

第2章 知识表示

- 如何表示知识是人工智能研究的一个重要议题
- 知识表示
 - 以形式化方式表示知识；
 - 供计算机自动处理；
- **【本章内容】**
 - 2.1 概述
 - 基本的知识表示方式
 - 2.2 谓词逻辑表示法★
 - 2.3 产生式表示法
 - 2.4 语义网络表示法★
 - 2.5 框架表示法
 - 补充 知识图谱

2.1 概述

- 知识表示是智能系统的**重要基础**，是人工智能中最活跃的研究部分之一。
- 为了使计算机具有智能，使它能模拟人类的智能行为，就必须使它具有知识。但知识是需要用适当的模式**表示出来才能存储到计算机中去的**，因此关于知识的表示问题就成为人工智能中一个十分重要的研究课题。

□ 知识的概念

知识就是人类认识自然界（包括社会和人）的精神产物，是人类进行智能活动的基础。

知识按其作用可大致分为三类

- **描述性知识**。表示对象及概念的特征及其相互关系的知识，以及问题求解状况的知识，也称为事实性知识。
- **判断性知识**。表示与领域有关的问题求解知识如推理规则等，也称为启发性知识
- **过程性知识**。表示问题求解的控制策略，即如何应用判断性知识进行推理的知识。

按照作用的层次，知识还可以分成以下两类

- **对象级知识**，直接描述有关领域对象的知识，或称为领域相关的知识。
- **元级知识**，描述对象级知识的知识，如关于领域知识的内容、特征、应用范围、可信程度的知识以及如何运用这些知识的知识，也称为关于知识的知识

注意

- **知识**是人通过实践，认识到的客观世界的规律性的东西。
- **知识**在信息的基础上增加了上下文信息，提供了更多的意义，因此也就更加有用和有价值。
- **知识**是随着时间的变化而动态变化的，新的知识可以根据规则和已有的知识推导出来。

□ **知识**是经过加工的信息，它包括事实、信念和启发式规则。

· **事实**：是关于对象和物体的知识。

· **规则**：是有关问题中与事物的行动、动作相联系的因果关系的知识。

- **知识表示**在人工智能体的建造中起到关键作用
 - 以适当方式表示知识，才导致智能体**展示**出智能行为
- **知识表示**是数据结构及其处理机制的综合
 - **知识表示=符号（结构）+处理机制**，其中
 - **恰当的符号（结构）**
 - 用于**存储要解决的问题、可能的中间解答和最终解答以及解决问题涉及的知识**；
 - **配套的处理机制**
 - **仅有符号（结构）不能体现出系统具有知识**；
 - **只有对其作适当的处理才构成意义**。

AI对知识表示方法的要求

- (1) **表示能力**，要求能够正确、有效地将问题求解所需要的各类知识都表示出来。
- (2) **可理解性**，所表示的知识应易懂、易读。
- (3) **便于知识的获取**，使得智能系统能够渐进地增加知识，逐步进化。
- (4) **便于搜索**，表示知识的符号结构和推理机制应支持对知识库的高效搜索，使得智能系统能够迅速地感知事物之间的关系和变化；同时很快地从知识库中找到有关的知识。
- (5) **便于推理**，要能够从已有的知识中推出需要的答案和结论。

2.2 谓词逻辑表示法

用形式逻辑（尤其是一阶谓词逻辑）表示知识是AI 研究中提出使用的一种普遍方法。

1. 命题逻辑和谓词逻辑

命题逻辑和谓词逻辑是最先应用于人工智能的两种逻辑，谓词逻辑是在命题逻辑基础上发展起来的，命题逻辑可以看作是谓词逻辑的一种特殊形式。

(1) 命题

[定义1] 命题是具有真假意义的语句。

命题代表人们进行思维时的一种判断，或者是肯定，或者是否定，只有这两种情况。若命题的意义为真，则称它的真值为真。记作T；若命题的意义为假，则称它的真值为假，记作F。一个命题不能同时既为真又为假，但在一定条件下为真，在另一条件下为假。

例如：“北京是中华人民共和国的首都”，

“ $3 < 5$ ” 都是真值为T的命题。

“太阳从西边升起”，“煤球是白的”都是真值为 **F** 的命题。

“**1+1=10**”在二进制情况下是真值为 **T** 的命题，在十进制下是真值为 **F** 的命题。

在命题逻辑中，命题通常用大写的英文字母表示：

例如，可以用 **P** 表示“西安是个古老的城市”。

命题有两种类型：

(1) **原子命题**：不能分解成更简单的陈述语句，称为**原子命题**。

(2) **复合命题**：由联结词、标点符号和原子命题等复合构成的命题，称为**复合命题**。

注意：所有这些命题都应具有确定的真值。

了解几个概念

命题常量：如果一个命题标识符表示确定的命题，就称为**命题常量**。

命题变元：如果命题标识符只表示任意命题的位置标志，就称为**命题变元**。

注意：

- (1) 因为命题变元可以表示任意命题，所以它不能确定真值，故**命题变元不是命题**。
- (2) 当命题变元P用一个特定的命题取代时，P才能确定真值，这时也称为**对P进行指派**。
- (3) 当命题变元表示原子命题时，该变元称为**原子变元**。

□ 命题逻辑的符号包括以下几种：

- (1) 命题常元：True (T) 和 False (F)；
- (2) 命题符号：P、Q、R、T等；
- (3) 联结词：① \neg ； ② \wedge ； ③ \vee ；
④ \rightarrow ； ⑤ \longleftrightarrow 。
- (4) 括号：()。

命题逻辑主要使用这5个联结词，通过这些联结词，可以由简单的命题构成复杂的复合命题。

连词优先级别是 \neg ， \wedge 、 \vee ， \Rightarrow 、 \Leftrightarrow ，但可通过括号改变优先级。

- \neg : 否定(Negation), 复合命题 $\neg Q$ 表示否定Q的真值的命题, 即“非Q”
- ② \wedge : 合取(Conjunction), 复合命题 $P \wedge Q$ 表示P和Q的合取, 即“P与Q”
- ③ \vee : 析取(Disjunction), 复合命题 $P \vee Q$ 表示P或Q的析取, 即“P或Q”

- ④ \rightarrow : 条件(Condition),也叫蕴含, 复合命题 $P \rightarrow Q$ 表示命题P是命题Q的条件,即“如果P,那么Q”
- ⑤ \leftrightarrow : 双条件(Bicondition),也叫等价, 复合命题 $P \leftrightarrow Q$ 表示命题P、命题Q相互作为条件,即“如果P,那么Q; 如果Q,那么P”

注意: 可以用真值表的方法表明联结词的功能: ★

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
F	F	T	F	F	T	T
F	T	T	F	T	T	F
T	F	F	F	T	F	F
T	T	F	T	T	T	T

命题这种表示法有较大的局限性，它无法把它所描述的客观事物的结构及逻辑特征反映出来，也不能把不同事物的共同特征描述出来。

例如：对“老李是小李的父亲”这一命题，若用英文字母 P 表示，怎么也看不出老李与小李的父子关系。

由于这些原因，在命题逻辑的基础上，发展起来了谓词逻辑。

谓词逻辑★

谓词逻辑：根据对象和对象上的谓词（即对象的属性和对象之间的关系），通过使用**连接词**和**量词**来表示世界。

主要思想：世界是由对象组成的，可以由标识符和属性来区分它们。在这些对象中，还包含着相互的关系。

注意：

- 在命题逻辑中，每个表达式都是句子，表示事实。
- 在谓词逻辑中，有句子，但是也有项，表示对象。常量符号、变量和函数符号用于表示项，量词和谓词符号用于构造句子。

1 语法

· 一阶谓词演算

- 标点符号、括号、逻辑联结词、常量符号集、变量符号集、 n 元函数符号集、 n 元谓词符号集、量词

· 谓词演算

- 合法表达式（原子公式、合式公式），表达式的演算化简方法，标准式（合取的前束范式或析取的前束范式）

语法元素

常量符号。

变量符号。

函数符号。

谓词符号。

联结词： \neg 、 \wedge 、 \vee 、 \rightarrow 、 \leftrightarrow 。（和命题相同）

量词：全称量词 \forall 、存在量词 \exists 。 \forall 和 \exists 后面跟着的 x 叫做量词的指导变元。

□ 量词★

■ 全称量词 \forall

□ 符号 $(\forall x)P(x)$: 表示对于某个论域中的**所有（任意一个）**个体 x , 都有 $P(x)$ 真值为 T 。

■ 存在量词 \exists

□ 符号 $(\exists x)P(x)$: 来表示某个论域中**至少存在一个**个体 x , 使 $P(x)$ 真值为 T 。

函数符号与谓词符号

- 若函数符号 f 中包含的个体数目为 n , 则称 f 为 n 元函数符号。
- 若谓词符号 P 中包含的个体数目为 n , 则称 P 为 n 元谓词符号。

如: $\text{father}(x)$ 是一元函数, $\text{Less}(x, y)$ 是二元谓词.
一般一元谓词表达了个体的性质, 而多元谓词表达了个体之间的关系.

谓词的阶

如果谓词P中的所有个体都是个体常量、变元、或函数，则该谓词为**一阶谓词**。

如果谓词P中某个个体本身又是一个一阶谓词，则称P为**二阶谓词**。

余者类推。

个体变元的取值范围称为**个体域**。个体域可以是有限的，也可以是无限的。把各种个体域综合在一起作为讨论的范围的域称为**全总个体域**。

- 在一阶谓词逻辑中，称 $\text{Teacher}(\text{father}(\text{Wang}))$ 中的 $\text{father}(\text{Wang})$ 为项，项可定义如下：
- **定义：**项可**递归**定义如下：
 - (1) 单独一个个体是项（包括常量和变量）。
 - (2) 若 f 是 n 元函数符号，而 t_1, \dots, t_n 是项，则 $f(t_1, \dots, t_n)$ 是项。
 - (3) 任何项仅由规则 (1) (2) 所生成。

原子公式

若 P 为 n 元谓词符号, t_1, \dots, t_n 都是项, 则称 $P(t_1, \dots, t_n)$ 为**原子公式**, 简称**原子**。

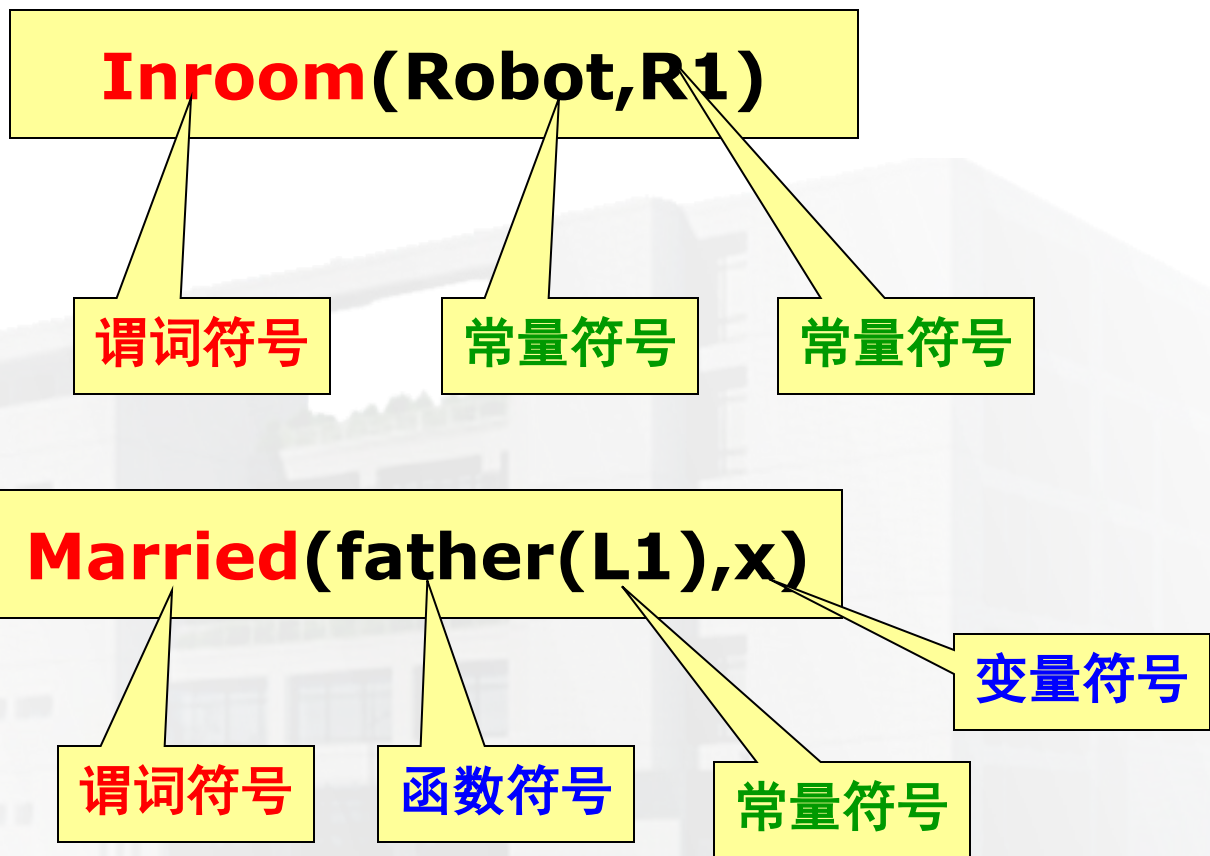
在原子中, 若 t_1, \dots, t_n 都不含变量, 则 $P(t_1, \dots, t_n)$ 是**命题**。

注意:

谓词逻辑可以由原子和5种逻辑连接词, 再加上量词来构造复杂的符号表达式。这就是所谓的谓词逻辑中的**公式**。

□ 原子公式举例

用**括号**和**逗号**隔开，以表示论域内的**关系**



Inroom(**Robot**,**R1**)

谓词符号

常量符号

常量符号

Married(**father**(**L1**),**x**)

谓词符号

常量符号

谓词符号、常量符号——首字母大写的形式来表示★

Married(**father**(L1),**x**)

函数符号

变量符号

函数符号、变量符号——小写字母的形式来表示★

□ 一阶谓词逻辑的合式公式（可简称公式）可递归定义如下：

(1) 原子谓词公式是合式公式（也称为原子公式）。

(2) 若 P 、 Q 是合式公式，则 $(\neg P)$ 、 $(P \wedge Q)$ 、 $(P \vee Q)$ 、 $(P \rightarrow Q)$ 、 $(P \leftrightarrow Q)$ 也是合式公式。

(3) 若 P 是合式公式， x 是任一个体变元，则 $(\forall x)P$ 、 $(\exists x)P$ 也是合式公式。

(4) 任何合式公式都由有限次应用(1)、(2)、(3)来产生。

□ 一阶谓词逻辑公式的**解释**：

设 D 为谓词公式 P 的非空个体域, 若对 P 中的个体常量、函数、谓词按如下规定赋值：

- (1) 为每个个体常量指派 D 中的一个元素。
- (2) 为每个 n 元函数指派一个从 D^n 到 D 的映射, 其中

$$D^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in D\}$$

- (3) 为每个 n 元谓词指派一个从 D^n 到 $\{T, F\}$ 的映射。

则称这些指派为公式 P 在 D 上的一个**解释**。

4. 注意:

(1) 在谓词逻辑中, 由于公式中可能含有个体常量、个体变元以及函数, 因此不能像命题公式那样直接通过真值指派给出解释, 必须首先考虑个体常量、和函数在个体域中的取值, 然后才能针对常量和函数的具体取值为谓词分别指派真值。

(2) 在给出一阶逻辑公式的一个解释时, 需要规定两件事情: 公式中个体的定义域和公式中出现的常量、函数符号、谓词符号的定义。

设个体域 $D = \{1, 2\}$ ，求公式

$$G = (\forall x)(\exists y)P(x, y)$$

在 D 上的解释，并指出在每一种解释下公式 G 的真值。

解：由于公式 G 没有包含个体常量和函数，因此可以直接为谓词指派真值，设

$P(1,1)$	$P(1,2)$	$P(2,1)$	$P(2,2)$
T	F	T	F

这就是公式G在D上的一个解释。从这个解释可以看出：

当 $x=1, y=1$ 时, $P(x, y)$ 的真值为T;

当 $x=2, y=1$ 时, $P(x, y)$ 的真值也为T;

即对 x 在D上任意取值, 都存在 $y=1$, 使得 $P(x, y)$ 的真值为T。因此, 在该解释下, 公式G的真值为T。

值得注意的是:

一个谓词公式在其个体域上的解释不是唯一的。
例如, 对公式G, 若给出另一组真值指派如下:

$P(1,1)$	$P(1,2)$	$P(2,1)$	$P(2,2)$
T	T	F	F

这也是公式G在D上的一个解释。从这个解释可以看出：

当 $x=1, y=1$ 时, $P(x, y)$ 的真值为T;

当 $x=2, y=1$ 时, $P(x, y)$ 的真值也为F;

同样

当 $x=1, y=2$ 时, $P(x, y)$ 的真值为T;

当 $x=2, y=2$ 时, $P(x, y)$ 的真值也为F;

即对 x 在 D 上任意取值，不存在一个 y ，使得 $P(x, y)$ 的真值为 T 。因此，在该解释下，公式 G 的真值为 F 。

实际上， G 在 D 上共有**16种**解释，这里就不一一列举了。

注意：

一个公式的解释通常有**任意多个**，由于个体域 D 可以随意规定，而对一个给定的个体域 D ，对公式中出现的常量、函数符号和谓词符号的定义也是随意的，因此公式的真值都是针对某一个解释而言，它可能在某一个解释下为真，而在另一个解释为假。

5. 谓词逻辑适应范围

谓词逻辑适合于表示事物的状态、属性、概念等**事实性知识**，也可以用来表示事物间具有确定因果关系的**规则性知识**。

1) 对**事实性知识**：可以使用谓词公式中的析取符号与合取符号连接起来的谓词公式来表示，如对下面句子★：

张三是一名计算机系的学生，他喜欢编程序。

可以用谓词公式表示为

$\text{Computer}(\text{张三}) \wedge \text{Like}(\text{张三}, \text{Programming})$

其中： $\text{Computer}(x)$ 表示 x 是计算机系的学生， $\text{Like}(x, y)$ 表示 x 喜欢 y ，都是谓词。

2) 对**规则性知识**：通常使用由蕴涵符号连接起来的谓词公式来表示★，例如，对于

如果x，则y

用谓词公式表示为

$$x \rightarrow y$$

复习

注意: 可以用真值表的方法表明联结词的功能: ★

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
F	F	T	F	F	T	T
F	T	T	F	T	T	F
T	F	F	F	T	F	F
T	T	F	T	T	T	T

