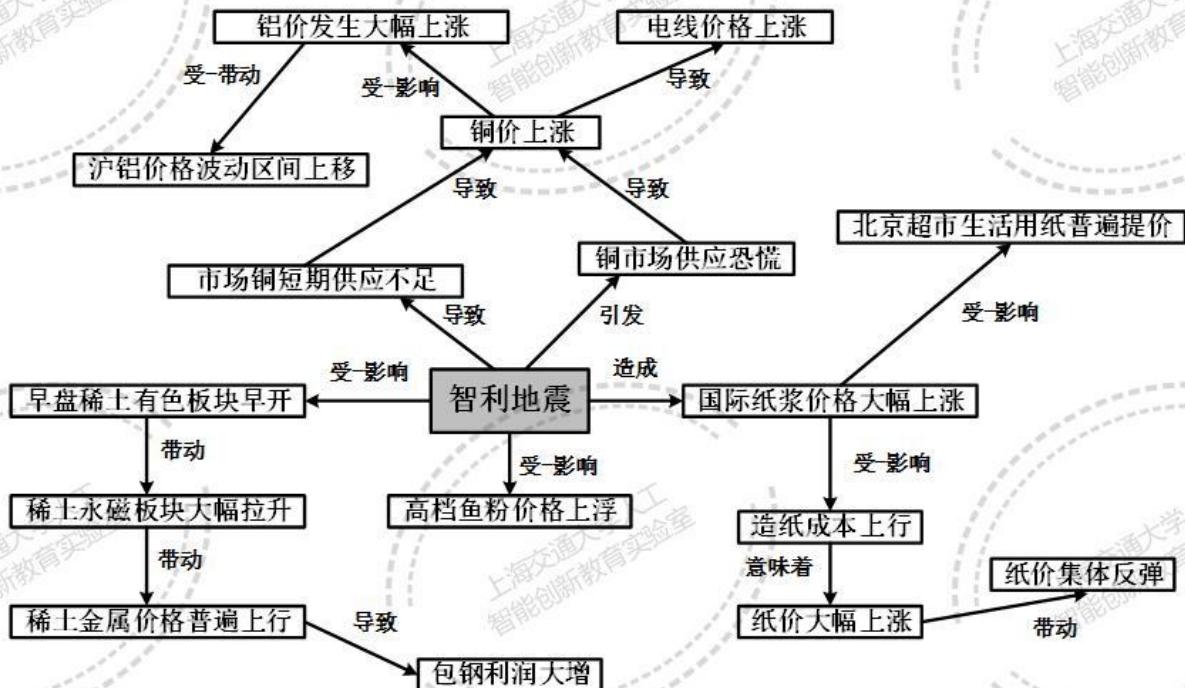
知识图谱**将信息表达成更接近人类认知世界的形式**，提供了组织、管理和理解海量信息的能力，拥有着丰富的应用场景。在出行领域，基于实体识别、实体链指和语义分析技术，整合关联其他外部公开的本体即可搭建起一个巨大的知识网，进一步释放知识图谱和大数据潜能。这无论是在安保、反欺诈、反身份盗用、异常分析还是在用户交互等方面，意义都非常重大。

知识图谱与事理图谱

(1) 事理图谱

事理图谱是新一代知识图谱未来方向的一个重要形态，可以将事理图谱定义为：“**事理图谱是以“事件”为核心的新一代动态知识图谱**，结构上具有抽象概念本体层和实例等多层结构；构成上包括静态实体图谱和动态事件逻辑图谱两部分；功能上注重描述事件及实体在时空域上的丰富逻辑事理关系(顺承、因果、反转、条件、上下位、组成等)；应用上可通过抽象、泛化等技术实现类人脑的知识建模、推理与分析决策”。

知识图谱与事理图谱



知识图谱与事理图谱

(2) 事理图谱与事件预警

事件预警是目前我们使用事理图谱的应用尝试，事件预警是一个面向商品领域的重要资讯预警产品。通过监控上千家全行业网站，实时采集相关资讯，通过抽取识别资讯中的事件，将事件与事理图谱中的事件进行链接，结合情感分析技术、文本标签技术、文本重要性判定技术对具有影响力的资讯进行过滤，最终为用户实现自定义标注的预警资讯筛选以及基于该预警资讯的影响寻迹探索。

知识图谱与事理图谱

19:42 蔬菜价格小幅上涨鸡蛋价格略有回升

本周在市物价局监测的张家口市主要食品价格中，蔬菜价格小幅上涨，鸡蛋价格略有回升，其它品种价格运行稳定。蔬菜价格：本周所监测的19种蔬菜价格与上周相比表现为7升1降11平，总体环比上升2.1%，其中价...