复习

Inroom(**Robot**,**R1**)

谓词符号

常量符号

常量符号

Married(**father**(**L1**),**x**)

谓词符号

常量符号

谓词符号、常量符号——首字母大写的形式来表示★

复习

Married(**father**(L1),**x**)

函数符号

变量符号

函数符号、变量符号——小写字母的形式来表示★

6. 谓词逻辑表示步骤★

从上述两个例子我们总结一下用谓词公式表示知识的一般步骤如下：

- (1) 定义谓词及个体，确定每个谓词及个体的确切含义；
- (2) 根据所要表达的事物或概念，为每个谓词中的变元赋以特定的值；
- (3) 根据所要表达的知识的语义，用适当的连接符将各个谓词连接起来形成谓词公式。

7. 谓词逻辑表示知识的举例★

例1：用谓词逻辑表示下列知识：

武汉是一个美丽的城市，但她不是一个沿海城市。

如果马亮是男孩，张红是女孩，则马亮比张红长得高。

解：按照知识表示步骤，用谓词公式表示上述知识。

第一步：定义谓词如下：

$BCity(x)$ ： x 是一个美丽的城市

$HCity(x)$ ： x 是一个沿海城市

$Boy(x)$ ： x 是男孩

$Girl(x)$ ： x 是女孩

$High(x, y)$ ： x 比 y 长得高

这里涉及的个体有：武汉（Wuhan），马亮（Ma l），张红（Zhangh）

第二步 将这些个体代入谓词中，得到
BCity(Wuhan)，HCity(Wuhan)，Boy(Ma l)，
Girl(Zhangh)，High(Ma l, Zhangh)

第三步 根据语义，用逻辑连接符将它们连接起来，就得到了表示上述知识的谓词公司。

$BCity(Wuhan) \wedge \sim HCity(Wuhan)$

$(Boy(Ma l) \wedge Girl(Zhangh)) \rightarrow High(Ma l, Zhangh)$

例2 用谓词逻辑表示下列知识：

所有学生都穿彩色制服。

任何整数或者为正数或者为负数。

自然数都是大于零的整数。

解：首先定义谓词如下：

Student(x) : x是学生 Uniform(x, y) : x穿y

N(x) : x是自然数 I(x) : 是整数 P(x) : x是正数

Q(x) : x是负数 L(x) : x大于零

按照第二步和第三步的要求，上述知识可以用谓词公式分别表示为：

$$(\forall x) (\text{Student}(x) \rightarrow \text{Uniform}(x, \text{Color}))$$

$$(\forall x) (I(x) \rightarrow P(x) \vee Q(x))$$

$$(\forall x) (N(x) \rightarrow L(x) \wedge I(x))$$

课堂测试

1. 每个人都有父亲
2. 所有的教师都有自己的学生

作答

课堂测试

将以下语句：

- (1) 会朗读者是识字的，
- (2) 海豚都不识字，
- (3) 有些海豚是很机灵的，
- (4) 有些很机灵的东西不会朗读。

形式化表示为合适公式。

注：令谓词 $R(x)$ 、 $L(x)$ 、 $D(x)$ 、 $I(x)$ 分别指示朗读、识字、海豚和机灵

禁止使用逻辑符号 \neg 以上公式中字母

作答

每个人都有一个父亲

■ 定义谓词：PERSON(x),表示x是人
HASFATHER(x,y),表示x有父亲y

■ 谓词表示：

$$(\forall x)(\exists y)(\text{PERSON}(x) \rightarrow \text{HASFATHER}(x,y))$$

所有的教师都有自己的学生

■ 定义谓词:

TEACHER(x),表示x是教师

STUDENT(y),表示y是学生

TEACHES(x,y),表示x是y的老师

■ 谓词表示:

$(\forall x)(\exists y)(\text{TEACHER}(x) \rightarrow \text{TEACHES}(x,y) \wedge \text{STUDENT}(y))$

将以下语句：

- (1) 会朗读者是识字的，
- (2) 海豚都不识字，
- (3) 有些海豚是很机灵的，
- (4) 有些很机灵的东西不会朗读。

形式化表示为合适公式。

令谓词R、L、D、I分别指示朗读、识字、海豚和机灵,则这些语句可形式化表示如下:

$$(1) (\forall_x)[R(x) \Rightarrow L(x)]$$

$$(2) (\forall_x)[D(x) \Rightarrow \neg L(x)]$$

$$(3) (\exists_x)[D(x) \wedge I(x)]$$

$$(4) (\exists_x)[I(x) \wedge \neg R(x)]$$

8.一阶谓词逻辑表示法的特点

1. 一阶谓词逻辑表示法的优点

- (1) **严密性:**可以保证其演绎推理结果的正确性，可以较精确地表达知识。
- (2) **自然性:**它的表现方式和人类自然语言非常接近。
- (3) **通用性:**拥有通用的逻辑演算方法和推理规则。
- (4) **知识易表达:**如果对逻辑的某些外延扩展后，则可以把大部分精确性知识表达成一阶谓词逻辑的形式。
- (5) **易于实现:**用它表示的知识易于模块化，便于知识的增删及修改，便于在计算机上实现。

2. 一阶谓词逻辑表示法的缺点

- (1) **效率低**:由于推理是根据形式逻辑进行的,把推理演算和知识含义截然分开,抛弃了表达内容所含的语义信息,往往是推理过程太冗长,降低系统效率。另一方面,谓词表示越细,表示越清楚,推理越慢、效率越低。
- (2) **灵活性差**:不便于表达和加入启发性知识和元知识。不便于表达不确定性的指示,但人类的知识大都具有不确定性和模糊性,这是使得它表示知识的范围受到了限制。
- (3) **组合爆炸**:在其推理过程中,随着事实数目的增大及盲目的使用推理规则,有可能产生组合爆炸。

2.3. 产生式表示法

- 产生式表示的起源：
 - 美国数学家波斯特（Post），1943年，产生式系统， 计算模型，
 - 以称为**产生式的规则描述符号串**替代运算(Post机的计算模型)
 - 用于描述形式语言的语法，表示人类心理活动的认知过程等。
- 现代产生式系统：
 - 与波斯特的模型已很不相同，
 - 基本概念相同，都使用产生式规则表示知识。
 - 便于模拟人求解问题的思维方式，系统模块性强，易于修改扩充，得到广泛应用；
 - 目前大多数专家系统（尤其是中小型系统）都采用产生式系统的结构方式来建立。 DENDRAL, MYCIN

在产生式系统中，把推理和行为的过程用产生式规则表示，所以又称**基于规则的系统**。

□ 1. 产生式规则

■ 通常用于表示事物间的**因果关系**；

■ 【基本形式】

□ IF **P** then **Q** 或 **P** → **Q**，其中

□ **P**表示规则的**条件**（或称**前提**）；

■ **谓词**、多元组、常量、变量、关系运算.....

□ **Q**表示规则激活时应该执行的**动作**（或得到的**结论**）；

■ **激活**——规则条件**P**满足；

■ 【规则分类】

□ ①**前提-结论型**

□ ②**条件-动作型**

规则的表示

一般地，一个规则由**前项**和**后项**两部分组成。**前项**表示前提条件，各个条件由逻辑连接词（合取、析取等）组成各种不同的组合。**后项**表示当前前提条件为真时，应采取的行为或所得的结论。产生式系统中每条规则是一个“前提→结论”或“条件→结论”的产生式，其简单形式为：

IF 〈前提〉 THEN 〈结论〉

IF 〈条件〉 THEN 〈动作〉

为了严格地描述产生式，下面用**巴科斯范式**给出它的形式描述和语义：

〈规则〉 ::= 〈前提〉 → 〈结论〉

〈前提〉 ::= 〈简单条件〉 | 〈复合条件〉

〈结论〉 ::= 〈事实〉 | 〈动作〉

〈复合条件〉 ::= 〈简单条件〉 And 〈简单条件〉 [(And 〈简单条件〉) ...] | 〈简单条件〉 Or 〈简单条件〉 [(Or 〈简单条件〉) ...]

〈动作〉 ::= 〈动作名〉 [(〈变元〉, ...)]

事实的表示

1. 确定性和不确定性规则知识的产生式表示

确定性规则知识可用前面介绍的产生式的简单形式表示即可。

不确定性规则知识对基本形式作一定的扩充，用如下形式表示

$P \rightarrow Q$ （可信度）

或者 $IF\ P\ THEN\ Q$ （可信度）

其中，P是产生式的**前提或条件**，用于指出该产生式是否是可用的条件；Q是一组结论或动作，用于指出该产生式的前提条件P被满足时，应该得出的结论或因该执行的操作。这一表示形式主要在不肯定推理中当已知事实与前提中的条件不能精确定匹配时，只要按照“可信度”的要求达到一定的相似度，就认为已知事实与前提条件匹配，再按照一定的算法将这种可能性（或不确定性）传递到结论。

2. 确定性和不确定性事实性知识的产生式表示

确定性事实性知识一般使用三元组

(对象, 属性, 值) 或

(关系, 对象1, 对象2)

来表示, 其中对象就是语言变量, 这种表示的机器内部实现就是一个表。如事实“老李年龄是35岁”, 便可以表示成

(Lee, Age, 35)

其中, Lee是事实性知识涉及的对象, Age是该对象的属性, 而35岁是该对象属性的值。而老李、老张是朋友, 可表示成

(Friend, Lee, Zhang)

而有些事实性知识带有**不确定性**和**模糊性**，若考虑不确定性，这种知识就可以用四元组的形式表示如下

(对象, 属性, 值, 不确定度量值) 或

(关系, 对象1, 对象2, 不确定度量值)

如不确定性事实性知识 “老李年龄可能是35岁”，这里老李是35岁的可能性取90%，便可以表示成

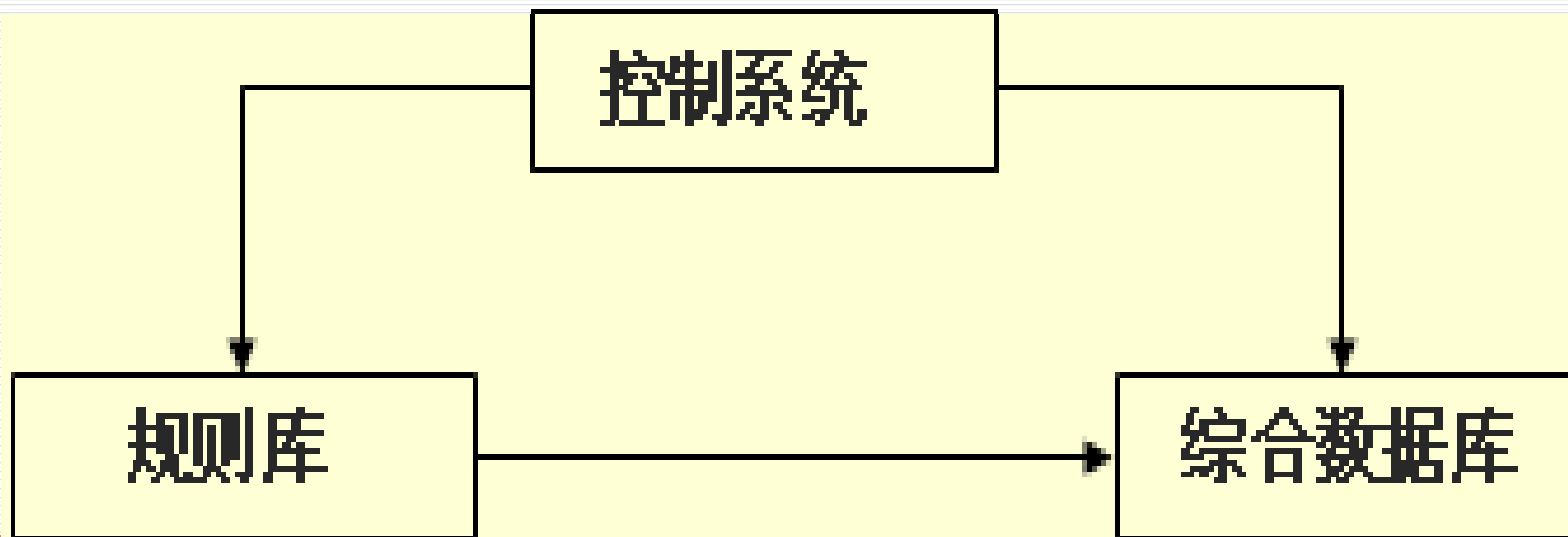
(Lee, Age, 35, 0.9)

而老李、老张是朋友的可能性不大，这里老李、老张是朋友的可能性取20%，可表示成

(Friend, Lee, Zhang, 0.2)

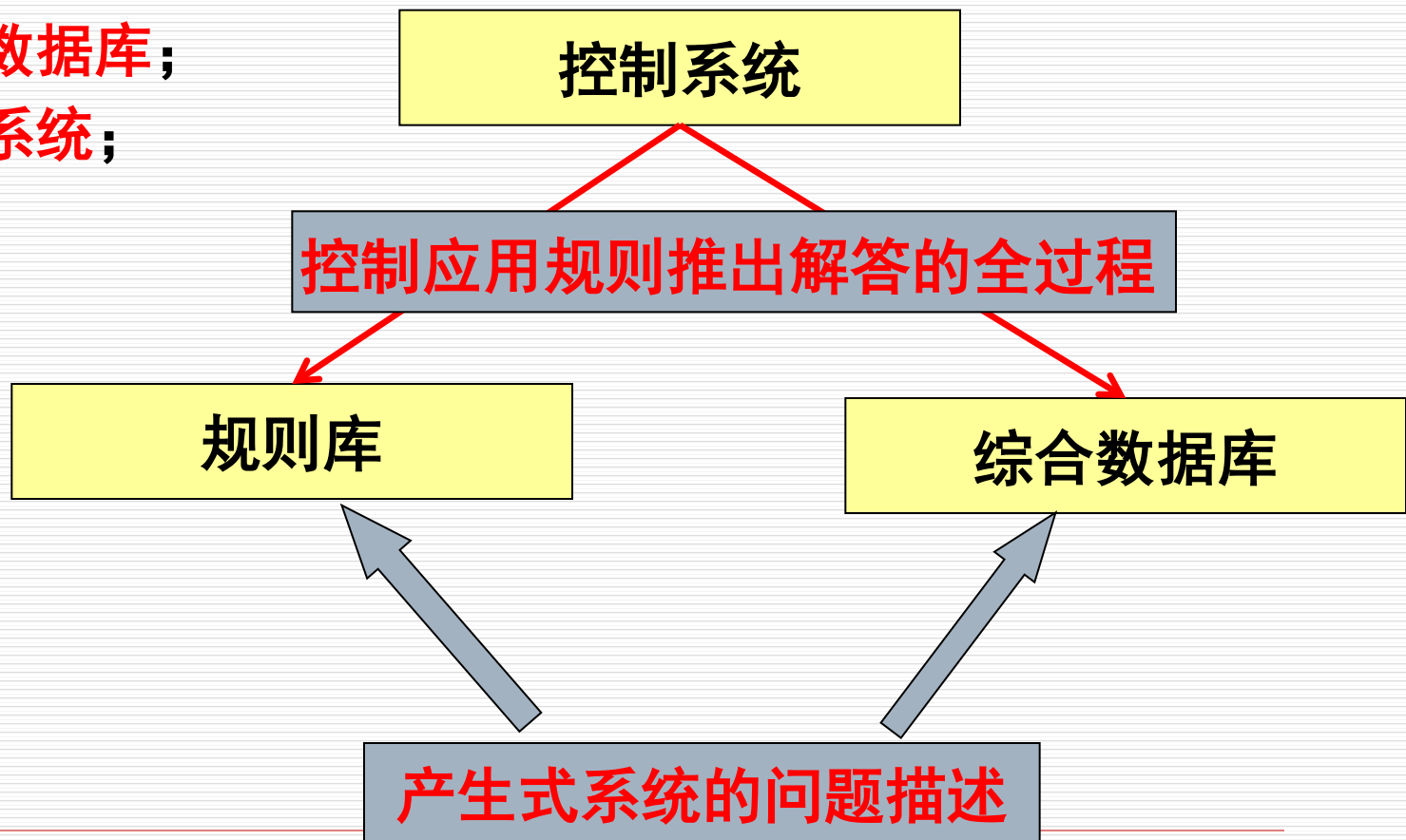
产生式系统的组成★

把一组产生式放在一起，让他们互相配合，协同作用，一个产生式生成的结论可以供另一个产生式作为已知事实使用，以求得问题的解决，这样的系统称为产生式系统。一般说来，一个产生式系统由以下三个基本部分组成：



■ 3个部分组成：

- 规则库；
- 综合数据库；
- 控制系统；



1. 规则库:

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。

规则库是专家系统的核心，也是一般产生式系统赖以进行问题求解的基础，其中知识的完整性和一致性、知识表达的准确性和灵活性以及知识组织的合理性，都将对产生式系统的性能和运行效率产生直接的影响。

2. 综合数据库：

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)和最后结果的工作区。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。

3. 推理机:

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。

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规则。

产生式系统的推理方式

产生式系统推理机的推理方式有**正向推理**、**反向推理**和**双向推理**三种。

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**。它的推理过程是：

- (1) 规则库中的规则与综合数据库中的事实进行匹配，得到匹配的规则集合；
 - (2) 使用冲突解决算法，从匹配规则集合中选择一条规则作为启用规则；
 - (3) 执行启用规则的操作部分，将该启用规则的操作结果送入综合数据库或对综合数据库进行必要的修改。
- 重复这个过程直至达到目标。

反向推理

反向推理是从目标（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**。其推理过程是：

- （1）规则库中的规则后件与目标事实进行匹配，得到匹配的规则集合；
 - （2）使用冲突解决算法，从匹配规则集合中选择一条规则作为启用规则；
 - （3）将启用规则的前件作为子目标。
- 重复这个过程直至各子目标均为已知事实，则反向推理的过程就算成功结束。

双向推理

双向推理是一种既**自顶向下**又**自底向上**的推理。推理从两个方向同时进行，直至某个中间界面上两方向结果相符便成功结束。不难想象，这种双向推理较正向或反向推理所形成的推理网络来得小，从而推理效率更高。

□ 例题

■ 动物识别系统的规则库。

这是一个用以识别虎、金钱豹、斑马、长颈鹿、企鹅、鸵鸟、信天翁等7种动物的产生式系统。为了实现对这些动物的识别，该系统建立了如下规则库：

R₁: IF 该动物有毛 THEN 该动物是哺乳动物

R₂: IF 该动物有奶 THEN 该动物是哺乳动物

R₃: IF 该动物有羽毛 THEN 该动物是鸟

R₄: IF 该动物会飞 AND 会下蛋 THEN 该动物是鸟

R₅: IF 该动物吃肉 THEN 该动物是食肉动物

R₆: IF 动物有犬齿 AND 有爪 AND 眼盯前方
THEN 该动物是食肉动物

R7: IF 该动物是哺乳动物 AND 有蹄 THEN 该动物是有蹄类动物

R8: IF 该动物是哺乳动物 AND 是嚼反刍动物
THEN 该动物是动物有蹄类动物

R9: IF 该动物是哺乳动物 AND 是食肉动物
AND 是黄褐色 AND 身上有暗斑点
THEN 该动物是金钱豹

R10: IF 该动物是哺乳动物 AND 是食肉动物
AND 是黄褐色 AND 身上有黑色条纹
THEN 该动物是虎

R11: IF 该动物是有蹄类动物 AND 有长脖子
AND 有长腿 AND 身上有暗斑点
THEN 该动物是长颈鹿

R₁₂: IF 该动物是有蹄类动物 AND 身上有黑色条纹
THEN 该动物是斑马

R₁₃: IF 该动物是鸟 AND 有长脖子 AND 有
长腿 AND 不会飞 THEN 该动物是鸵鸟

R₁₄: IF 该动物是鸟 AND 会游泳
AND 不会飞 AND 黑白二色
THEN 该动物是企鹅

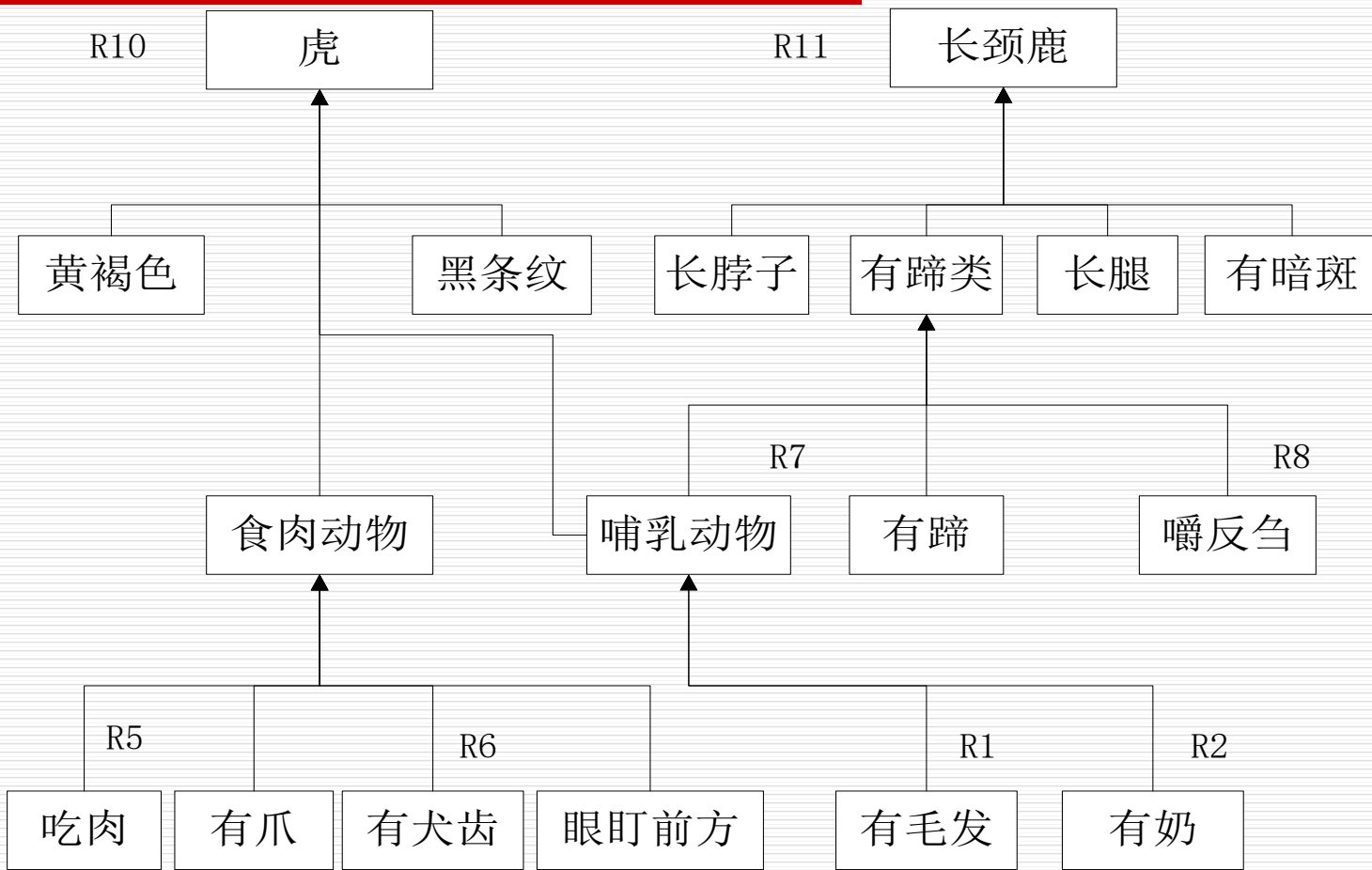
R₁₅: IF 该动物是鸟 AND 善飞
THEN 该动物是信天翁

在上例中, R_1-R_{15} 分别是对各产生式规则所做的编号, 以便于对它们的引用. 同时, 我们还看出, 虽然该系统是用来识别7种动物的, 但它并没有简单地设计7条规则, 而是设计了15条.

识别动物基本想法是: 首先根据一些比较简单的条件, 如“有毛发” “有羽毛” “会飞” 等对动物进行比较粗的分类, 如“哺乳动物” “鸟类” 等, 然后随着条件的增多, 逐步缩小分类范围, 最后给出分别识别7种动物的规则.

这样做有下列好处:

- (1) 当已知的事实不完全时, 虽不能推出最终结论, 但可以得到分类结果;
- (2) 当需要增加对其它动物(如牛马等)的识别时, 规则中只需增加关于这些动物个性方面的知识, 如 R_9-R_{15} 那样, 而对 R_1-R_{10} 可直接利用, 这样增加的规则就不会太多;
- (3) 由上述规则, 很容易形成各种动物的推理链, 例如虎及长颈鹿的推理过程如图所示.



产生式系统的优缺点

优点:

- (1) **清晰性**：产生式表示格式固定、形式简单，规则（知识单位）间相互较为独立，没有直接关系，使知识库的建立较为容易，处理较为简单。
 - (2) **模块性**：知识库与推理机是分离的，这种结构给知识库的修改带来方便，无须修改程序，对系统的推理路径也容易做出解释。
 - (3) **自然性**：符合人类的思维习惯，是人们常用的一种表达因果关系的知识表示形式，既直观自然，又便于推理。另外，产生式表示法既可以表示确定性知识，又可以表示不确定性知识，更符合人们日常见到的问题。
- 因此，产生式方法是当今最流行的专家系统设计模式。

缺点或不足：

(1) **难以扩展**：尽管规则形式上相互独立，但实际问题中往往彼此是相关的。这样当知识库不断扩大时，要保证新的规则和已有的规则没有矛盾就会越来越困难，知识库的一致性越来越难以实现。

(2) **规则选择效率较低**：在推理过程中，每一步都要和规则库中的规则做匹配检查。如果知识库中规则数目很大，显然效率会降低。

(3) **控制策略不灵活**：产生式系统往往采用单一的控制策略，如顺序考察规则库中的每一条规则，这同样会降低系统的效率。

(4) **知识表示形式单一**：产生式系统比较适合于表示非结构化的知识，对于结构化的知识可能用语义网络或框架或面向对象的表示方式更为合适。

注意：以纯粹的产生式系统表示复杂的知识结构比较困难，因此发展了一系列知识的结构化表示方法，如框架和语义网络等，知识以这种形式表示的系统，称为**基于知识的系统**。

结构化表示

□ 需求：

- 世界是复杂的，充满万物，事物间存在着千丝万缕的联系；
- 一阶谓词逻辑离散地表示事物和事物间的关系，表示法效用低；
- 结构化的手段高效和集中地描述特定的事物和事物间的关系。

□ 主要内容：

- 语义网络
 - 事物间的关系
- 框架表示法
 - 事物内部的结构

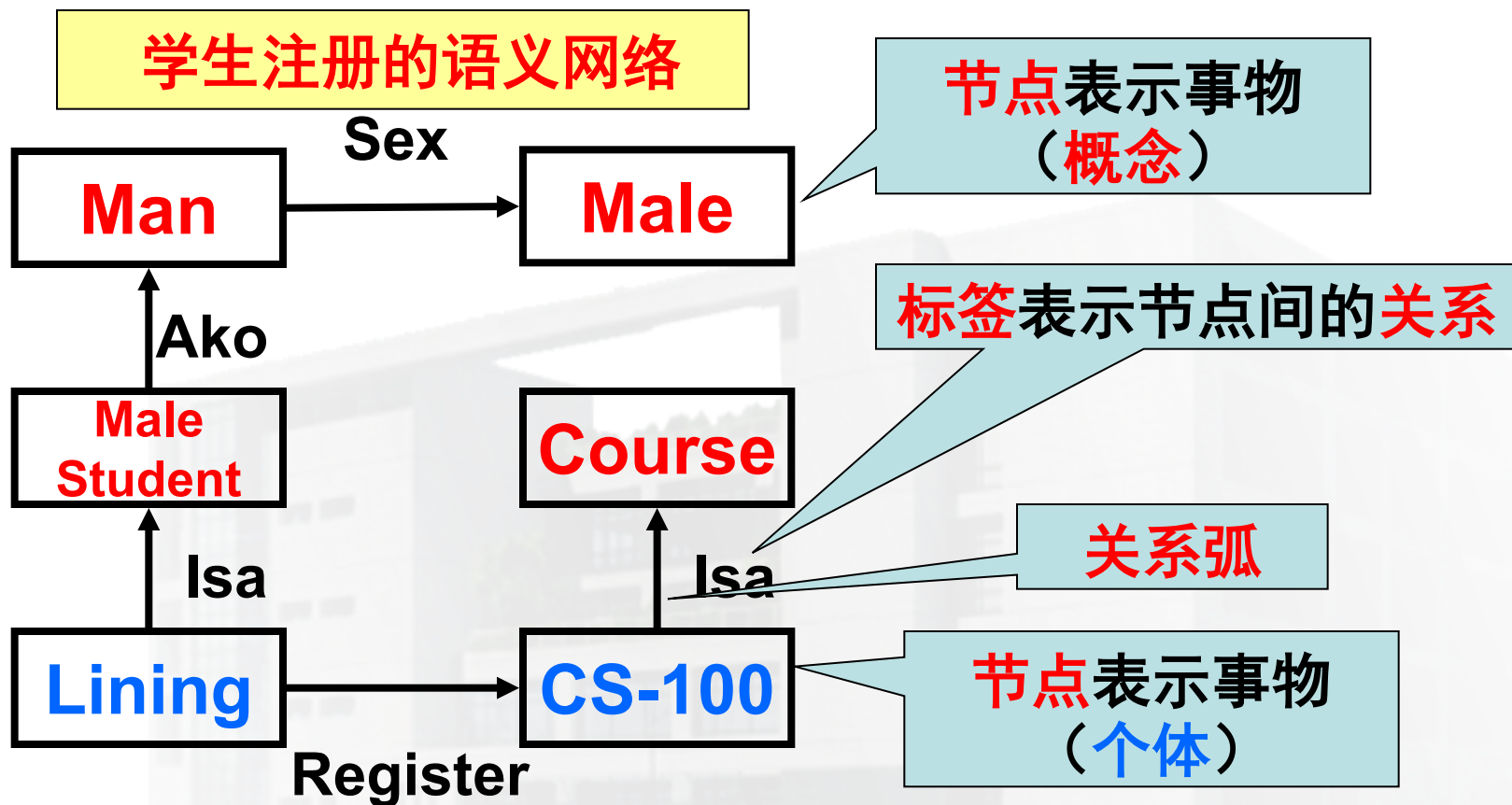
2.4 语义网络

Semantic Network

- **Quillian**（奎廉），M. R. (1968). *Semantic memory*. In Minsky, M., editor, *Semantic Information Processing*, pages 216--270. MIT Press.
- 模拟人对事物的认识，表示事物之间的关系。
- 主要应用于自然语言理解系统中。
- 强大和直观的表示能力，广泛应用。
- 命题语义网络、数据语义网络、语言语义网络等。

1、语义网络的一般概念

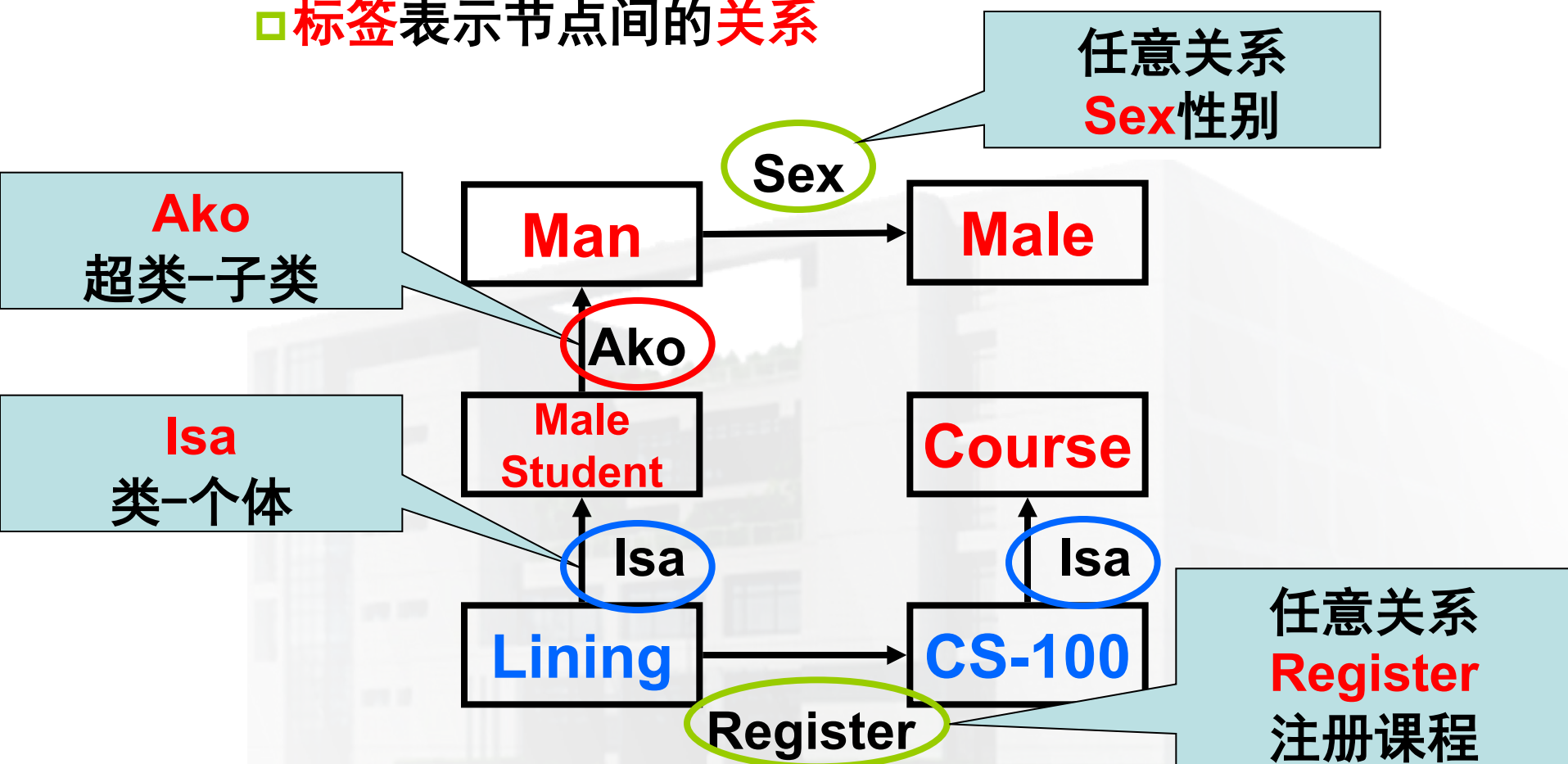
■ 表示为描述事物间关系的有向图



1、语义网络的一般概念

表示为描述事物间关系的有向图

标签表示节点间的关系



从结构上来看，语义网络一般由一些最基本的语义单元组成。这些最基本的语义单元被称为语义基元，可用如下三元组来表示为

(节点1, 弧, 节点2)