影响：农产品生产价格上涨

食品科技网

事理图谱



19:42 全球70%的央行参与央行数字货币的研究

高雄市将发行区域性商圈货币“高雄币”，核算来自区块链；中国移动正开发区块链净水器，用于收集用户行为数据；柯达授权区块链测试版已处理逾百万美元图片授权.....《得得时刻》分享此时此刻。大家好，今天...

华尔街见闻

19:41 1月9日国内经济动态汇总：2018年国内手机市场出货量同比下降15.6%

【宏观动态】1、国务院总理：要优化科技发展战略布局，调整优化重大科技项目，把基础研究摆在更加突出的位置，加大长期稳定支持，推动基础研究、应用研究和产业化协调发展；要大力营造公平包容的创新创业环境，降...

东方财富

事理图谱



影响：企业订单数量减少

19:41 展望2019只差临门一脚，覆盖地球近半人口的自贸区即可问世

2019年，一个包含约36亿人口的“超级自由贸易圈”料将诞生。在去年11月中旬于新加坡举办的第二次“区域全面经济伙伴关系协定”(RCEP)首脑会议期间，与会各方均认为，RCEP谈判已取得实质性进展...

一财网

19:40 风险情绪回归，原油报不停，黄金没戏？

美原油：基本面：受到中美贸易谈判的乐观情绪影响，增加了原油需求的预期，投资者的风险情绪再次回归。目前，沙特再次放空，表明可能会持续削减原油的出口量，这意味着后期进一步减产的概率极大，对油价也起到了...

新浪财经

事理图谱



影响：有色金属集体飙升 | 美国汽油价格飞涨 | 金价反弹 | 金价刷新日高点多头 | 黄金类资产投资热情高涨 | 原油大涨反...

知识图谱与事理图谱

风险情绪回归，原油涨不停，黄金没戏？

2019-01-09 19:40 新浪财经

美原油：

基本面：

受到中美贸易谈判的乐观情绪影响，增加了原油需求的预期，投资者的风险情绪再次回归。目前，沙特两次放话，表明可能会持续消减原油的出口量，这意味着后期进一步减产的概率极大，对...

10

全屏观看

● 本資訊事件

● 以事件为 中心的新闻



历史相关资讯

- 2018-09-20 皓月贵金属指继续维持疲软态势黄金走势弱势上扬

2018-09-20 美指继续维持疲软态势9.20黄金走势弱势上扬

2018-09-20 黄金投资晨报：黄金再度收复千二关口今晚能否再接再厉

2018-09-20 美元继续维持疲软态势黄金再度收复千二关口

2018-09-20 万凯伦：美元黄金走势反复千二之上多头有望延伸

2018-07-13 美日涨至半年高位。因日本股市上涨削弱避险

2018-07-12 牛汇：美元兑元创半年新高解构背后三大因素

2018-07-12 美元兑元创半年新高，解构背后三大因素

2018-05-25 长江期货：中美贸易关系改善本周金价整体走高(第2...)

2018-10-25 中行(广东省分行)：对美联储加息预期的逼近令黄金多头谨慎

2018-07-18 黄金不为悲袭所劫止正亟亟反弹趋势

2018-07-16 黄金反弹或蓄势待发

2018-06-21 DailyFX：日内黄金和原油行情分析

2018-03-02 投资海外基金渐受追捧

2018-02-16 2月16日英镑机构观点汇总（亚市）

2018-02-16 邦达亚洲：德拉基发表鸽派言论施压欧元小幅下滑

2018-02-16 澳盛集团：美元指数强势回升OPEC减产传闻再度升温

2018-02-16 证金贵金属：避险情绪消退黄金三连阴

2018-02-16 高宏亮：金银高位压力沉重弱势格局或继续

2018-02-16 证金贵金属：险后反弹动能不足警惕基耶斯纳