可用如图2.3所示的有向图来表示。其中A和B分别代表节点，而R则表示A和B之间的某种语义联系。

当把多个语义基元用相应的语义联系关联在一起的时候，就形成了一个语义网络。如图2.4所示。

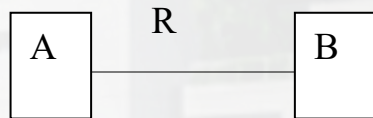


图2.3 语义基元结构

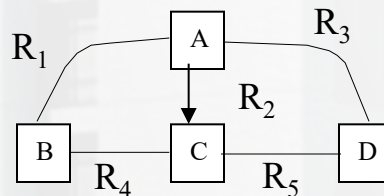


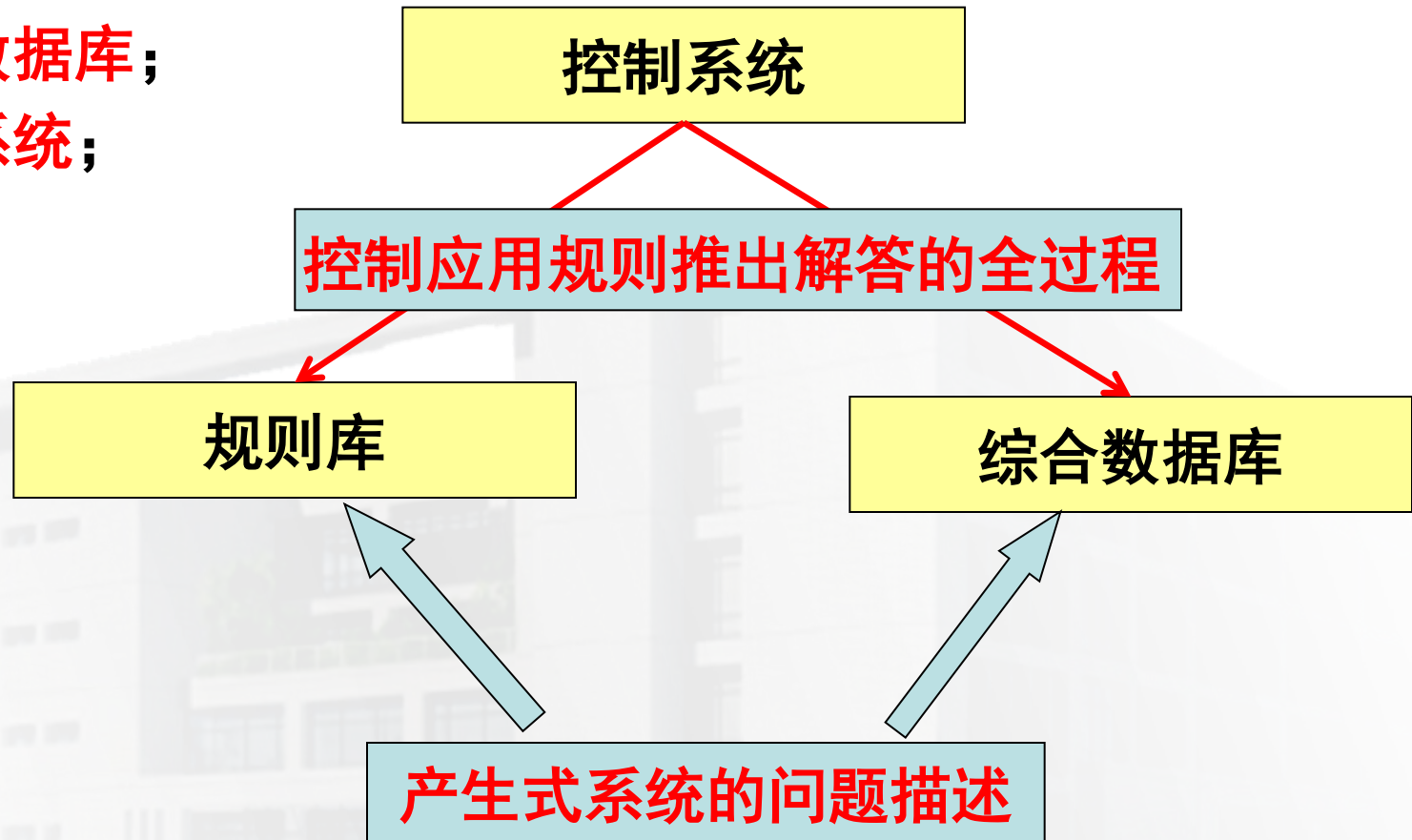
图2.4 语义网络结构

□ 1. 产生式规则

- 通常用于表示事物间的**因果关系**；
- **【基本形式】**
 - IF **P** then **Q** 或 **P** → **Q**，其中
 - **P**表示规则的**条件**（或称**前提**）；
 - **谓词**、多元组、常量、变量、关系运算……
 - **Q**表示规则激活时应该执行的**动作**（或得到的**结论**）；
 - **激活**——规则条件**P**满足；
- **【规则分类】**
 - ①**前提-结论型**
 - ②**条件-动作型**

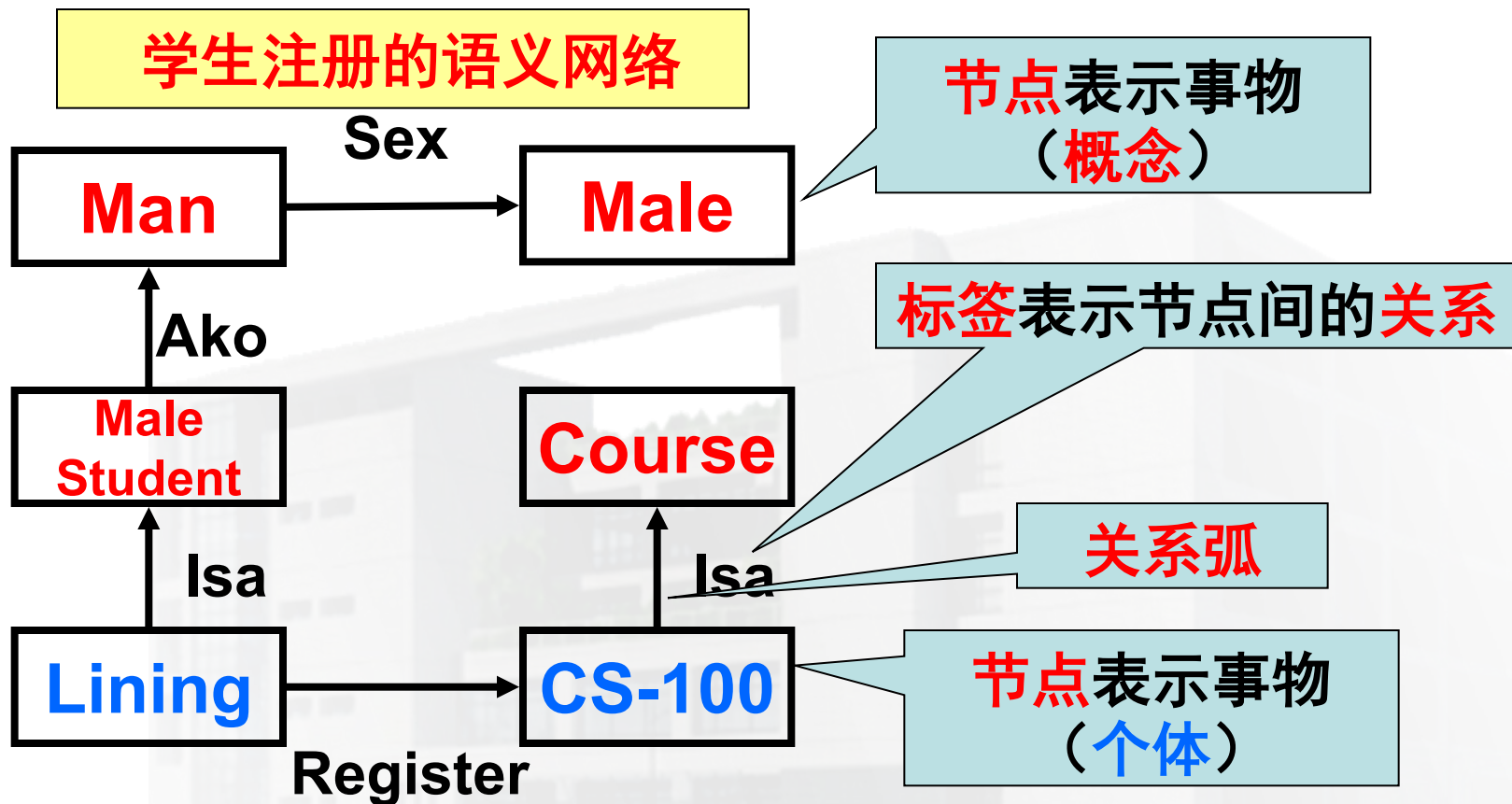
■ 产生式系统的3个部分组成：

- 规则库；
- 综合数据库；
- 控制系统；



1、语义网络的一般概念

- 表示为描述事物间关系的有向图



语义网络的基本语义联系

语义网络除了可以描述事物本身之外，还可以描述事物之间的错综复杂的关系。基本语义联系是构成复杂语义联系的基本单元，也是语义网络表示知识的基础，因此从一些基本的语义联系组合成任意复杂的语义联系是可以实现的。这里只给出一些经常使用的最基本语义关系。

1. 类属关系

类属关系是指具有共同属性的不同事物间的分类关系、成员关系或实例关系，它体现的是“具体与抽象”、“个体与集体”的层次分类。其直观意义是“是一个”，“是一种”，“是一只”……。在类属关系中，其一个**最主要特征是属性的继承性**，处在具体层的结点可以继承抽象层结点的所有属性。常用的类属关系有：

AKO(A-Kind-of):表示一个事物是另一个事物的一种类型。 ★

AMO(A-Member-of):表示一个事物是另一个事物的成员。

ISA(Is-a):表示一个事物是另一个事物的实例。 ★

3. 属性关系

属性关系是指事物和其属性之间的关系。常用的属性关系有：

Have：表示一个结点具有另一个结点所描述的属性。★

Can：表示一个结点能做另一个结点的事情。★
例如，“鸟有翅膀”，“电视机可以放电视节目”。
其对应的语义网络表示如图2.8所示。



图2.8 属性关系实例

4. 时间关系

时间关系是指不同事件在其发生时间方面的先后关系，节点间的**不具备属性继承性**。常用的时间关系有：

Before：表示一个事件在一个事件之前发生。

After：表示一个事件在一个事件之后发生。

例如，“香港回归之后，澳门也会回归了”，“王芳在黎明之前毕业”。其对应的语义网络表示如图2.9所示



图2.9 时间关系实例

5. 位置关系

位置关系是指不同事物在位置方面的关系。节点间的不具备属性继承性。常用的位置关系有：

Located-on:表示一物体在另一物体之上。

Located-at:表示一物体在某一位置。

Located-under:表示一物体在另一物体之下。

Located-inside:表示一物体在另一物体之中。

Located-outside:表示一物体在另一物体之外。

例如，“华中师范大学坐落于桂子山上”，其对应的语义网络表示如图2.10所示。

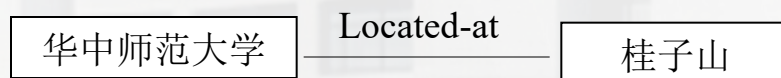


图2.10 位置关系实例

6. 相近关系

相近关系,又称相似关系,是指不同事物在形状、内容等方面相似和接近。常用的相近关系有:

Similar-to:表示一事物与另一事物相似。

Near-to:表示一事物与另一事物接近。

例如,“狗长得像狼” 其对应的语义网络表示如图2.11所示。

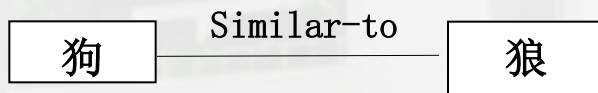


图2.11 相近关系实例

7. 因果关系

因果关系是指由于某一事件的发生而导致另一事物的发生，适合表示规则性知识。通常用If-then联系表示两个节点之间的因果关系，其含义是“如果……，那么……”。例如，“如果天晴，小明骑自行车上班”，其对应的语义网络如图2.12所示。



图2.12 因果关系实例

8. 组成关系

组成关系一种一对多的联系，用于表示某一事物由其它一些事物构成，通常用Compsloed-of联系表示。Compsloed-of联系所连接的节点间**不具备属性继承性**。例如，“整数由正整数、负整数和零组成”可用图2.13表示。

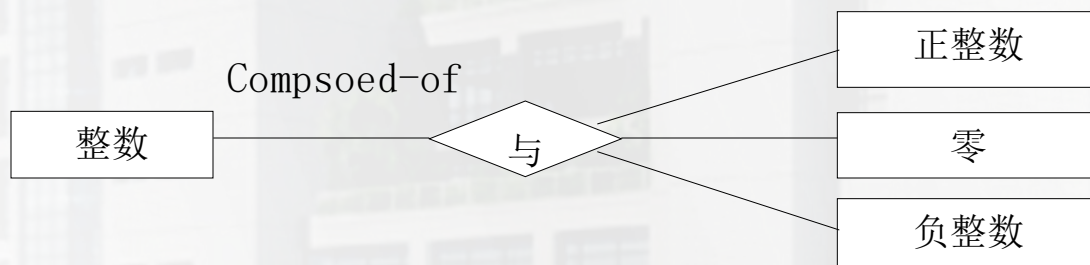


图2.13 组成关系实例

1. 事实性知识的表示

对于一些简单的事实，例如“鸟有翅膀”，“轮胎是汽车的一部分”，这里要描述这些事实需要两个节点，用前面给出的基本语义联系或自定义的基本语义联系就可以表示了。对于稍微复杂一点的事实，比如在一个事实中涉及到多个事物时，如果语义网络只被用来表示一个特定的事物或概念，那么当有更多的实例时，就需要更多的语义网络，这样就使问题复杂化了。

通常把有关一个事物或一组相关事物的知识用一个语义网络来表示。

例如，用一个语义网络来表示事实“苹果树是一种果树，果树又是树的一种，树有根、有叶”。

这一事实涉及“苹果树”、“果树”和“树”这3个对象，树两个属性“有根”、“有叶”。首先建立“苹果树”节点，为了进一步说明苹果树是一种果树，增加一个“果树”节点，并用AKO联系连接着两个节点。为了说明果树是树的一种，增加一个“树”节点，并用AKO联系连接着两个节点。为了进一步描述树“有根”、“有叶”的属性，引入两个“根”节点和“叶”节点，并分别用HAVE联系与“树”节点连接。这个事实可用如图2.14所示的语义网络表示。



图2.14 有关苹果树的语义网络

2. 情况、动作和事件的表示

为了描述那些复杂的知识，在语义网络的知识表示法中，通常采用引进**附加节点**的方法来解决。西蒙（Simon）在提出的表示方法中增加了**情况节点、动作节点和事件节点**，允许用一个节点来表示情况、动作和事件。

(1) 情况的表示

在用语义网络表示那些**不及物动词**表示的语句或**没有间接宾语的及物动词**表示的语句时，如果该语句的动作表示了一些其它情况，如动作作用的时间等，则需要增加一个**情况节点**用于指出各种不同的情况。

例如，用语义网络表示知识“请在2006年6月前归还图书”。这条知识只涉及到一个对象就是“图书”，它表示了 在2006年6月前“归还”图书这一种情况。为了表示归还的时间，可以增加一个“归还”节点和一个情况节点，这样不仅说明了归还的对象是图书，而且很好地表示了归还图书的时间。其语义网络表示如图2.15所示。归还图书情况2006年6月AKO, Object, Before, 图2.15为带有情况节点的语义网络:

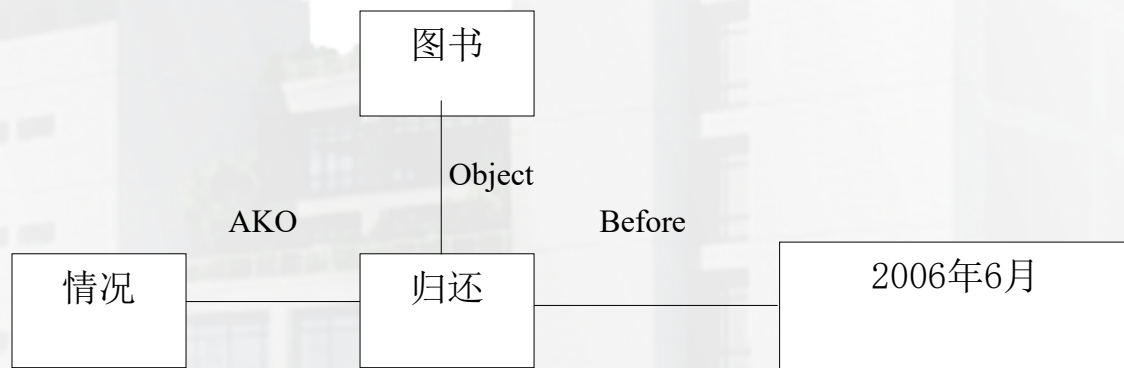


图2.15 带有情况节点的语义网络

(2) 动作的表示

有些表示知识的语句既有发出动作的主体，又有接受动作的客体(其实质表示的是一个三元关系)。在用语义网络表示这样的知识时，可以增加一个动作节点用于指出动作的主体和客体。

例如：用语义网络表示知识“校长送给李老师一本书”。这条知识只涉及到两个对象就是“书”和“校长”，为了表示这个事实，增加一个“送给”节点。其语义网络表示如图2.16所示。

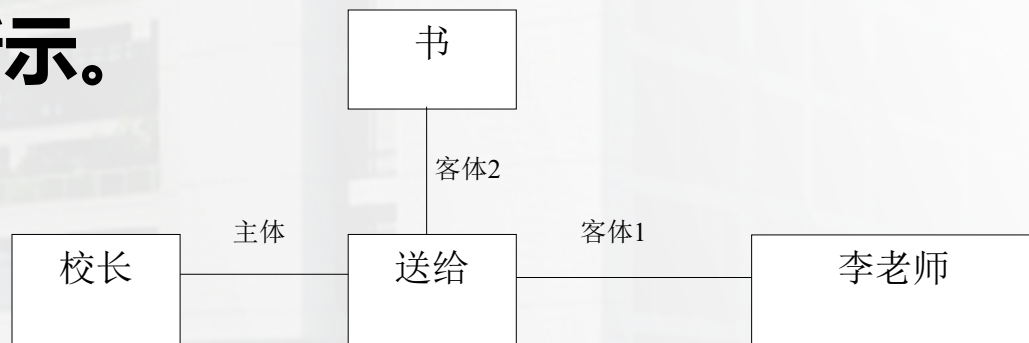


图2.16 带有动作节点的语义网络

- 关系弧只能表示二元关系

- 多元语义网络表示的实质

通过增加事件节点，把多元关系转化为一组二元关系的组合，或二元关系的合取。

可转换为

$$R(X_1, X_2, \dots, X_n)$$

$$R_{12}(X_1, X_2) \wedge R_{13}(X_1, X_3) \wedge \dots \wedge R_{1n}(X_1, X_n)$$

$$\dots \dots \dots$$
$$R_{n-1\ n}(X_{n-1}, X_n)$$

(3) 事件的表示

如果要表示的知识可以看成是发生的一个事件，那么可以增加一个事件节点来描述这条知识。

例如：用语义网络表示知识“中国队与日本队两国的国家足球队在中国进行一场比赛，结局的比分是3:2”。其语义网络表示如图2.17所示。



图2.17 带有事件节点的语义网络

3. 连词和量词的表示★

在稍微复杂一点的知识中，经常用到象“并且”、“或者”、“所有的”、“有一些”等这样的联结词或量词，在谓词逻辑表示法中，很容易就可以表示这类知识。而谓词逻辑中的连词和量词可以用语义网络来表示。因此，语义网络也能表示这类知识。

(1) 合取与析取的表示

当用语义网络来表示知识时，为了能表示知识中体现出来的“合取与析取”的语义联系，可通过增加**合取节点**与**析取节点**来表示。只是在使用时要注意其语义，不应出现不合理的组合情况。

例如，对事实“参观者有男有女，有年老的，有年轻的”。可用图2.18所示的语义网络表示。其中，A、B、C、D分别代表4种情况的参观者。

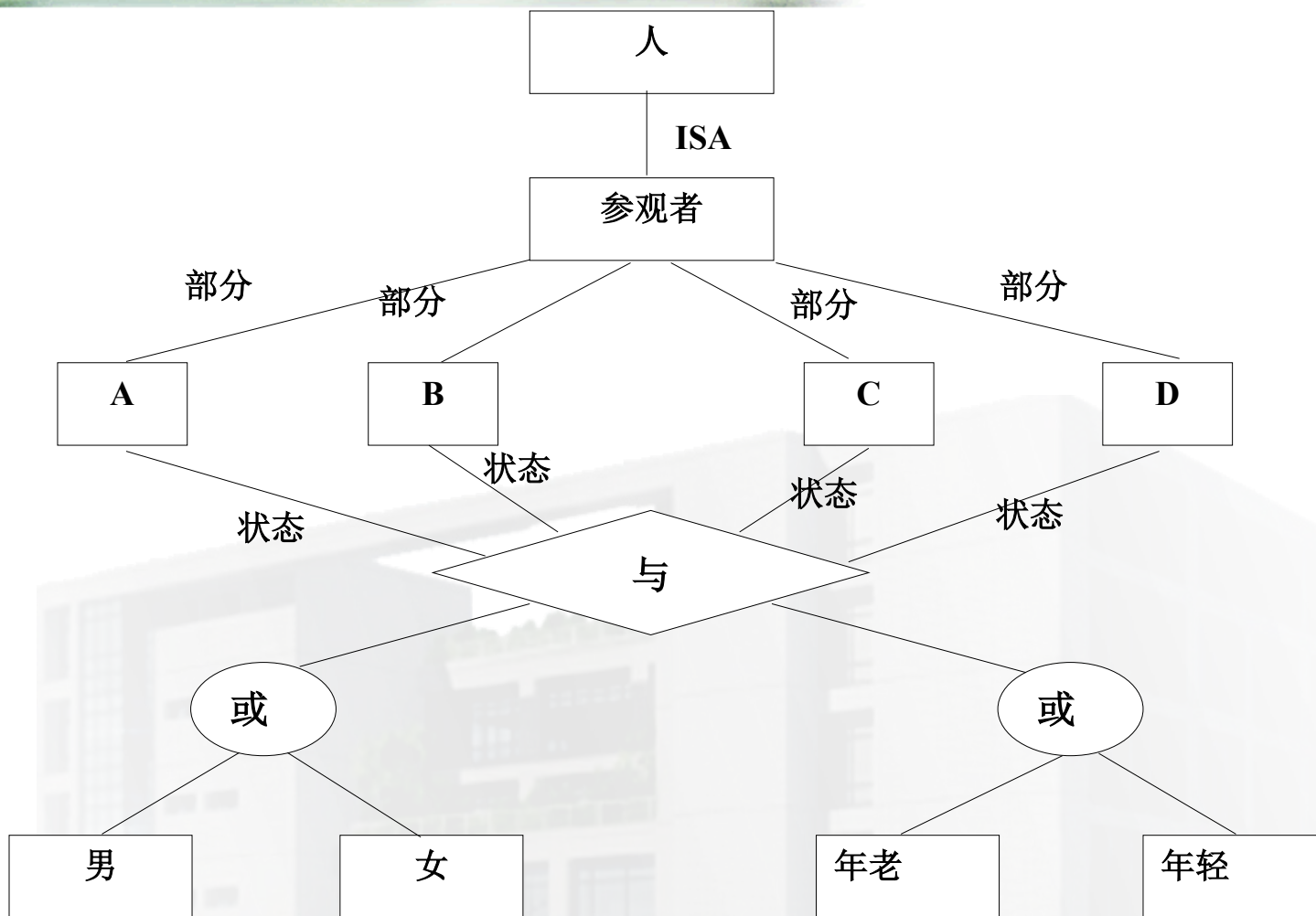


图2.18 具有合取与析取关系的语义网络

(2) 存在量词与全称量词的表示★

在用语义网络表示知识时，对存在量词可以直接用“是一种”、“是一个”等语义关系来表示。

对全称量词可以采用亨德里克(G.G.Hendrix)提出的语义网络分区技术来表示，也称为分块语义网络(Partitioned Semantic Net)，以解决量词的表示问题。该技术的基本思想是：把一个复杂的命题划分成若干个子命题，每一个子命题用一个简单的语义网络来表示，称为一个子空间，多个子空间构成一个大空间。每个子空间看作是大空间中的一个节点，称为超节点。空间可以逐层嵌套，子空间之间用弧相互连接。

例如，对事实“每个学生都学习了一门外语”，可用图2.19所示的语义网络表示。

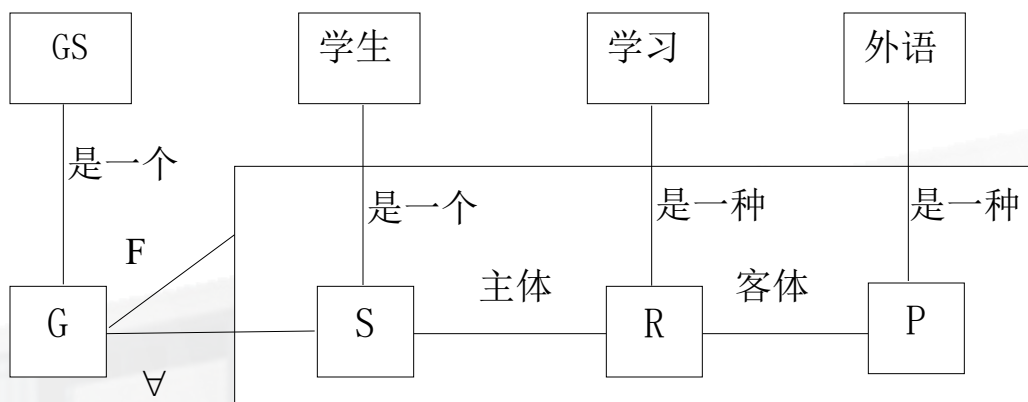


图2.19 具有全称变量的语义网络（分块语义网络）

其中G代表整个陈述句，它是一般陈述句GS的一个实例。G中的每一个元素至少有两个特性：Form(F)，即句中的关系和全称量词(\forall)。在这个例子中只有一个变量S具有全称量词，Form中其余两个变量R, P看成具有存在量词。

另一个例子，对事实“**每个学生都学习了每门外语**”，只需对图2.19作简单的修改，唯一要做的是用 \forall 链与节点P相连，如图2.20所示的语义网络表示。

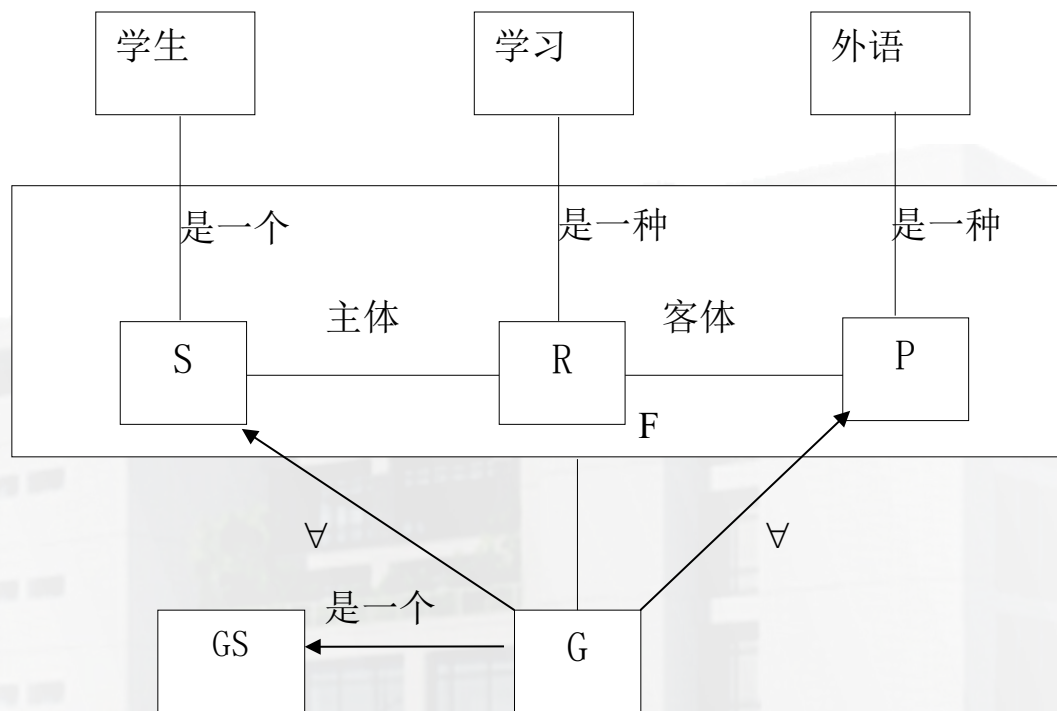


图2.20 全称变量在语义网络中的表示

4.用语义网络表示知识的步骤如下：★

- (1) 确定问题中所有对象和每个对象的属性。
- (2) 确定所讨论对象间的关系。
- (3) 根据语义网络中所涉及的关系，对语义网络中的节点及弧进行整理，包括**增加节点**、弧和归并节点等。

① 在语义网络中，如果节点中的联系是ISA、AKO、AMO等类属关系，则下层节点对上层节点具有属性继承性。整理同一层节点的共同属性，并抽出这些属性，加入上层节点中，以免造成信息冗余。

② 如果要表示的知识中含有因果关系，则增加情况节点，并从该节点引出多条弧将原因节点和结果节点连接起来。

③ 如果要表示的知识中含有动作关系，则增加动作节点，并从该节点引出多条弧将动作的主体节点和客体节点连接起来。

④ 如果要表示的知识中含有“与”和“或”关系时，可在语义网络中增加“与”节点和“或”节点，并用弧将这些“与”、“或”与其它节点连接起来表示知识中的语义关系。

(或者采用虚线框法)

⑤ 如果要表示的知识是含有全称量词和存在量词的复杂问题，则采用前面介绍的亨德里克(G .G .Hendrix)提出的语义网络分区技术来表示。

⑥ 如果要表示的知识是规则性的知识，则应仔细分析问题中的条件与结论，并将它们作为语义网络中的两个节点，然后用IF-THEN弧将它们连接起来。

(4) 将各对象作为语义网络的一个节点，而各对象间的关系作为网络中各节点的弧，连接形成语义网络。

语义网络知识表示举例★

例 把下列命题用一个语义网络表示出来：

- (1) 猪和羊都是动物；
- (2) 猪和羊都是哺乳动物；
- (3) 野猪是猪，但生长在森林中；
- (4) 山羊是羊，头上长着角；
- (5) 绵羊是一种羊，它能生产羊毛。

解题分析

问题涉及的对象有猪、羊、动物、哺乳动物、野猪、山羊、绵羊、森林、羊毛、角等。

然后分析它们之间的语义关系，“动物”和“哺乳动物”、“哺乳动物”和“猪”、“哺乳动物”和“羊”、“羊”和“山羊”及“绵羊”、“野猪”和“猪”之间的关系是“是一种”的关系，可用**AKO**来表示。“山羊”和“头上有角”之间是一种属性关系，可用**HAVE**来描述；“绵羊”和“羊毛”之间是一种属性关系，可用**HAVE**来描述；“野猪”和“森林”之间是位置关系，可用**Locate-at**来表示。其语义网络如图2.21所示。

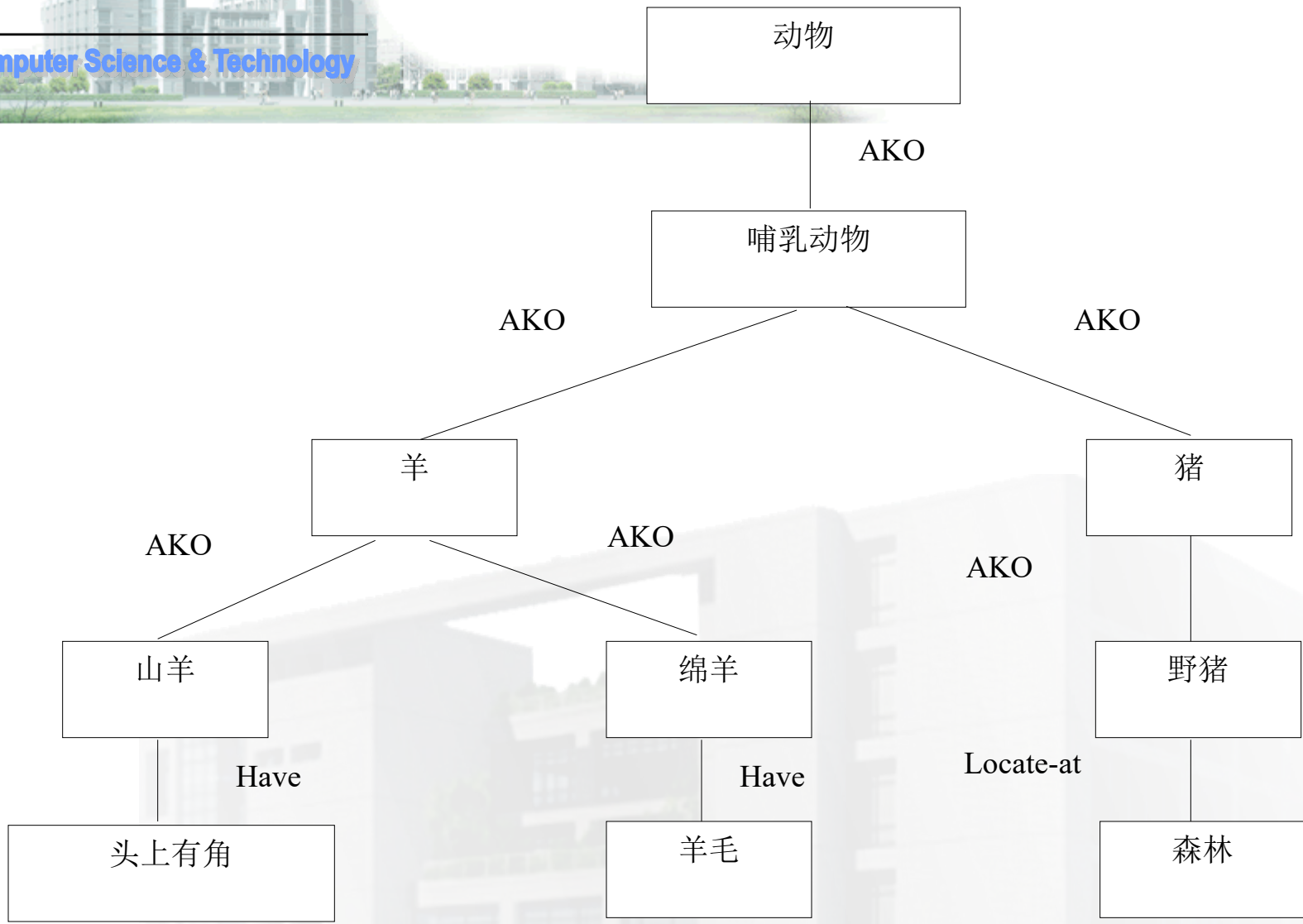


图2.21 有关猪和羊的语义网络

- 注意：对于一句复杂的话画语义网络，应该先抓主干，再考虑修饰。如果主干涉及到三元关系，则采用增加节点的方法化成多个二元关系。 ★

例 用语义网络表示下列知识：教师张明在本年度第二学期给计算机应用专业的学生讲授“人工智能”这一门课程。

解题分析：本质是涉及的对象包括：**教师、张明、学生、计算机应用、人工智能、本年度第二学期**等。然后确定各对象间的关系。“张明”与“教师”之间是一种类属关系，可用ISA表示；“学生”和“计算机应用”间的关系是一种属性关系可以用Major表示。“张明”、“学生”和“人工智能”则是通过“讲课”这一动作联系在一起。从上面的分析可知，必须**增加一个动作节点“讲课”**，“张明”是这一动作的主体，而“学生”和“人工智能”是这一动作的两个客体。“本年度第二学期”则是这个动作的作用时间，属于一种时间关系。因此，通过增加这个动作节点“讲课”将网络中的各节点联系起来。由“讲课”节点引出的弧不仅指出了讲课的主体和客体，还指出了讲课的时间。

通过分析可得其对应的语义网络如图2.22所示。

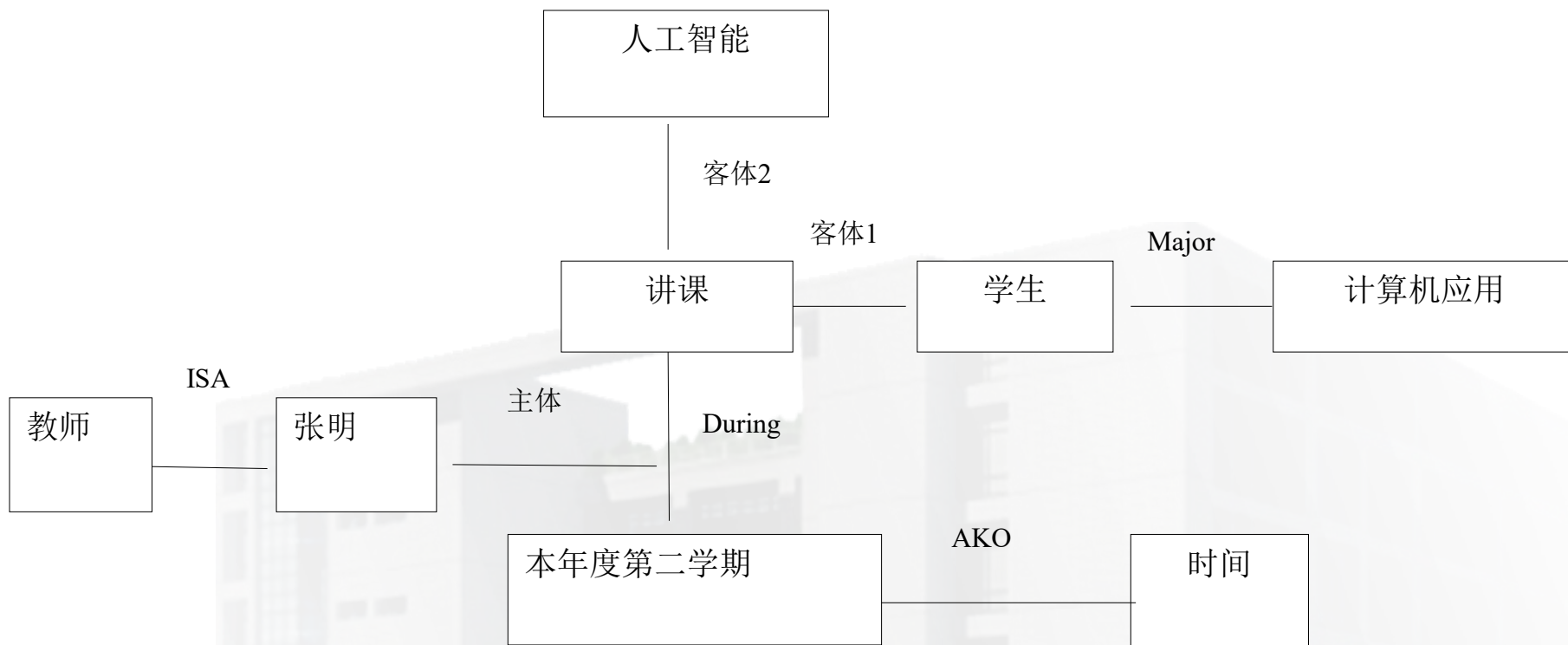


图2.22 有关讲课的语义网络

用语义网络表示知识的问题求解系统主要有两大部分组成，一部分是由语义网络构成的知识库，另一部分是用于问题求解的推理机。语义网络的推理过程主要有两种，一种是继承，另一种是匹配。

1. 继承推理

继承是指把对事物的描述从抽象结点传递到具体结点。通过继承可以得到所需结点的一些属性值，它通常是沿着ISA、AKO、AMO等继承弧进行的。继承的一般过程为：

- (1) 建立结点表，存放待求结点和所有以ISA、AKO、AMO等继承弧与此结点相连的那些结点。初始情况下，只有待求解的结点
- (2) 检查表中的第一个是否有继承弧。如果有，就从该弧所指的所有结点放入结点表的末尾，记录这些结点的所有属性，并从结点表中删除第一个结点。如果没有，仅从结点表中删除第一个结点。
- (3) 重复检查表中的第一个是否有继承弧，直到结点表为空。记录下来的属性就是待求结点的所有属性。

2. 匹配推理

语义网络问题的求解一般是通过匹配来实现的。所谓匹配就是在知识库的语义网络中寻找与待求问题相符的语义网络模式。其主要过程为：

- (1) 根据问题的要求构造网络片断，该网络片断中有些结点或弧为空，标记待求解的问题（询问处）。
- (2) 根据该语义网络片断在知识库中寻找相应的信息。
- (3) 当待求解的语义网络片断和知识库中的语义网络片断相匹配时，则与询问处（也就是待求解的地方）相匹配的事实就是问题的解。

注意：

在语义网络知识表达方法中，**没有形式语义**，也就是说，和谓词逻辑不同，对所给定的表达表示什么语义**没有统一**的表示法。

赋予网络结构的含义完全决定于管理这个网络的过程的特性。

在已经设计出来的以语义网络为基础的系统，它们各自采用不同的推理过程。但推理的核心思想无非是继承和匹配。

2.4.6 语义网络表示法的特点

(1) 结构性：语义网络把事物的属性以及事物间的各种语义联系显式地表现出来，是一种结构化的知识表示法。在这种方法中，下层结点可以继承、新增和修改上层结点的属性，从而实现信息共享。

(2) 联想性：着重强调事物间的语义联系，体现了人类思维的联想过程。

(3) 自索引性：语义网络表示把各结点之间的联系以明确、简洁的方式表示出来，通过与某一结点连接的弧很容易找出相关信息，而不必查找整个知识库。可以有效地避免搜索时的组合爆炸问题。

(4) 自然性：是一种直观的知识表示方法，符合人们表达事物间关系的习惯，而且把自然语言转换成语义网络也较为容易。

(5) 非严格性：语义网络没有公认的形势表示体系，它没有给其节点和弧赋予确切的含义，有时标签语义模糊。在推理过程中有时不能区分物体的“类”和“实例”的特点。因此，通过语义网络实现的推理不能保证其推理结果的正确性。

另外，语义网络表示法的推理规则不十分明了。其表达范围也受到一定限制，一旦当语义网络中结点个数比较多时，网络结构复杂，推理就难以进行。

□ 表示包含下面句子含义的**语义网络**：

- (1) 典型的**哺乳动物**有**毛发**。

- **Have**(哺乳动物,毛发)

- (2) **狗**是哺乳动物，且吃**肉**。

- **Ako**(狗,哺乳动物) \wedge **Eat**(狗,肉)

- (3) **Fido**是**John**的狗。

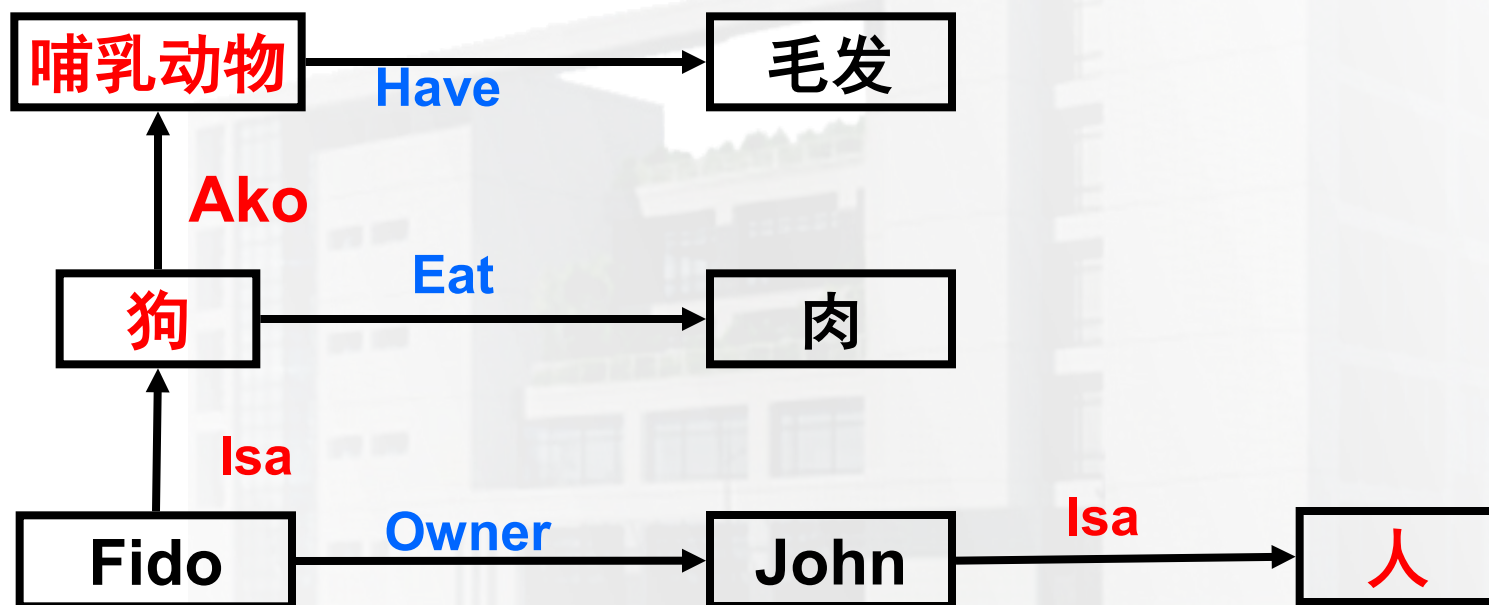
- **Own**(John,Fido)

正常使用主观题需2.0以上版本雨课堂

作答

□ 表示包含下面句子含义的**语义网络**：

- (1) 典型的**哺乳动物**有**毛发**。
 - **Have**(哺乳动物,毛发)
- (2) **狗**是哺乳动物, 且**吃肉**。
 - **Ako**(狗,哺乳动物) \wedge **Eat**(狗,肉)
- (3) **Fido**是**John**的**狗**。
 - **Isa**(Fido,狗) \wedge **Owner**(Fido,John)



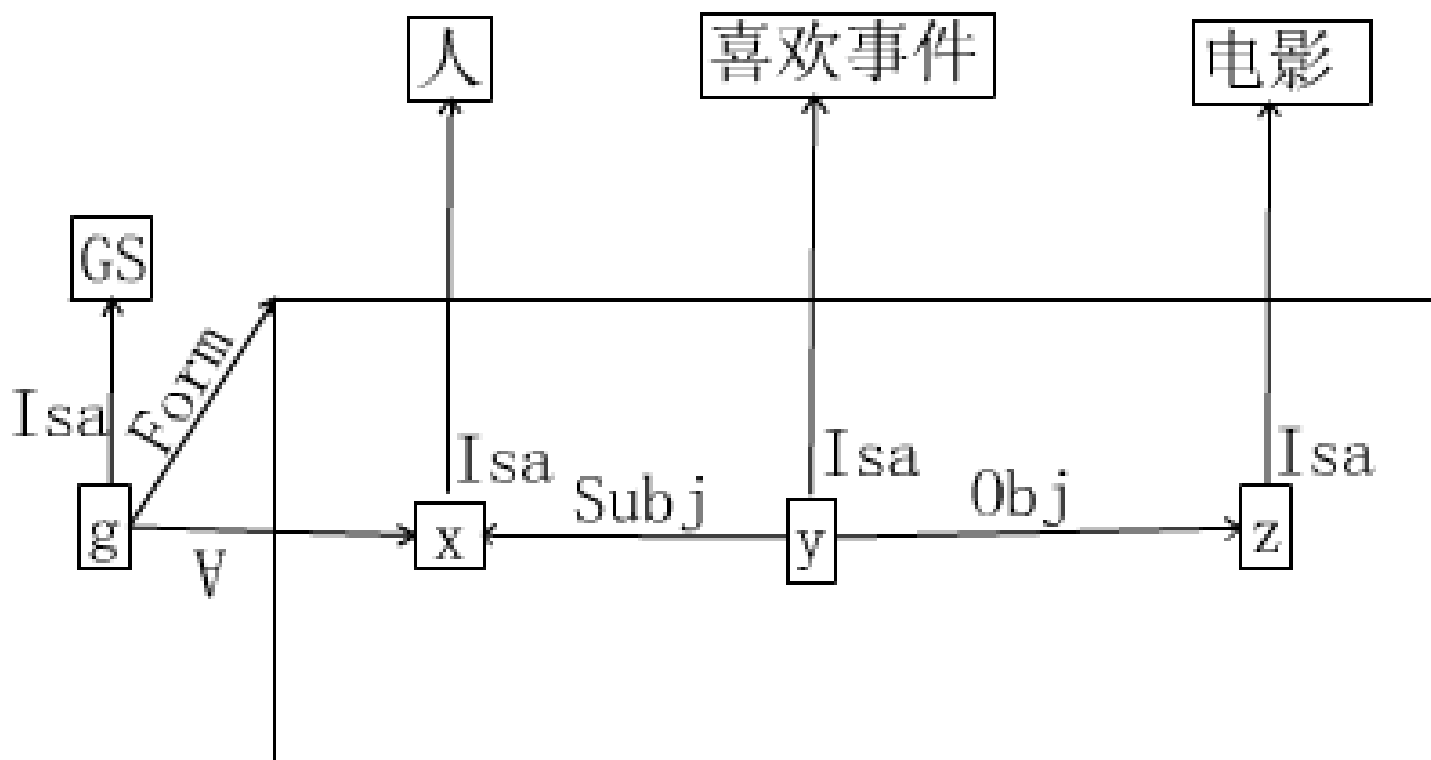
□ 用语义网络表示
每个人都喜欢电影。

正常使用主观题需2.0以上版本雨课堂

作答

□ 用语义网络表示

每个人都喜欢电影。



2.5 框架表示法

□ 【观察事物时的思维】

- 人头脑中形成的对于事物的**抽象模型（框架）**；
- 用“**自顶向下**”的方法
 - 先**匹配**一个现有的**抽象模型（框架）**；
 - 再确定**抽象模型（框架）**中的细节。

□ 框架表示法

- 明斯基，1975，强调**事物内部**的结构化描述；
- 较好地反映人**观察事物**的思维方式；
- 应用于**机器人识别领域**。

□ 1、框架的一般表示结构

- 框架由描述事物各个方面属性的槽 (slot) 组成

<框架> := (Frame <框架名> {<槽>}+)

- 槽有多侧面 (aspect)

<槽> := (<槽名> {<侧面>} +)

<侧面> := <侧面名> : <侧面值>

- 与语义网络节点的表示结构相比较:

- 表示结构接近, 但框架更丰富 (引入侧面)

- 语义网络节点: <槽> := {<槽名> : <槽内容>}

- 框架更强调表示事物的内部结构;

- 语义网络节点更强调表示事物间的关系;

<框架名>

槽名 1	侧面名 1	值1, 值2, ... 值 p_1
	侧面名 2	值1, 值2, ... 值 p_2
	
槽名 2	侧面名 m_1	值1, 值2, ... 值 p_{m_1}
	侧面名 1	值1, 值2, ... 值 q_1
	侧面名 2	值1, 值2, ... 值 q_2
	
槽名 n	侧面名 m_2	值1, 值2, ... 值 q_{m_2}
	
	侧面名 1	值1, 值2, ... 值 r_1
	侧面名 2	值1, 值2, ... 值 r_2
约束:	
	侧面名 m_n	值1, 值2, ... 值 r_{m_n}
	约束条件 1	
	约束条件 2	
	
	约束条件 n	

例如，要用框架来描述“优质商品”这一概念。首先分析商品所具有的属性，一个商品可能具有的属性有：商品名称、生产厂商、生产日期、获奖情况等等，这里只考虑这几个属性。这几个属性可以定义为“优质商品”框架的槽，而“获奖情况”这个属性还可以从获奖等级、颁奖部门和获奖时间这3个侧面来加以描述。如果给各个槽和侧面赋予具体的值，就得到了“优质商品”这一概念的一个实例框架。

框架名： <优质商品>

商品名称： 红桃K口服液

生产厂商： 红桃K集团

生产日期： 1998年6月17日

获奖情况： **获奖等级：** 省级

颁奖单位： 湖北省卫生厅

获奖时间： 2000年5月

通常在框架系统中定义一些公用、常用且标准的槽名，并把这些槽名称为系统预定义槽名。人们在使用这些槽名时，不用说明就知道它表示何种联系。下面给出几个比较常用的、用来表示对象间关系的槽名：

(1) ISA槽

ISA槽用于指出对象间抽象概念上的类属关系。其直观意义是“是一个”，“是一种”，“是一只”……。在一般情况下，用ISA槽指出的联系都具有继承性。

所谓框架的继承性就是指当下层框架中的某些槽值或侧面值没有被直接给定时，可以从其上层框架中继承这些值或属性。

例如，椅子一般有4条腿，如果一把具体的椅子没有指出它有几条腿时，则可以通过一般椅子的特性，得出它有4条腿。

(2) AKO槽

AKO槽用于具体地指出对象间的类属关系。其直观意义是“是一种”。当用它作为某下层框架的槽时，就明确地指出了该下层框架所描述的事物是其上层框架所描述事物中的一种，下层框架可继承上层框架中值或属性。

(3) Instance槽

Instance槽用来表示AKO槽的逆关系。当用它作为某上层框架的槽时，可在该槽中指出它所联系的下层框架。用Instance槽指出的联系都具有继承性，即下层框架可继承上层框架中所描述的属性或值。

(4) Part-of槽

Part-of槽用于指出部分和全体的关系。当用其作为某框架的一个槽时，槽中所填的值称为该框架的上层框架名，该框架所描述的对象只是其上层框架所描述对象的一部分。

例如，“两条腿”是“人体”的一部分。可以将“两条腿”和“人体”分别定义成框架，“两条腿”为下层框架，“人体”为其上层框架。在“两条腿”的框架中设置一个**Part-of**槽，槽值填入<人体>这个框架名。

显然，用**Part-of**槽指出的联系所描述的下层框架和上层框架之间**不具有继承性**。

1. 事实性知识的表示

对于一些简单的事实，例如“鸟有翅膀”，“轮胎是汽车的一部分”，这里要描述这些事实需要两个节点，用前面给出的基本语义联系或自定义的基本语义联系就可以表示了。对于稍微复杂一点的事实，比如在一个事实中涉及到多个事物时，如果语义网络只被用来表示一个特定的事物或概念，那么当有更多的实例时，就需要更多的语义网络，这样就使问题复杂化了。

通常把有关一个事物或一组相关事物的知识用一个语义网络来表示。

2.情况、动作和事件的表示

为了描述那些复杂的知识，在语义网络的知识表示法中，通常采用引进**附加节点**的方法来解决。西蒙（Simon）在提出的表示方法中增加了**情况节点、动作节点和事件节点**，允许用一个节点来表示情况、动作和事件。

例如，对事实“每个学生都学习了一门外语”，可用图2.19所示的语义网络表示。

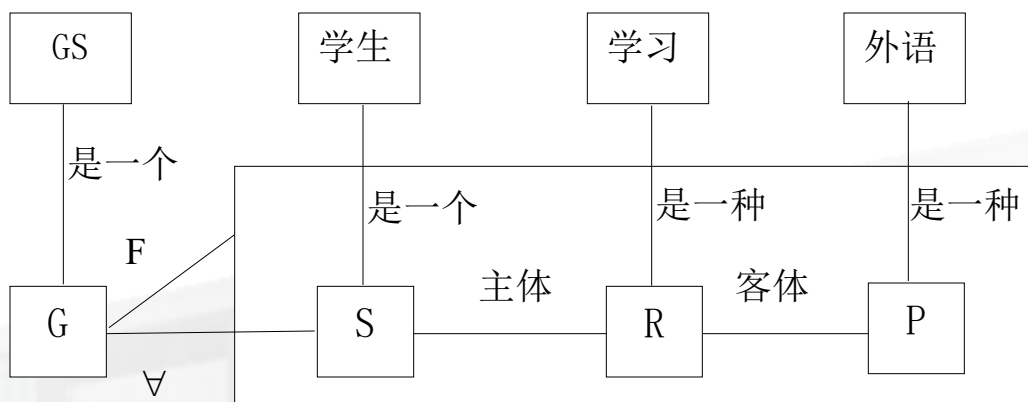


图2.19 具有全称变量的语义网络（分块语义网络）

其中G代表整个陈述句，它是一般陈述句GS的一个实例。G中的每一个元素至少有两个特性：Form(F)，即句中的关系和全称量词(\forall)。在这个例子中只有一个变量S具有全称量词，Form中其余两个变量R, P看成具有存在量词。

□ 1、框架的一般表示结构

- 框架由描述事物各个方面属性的槽 (slot) 组成

<框架> := (Frame <框架名> {<槽>}+)

- 槽有多侧面 (aspect)

<槽> := (<槽名> {<侧面>} +)

<侧面> := <侧面名> : <侧面值>

- 与语义网络节点的表示结构相比较：

- 表示结构接近，但框架更丰富（引入侧面）

- 语义网络节点：<槽> := {<槽名> : <槽内容>}

- 框架更强调表示事物的内部结构；

- 语义网络节点更强调表示事物间的关系；

例 下面一是一个描述“教师”的框架。

框架名：〈教师〉

类属：〈知识分子〉

工作：范围：（教学，科研）

缺省：教学

性别：（男，女）

学历：（中专，大学）

类别：（〈小学教师〉，〈中学教师〉，〈大学教师〉）

在这个框架中，框架名为“教师”，它含有5个槽，槽名分别是“类属”、“工作”、“性别”、“学历”和“类别”。这些槽名后面就是其槽值，而槽值“〈知识分子〉”又是一个框架名。“范围”、“缺省”是槽“工作”的两个不同的侧面，其后是侧面值。

例 下面是描述“大学教师”的框架。

框架名：〈大学教师〉

类属：〈教师〉

学位：范围：（学士，硕士，博士）

缺省：硕士

专业：〈学科专业〉

职称：范围：（助教，讲师，副教授，教授）

缺省：讲师

水平：范围：（优，良，中，差）

缺省：良

从上述两例可以看出，这两个框架之间存在一种层次关系，称前者为上层框架（或父框架），后者为下层框架（或子框架）。

例 下面是描述一个具体教师的框架。

框架名：〈教师-1〉

类属：〈大学教师〉

姓名：张宇

性别：男

年龄：32

职业：〈教师〉

职称：副教授

部门：计算机系

研究方向：计算机软件与理论

工作：参加时间：2000年7月

工龄：当前年份-2000

工资：〈工资单〉

比较上面几个例子，可以发现“教师-1”是“大学教师”的下层框架，而“大学教师”又是“教师”的下层框架，“教师”又是“知识分子”的下层框架。框架之间的这种层次关系是相对而言的，下层框架可以从上层框架继承某些属性或值。这样，一些相同的信息可以不必重复存储，节省了存储空间，这种层次结构对减少冗余信息有重要意义。

例 下面有关地震的新闻报道，请用框架结构表示这段报道。

“今天，一次强度为里氏8.5级的强烈地震袭击了下斯洛文尼亚（Low Slabovia）地区，造成25人死亡和5亿美元的财产损失。下斯洛文尼亚地区主席说：多年来，靠近萨迪壕金斯断层的重灾区一直是一个危险地区。这是本地区发生的第3号地震。”

解：首先分析关于地震报道中所涉及的一些有关地震的关键属性，这些属性是地震发生的地点、时间、伤亡人数、财产损失数量、地震强度的震级和断层情况。这些属性可以作为该框架的各个槽。

接下来，将本报道中的有关数据填入相应的槽后就得到了给第3号地震的框架。

框架名：〈第3号地震〉

地点：Low Slabovia

时间：今天

伤亡人数：25

财产损失：5亿美元

震级：8.5级

断层：萨迪壕金斯

该框架也可以用下图表示。

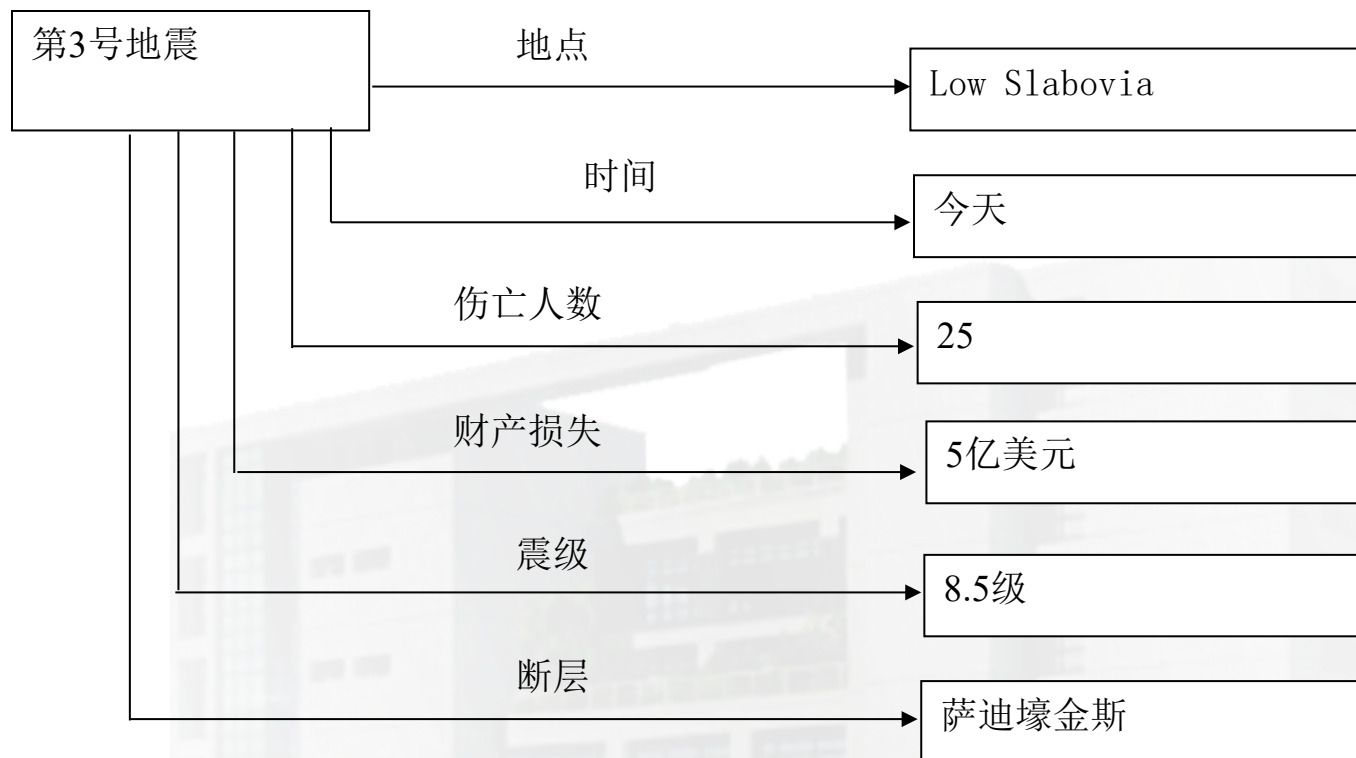


图2.2 斯洛文尼亚第3号地震框架图

在框架表示的知识库中，主要有两种活动：一是填槽，即框架中未知内容的槽需要填写；而是“匹配”，根据已知事件寻找合适的框架，并将该内容填入槽中。

上述两种操作均将引起推理，其主要推理形式有：

1、默认推理

在框架网络中，各框架之间通过ISA链（槽）构成半序的继承关系。在填槽过程中，如果没有特别的说明，子框架的槽值将继承父框架相应的槽值，称为默认推理。

2、匹配

由框架所构成的知识库，当利用它进行推理、形成概念和作出决策、判断时，其过程往往是根据已知的信息，通过与知识库中预先存储的框架进行匹配，找出一个或几个与该信息所提供的情况最适合的预选框架，形成初步假设，即由输入信息激活相应的框架。然后在在该假设框架引导下，收集进一步的信息。按某种评价原则，对预选的框架进行评价，以决定最后接受或放弃预选的框架，即在框架引导下的推理。这个过程可以用来模拟人类利用已有的经验进行思考、决策，以及形成概念、假设的过程。

2.5.4 框架表示法的特点

(1) 继承性

是框架的一个很重要的性质，下层框架可以从上层框架继承某些属性或值，也可以进行补充和修改。这样一些相同的信息可以不必重复存储，减少冗余信息节省了存储空间。

(2) 结构化

框架表示法是一种结构化的知识表示方法。不但把知识的内部结构表示出来还可以把知识之间的联系也表示出来，是一种表达能很强的知识表示方法。

2.5.4 框架表示法的特点

(3) 自然性 在人类思维和理解活动中分析和解释遇到的情况时，就从记忆中选择一个类似事物的框架，通过对其细节进行修改或补充，形成对新事物的认识，这与人们的认识活动是一致的。

(4) 推理灵活多变 框架表示法没有固定的推理机制，它可以根据待求解问题的特点采取灵活地采取多种推理方法。

(5) 框架表示法的主要不足之处 在于它不善于表达过程性知识。因此它经常与产生式表示法结合起来使用，以取得互补效果。

目录



CONTENTS

01

知识图谱概览

02

典型的知识库项目介绍

03

知识图谱构建

目录



01

知识图谱概览

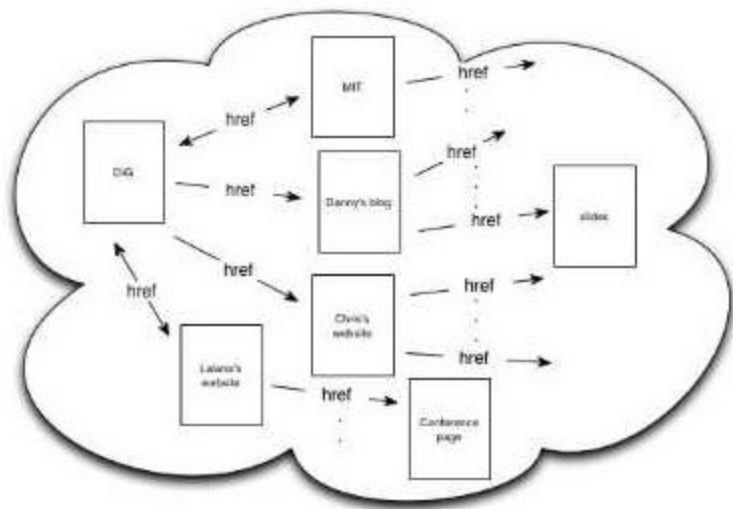
CONTENTS



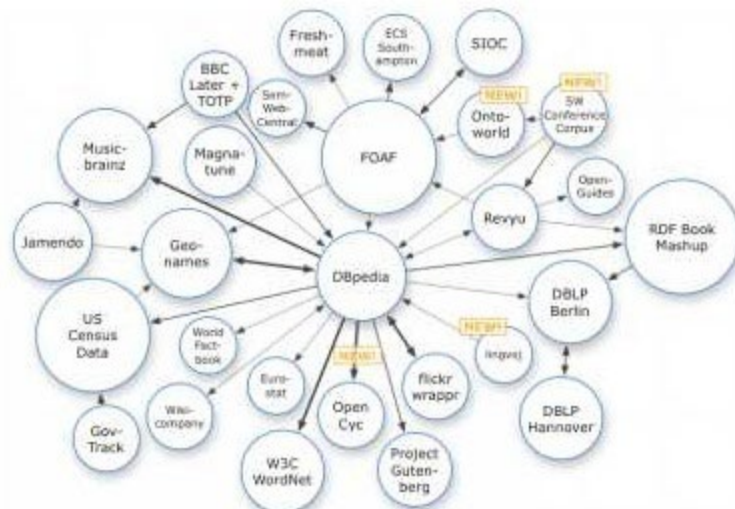
产生背景

大量数据多源异质，组织结构松散，不利于接受和使用。

➡ 知识互联，构建人与机器都可理解的万维网。



Web of Document
人类可理解

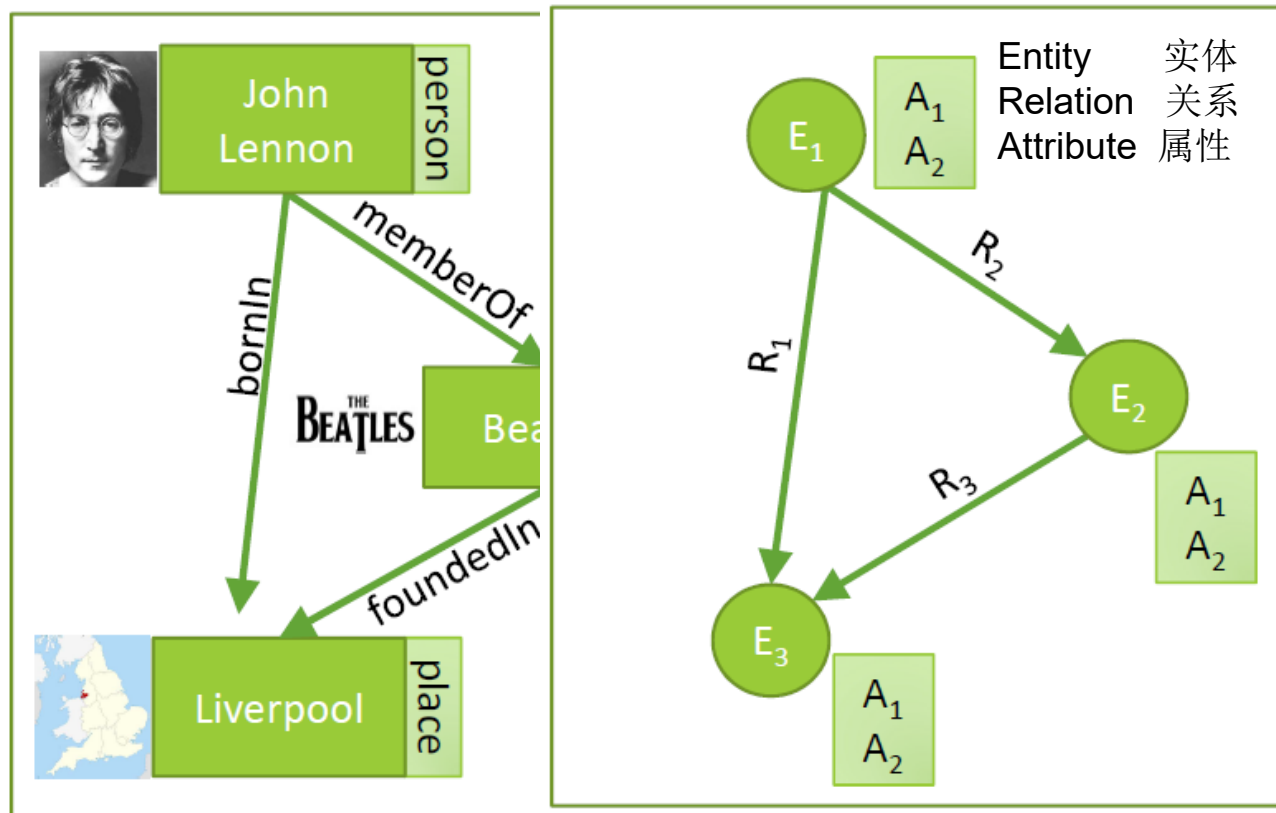


Web of Data
机器可理解

Things, not strings!

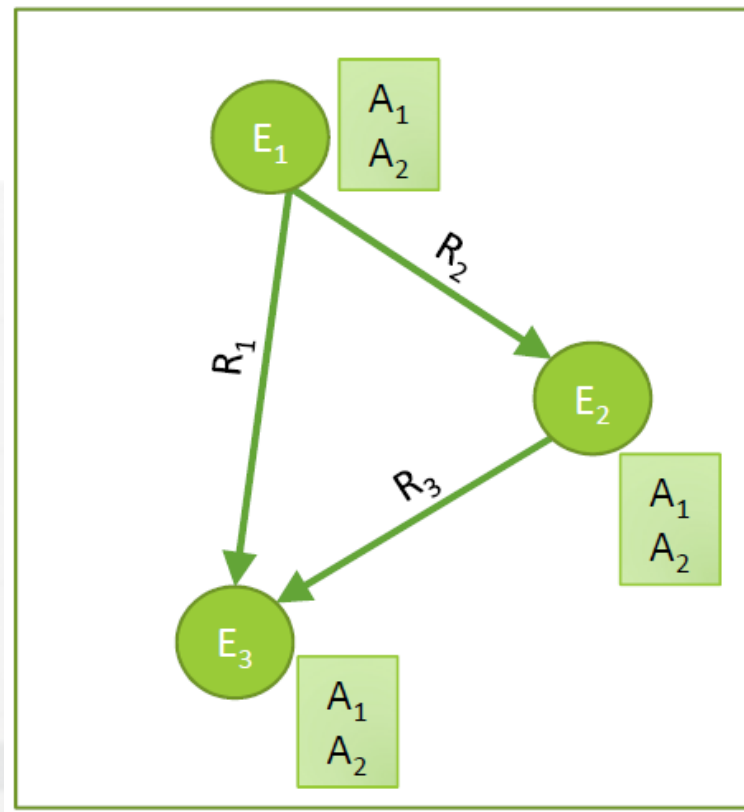
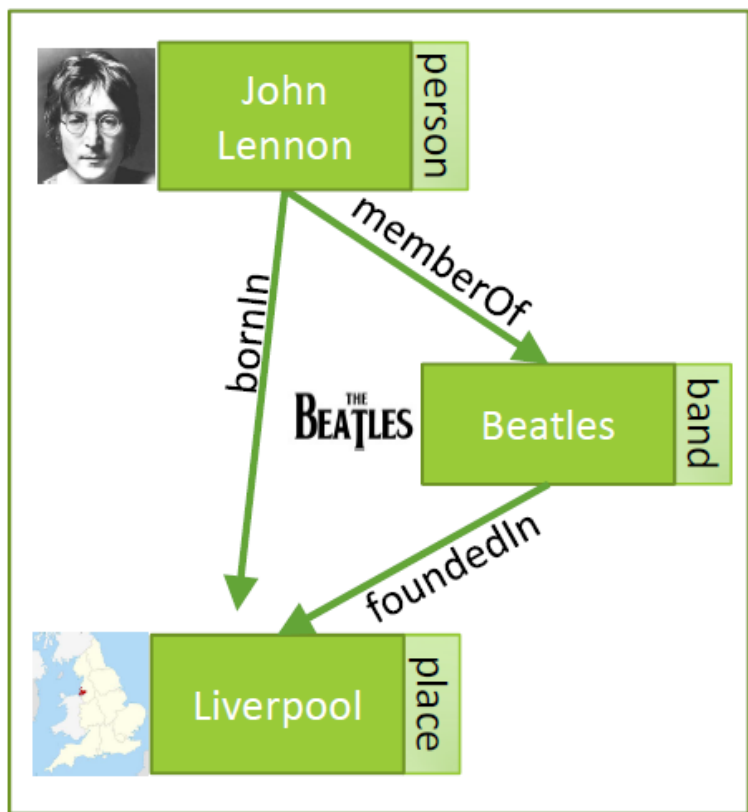
什么是知识图谱

结构化的语义知识库，以符号形式描述物理世界中的概念及其相互关系。即，图谱形式的知识信息。

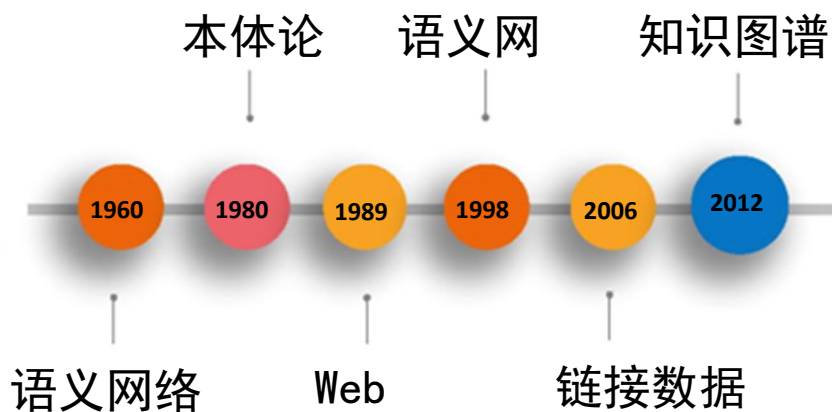
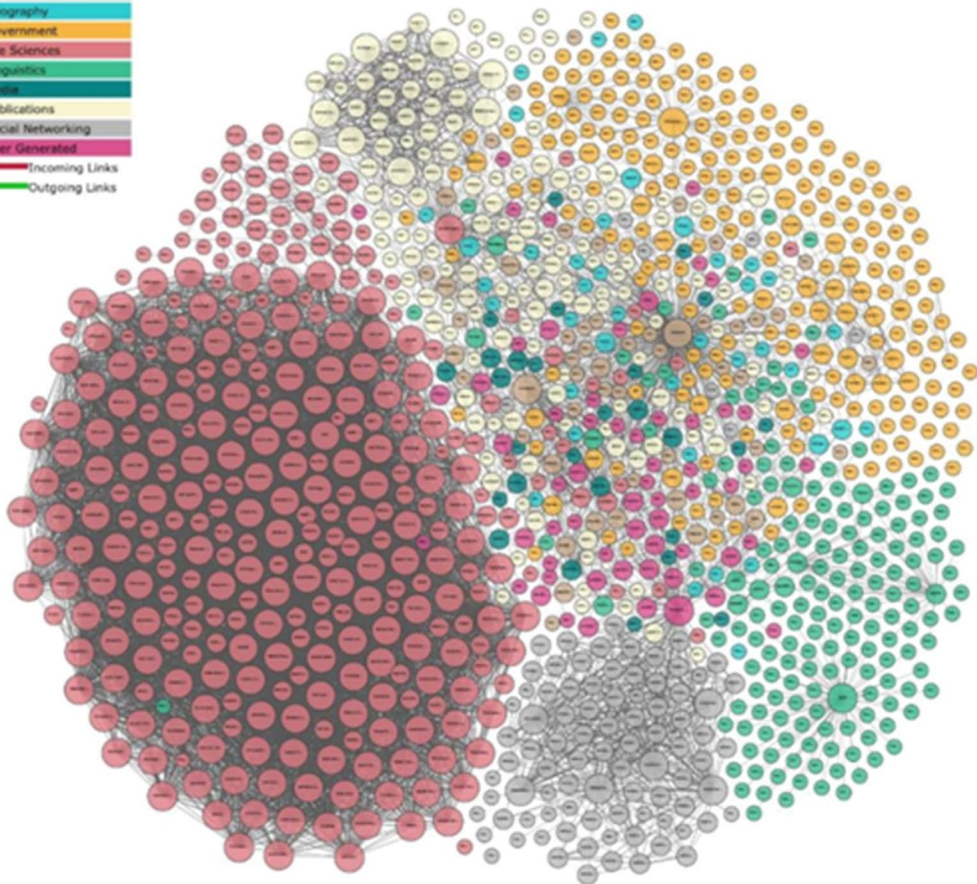


什么是知识图谱

本质：以图形的方式向用户返回经过加工和推理的知识



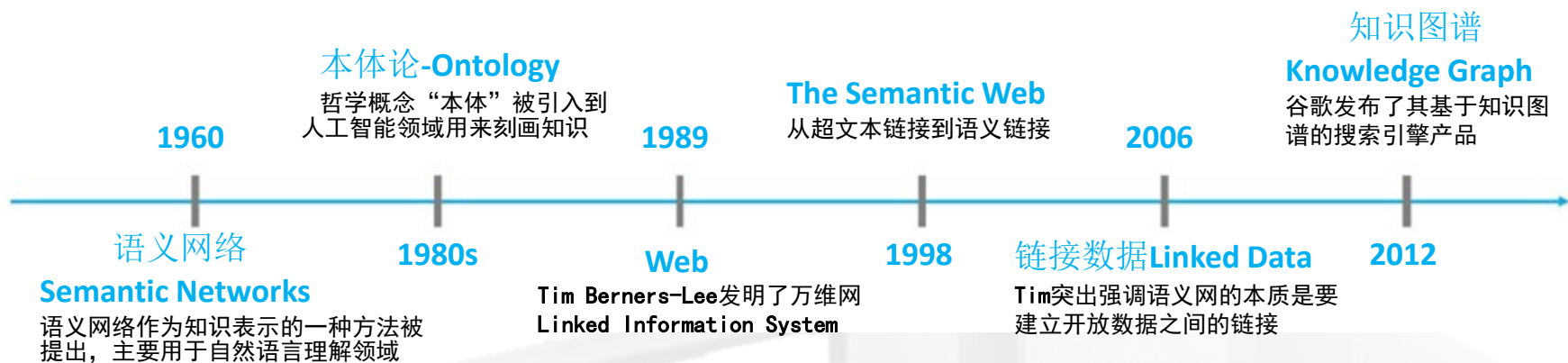
知识图谱 (Knowledge Graph, KG)



"Linking Open Data cloud diagram 2017, by
Andrejs Abele, John P. McCrae, Paul
Buitelaar, Anja Jentzsch and Richard
Cyganiak. <http://lod-cloud.net/>"

KG的概念演化

Computer Science & Technology



知识表示与知识库- Knowledge Representation / Knowledge Base

人工智能研究者陆续提出了大量知识表示的方法，如框架系统、产生式规则、描述逻辑等。

知识图谱得益于Web的发展（更多的是数据层面），有着来源于KR、NLP、Web、AI多个方面的基因。

Web的理想是万物的链接，搜索的理想是事物的搜索

Google obama

全部 图片 新闻 视频 地图 更多 设置 工具

找到约 104,000,000 条结果 (用时 0.49 秒)

Barack Obama - Wikipedia
https://en.wikipedia.org/wiki/Barack_Obama
 Barack Hussein **Obama II** is an American politician who served as the 44th President of the United States from 2009 to 2017. He is the first African American to ...
[Michelle Obama](#) · [Ann Dunham](#) · [Family of Barack Obama](#) · [Barack Obama Sr.](#)

The Office of Barack and Michelle Obama
<https://www.barackobama.com/> · [翻译此页](#)
 Welcome to the Office of Barack and Michelle **Obama**. We Love You Back. Play video. The Office of Barack and Michelle **Obama**. © 2017 | [Legal & Privacy](#).

Barack Obama (@BarackObama) | Twitter
<https://twitter.com/barackobama> · [翻译此页](#)
 15.5K tweets · 2068 photos/videos · 94.5M followers. "To target hopeful young strivers who grew up here is wrong, because they've done nothing wrong."

Barack Obama - Home | Facebook
<https://www.facebook.com/barackobama/> · [翻译此页](#)
 Barack **Obama**, Washington, DC. 55M likes. Dad, husband, former President, citizen.

President Obama's Farewell Address: Full Video and Text - The New ...
<https://www.nytimes.com/2017/01/10/.../obama-farewell-address-speech.htm> · [翻译此页](#)
 2017年1月10日 - President **Obama** delivered his farewell address in Chicago on Tuesday. The following is the complete transcript, as provided by the Federal ...

Barack Obama response reflects ex-president balancing act ...
www.cnn.com/2017/08/17/politics/barack-obama...response/index.html · [翻译此页](#)
 2017年8月17日 - Since leaving office in January, President Barack **Obama** has used written statements to defend the Affordable Care Act, denounce a decision ...

Barack Obama: Latest News, Top Stories & Analysis - POLITICO
www.politico.com/news/barack-obama · [翻译此页](#)
 Latest news, headlines, analysis, photos and videos on Barack **Obama**.

Barack Obama Articles, Photos, and Videos - Los Angeles Times
www.latimes.com > [Topics](#) > [Politics and Government](#) > [Government](#) · [翻译此页](#)

原来的结果：
相关的网页

巴拉克·奥巴马
前美国总统

贝拉克·奥巴马，全名贝拉克·侯赛因·奥巴马二世，是美国政治人物，从2009年至2017年任第44任美国总统。他是第一位美国非裔总统，也是第一位出生于美国本土之外的美国总统。在这之前，他从2005年至2008年代表伊利诺伊州担任美国联邦参议员，从1997年至2004年担任伊利诺伊州参议员。 [维基百科](#)

生于：1961年8月4日（56岁），夏威夷州檀香山Kapiolani Medical Center for Women and Children

身高：6' 1"

子女：玛莉亚·安·奥巴马，萨莎·奥巴马

家长：安·邓纳姆，老贝拉克·奥巴马

教育背景：哈佛法学院（1988年–1991年），更多 >

图书 还有30+项

奥巴马的梦想之路-以父之名 1995年	无畏的希望 2006年	奥巴马为你歌唱：给女儿的信 2010年	Change We Can Believe In 2008年

用户还搜索了 还有15+项

乔·拜登 最新热搜榜	唐纳德·特朗普	伊万卡·特朗普 最新热搜榜	希拉里·克林顿	米歇尔·奥巴马

现在的结果：
精准的知识



三百六十行 行行有百科

请输入搜索内容

搜索

百科分类

生活

美食 宠物 家庭 购物 服饰 家居 情感
出行 美容 日用品 手工 鲜花礼品

体育

球类运动 休闲运动 比赛 俱乐部 彩票
武术 养生运动 体育器材 智力体育

娱乐

电影 电视剧 音乐 舞蹈 综艺节目 演出
明星

财经

股票 房产 投资 理财 汽车 基金 贷款
外汇 财税 保险 经营 银行 债券

农业种植

蔬菜 水果 园林花卉 农药肥料 水产
畜禽 饲料兽药

旅游

世界名城 中国城市 中国古迹 中国旅游景点
名山名河 户外 自驾游 自驾游

健康

疾病 两性 整形 心理 保健 饮食健康
母婴 养生 药品 中老年 中医 美容
医疗器械

社会

政治 军事 法律 社会问题 战争
灾难事故 民生 时事评论 公益慈善
社会现象 服务 企业 福利保障

科技

通信 IT 技术 数码 家电 手机 互联网
软件 硬件 能源 电子商务 科学探索

商务服务

翻译 物流 中介 酒店餐饮 印刷包装
广告公关 展览 维修

教育

高校 幼儿教育 中小学教育 技能培训
校园 网络教育 资格认证 高等教育 考试
职场学习 留学

文化

历史 文学 宗教 民族 民俗 戏曲 工艺
语言 文物古迹 国学 文化遗产 绘画
书法 雕塑 设计 视觉 建筑 综艺
艺术鉴定 拍卖收藏 科普

游戏/动漫

电子游戏 电脑游戏 手机游戏
网络游戏 桌面游戏 卡牌游戏 网页游戏
运动游戏 语言游戏 益智游戏 漫画 动画

工业技术

原材料 电子电工 通信技术 建筑材料
五金 机械 石油化工 仪器仪表 安全防护

电影



Seediq Bale百科

《赛德克·巴莱太阳旗》和《赛德克·巴莱彩虹桥》两部分。魏德圣导演重新剪辑的版本。影片讲述的是1930年，台

赛德克·巴莱业内人士好评为何如潮？
骄傲才是生命，我们就要民族骄傲性！！
赛德克·巴莱为何又叫座又叫卖？



香港电影金像奖百科

香港电影金像奖标志香港电影金像奖于1982年由《电影双周刊》创办。这是从《电影双周刊》每年邀请影评人评选

是变态还是复仇？新闻与【微电影】引人深思
为什么说金像奖没有争议有些遗憾？
金像奖影后叶德嫻还有哪些影视作品？



新少林寺百科

《新少林寺》由香港动作片导演陈木胜执导，汇聚了刘德华、谢霆锋、成龙、范冰冰等两岸三地实力战将，影片总投资

你了解少林寺武功吗？
成龙为何在《新少林寺》中改变戏路？
谢霆锋在《新少林寺》中是大反派吗？



新少林寺

《新少林寺》是由中国大陆和香港英皇电影公司联合制片出品，在获得少林寺正式授权与大力支持下，由陈木胜执导

可以到少林寺体验禅修吗？
范冰冰是《新少林寺》的出品人吗？
新建的少林寺真的被剧组烧掉了吗？



韩国影视百科

韩剧影视资讯介绍，逐步完善中。

杜珂
《加油金先生》
当下韩剧也玩穿越 细数韩古装剧三大特点



黄金大劫案百科

《黄金大劫案》是由李健，王文，黄勃主演的喜剧动作片。讲述了一个出身贫寒的小混混无意中卷进了一个地下党策

《黄金大劫案》里东北方言会成亮点吗？
大劫案，不想当英雄的混混不是情圣吗？
说说《黄金大劫案》的观后感



匹夫百科

《匹夫》是由杨树雁执导，黄晓明、张译、张歆艺等主演的热血匪帮传奇电影，于2012年4月24日上映。匹夫》都

《匹夫》黄晓明为什么要大牌？
《匹夫》黄晓明为什么拍激情？
《匹夫》张歆艺腿青手肿怎么不叫苦？



电影百科

汇集包括十大经典电影在内的电影漫评和电影观后感，解析科幻电影悬疑电影的前台幕后，罗列电影明星生活趣闻和

3·15电影质检报告
《遗落战境》之丢失的世界！
霍比特人3---中土世界的告别



机器学习实践应用

人工智能触手可及 让数据起舞 用算法扩展业务边界 阿里机器学习专家力作 实战经验分享

李博 著

京东价: **¥57.80** [8.4折] [定价: ¥69.00] (降价通知)

累计评价
900+

促销信息: 以下促销可在购物车任选其一

满减 每满100元, 可减30元现金 详情 >>

加价购 满10元另加16.90元, 或满20元另加9.90元, 或满30元另加6.90元, 即可在购物车换购热销商品 详情 >>

优惠券: **满105减6** **满200减50** **满200减16** 更多 >>

增值业务: **礼品包装**

配送至: 北京海淀区三环到四环之间 **有货**, 支持 99元免基础运费 | 货到付款

服务: 由 **京东** 发货, 并提供售后服务。23:00前下单, 预计明天(09月08日)送达

人气单品



机器学习实战

¥48.70



Python机器学习算法

¥57.80



统计学习方法

¥27.20



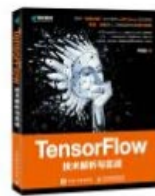
机器学习【首届京东文学奖-年度新锐入围作】

¥77.00



终极算法: 机器学习和人工智能如何重塑世界

¥44.20



TensorFlow技术解析与实战

¥66.20



大S前夫的前妻的现任丈夫是谁?



百度一下

网页

视频

资讯

图片

知道

贴吧

文库

地图

采购

更多

百度为您找到相关结果约6,730,000个

搜索工具



徐熙媛丈夫:

来自百度百科

具俊晔

具俊晔(DJ.koo), 1969年9月11日出生于韩国, 韩国歌手。与姜元来在大学时举行的一次夜总会迪斯科竞演大会上得了第一名, 1989年底加入了“WAWA”, 开始正式迈入歌坛。 [详情 >](#)

大s再婚对象是谁 前任丈夫是谁什么时候离婚的_手机搜狐网



从获悉的资料来看,大S离婚到再婚,中间隔了也就3个多月的时间,这再婚的速度也太快了。而且**现任丈夫**还是初恋,不得不说,这缘分实在是妙不可言。也有网友表示,看来初恋的魅力总是不可挡的,当对方回头...

搜狐网 百度快照

KG辅助问答

Computer Science & Technology

机器人及IoT设备的智能化：给万物都挂接一个背景知识库

对话式的信息获取更加需要`精准度和可靠度，知识图谱对于提升用户体验更加不可少。



DBpedia Yago

Wordnet



True Knowledge/Evi
WolframAlpha

DBpedia



zhishi.me



智能厨房



智能驾驶



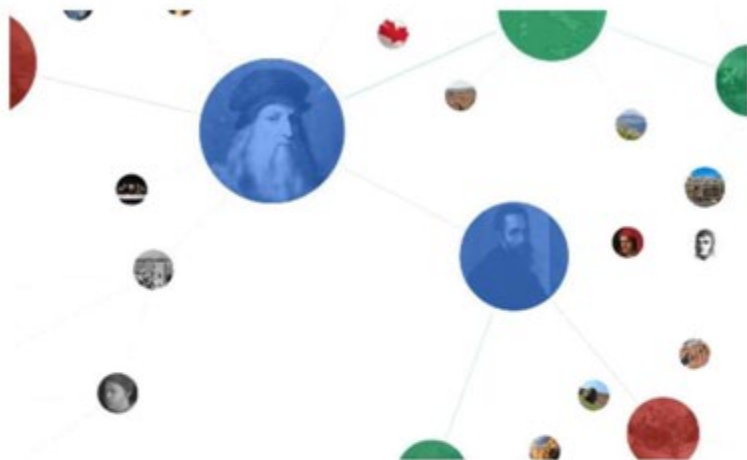
智能家居

DUER OS

Conversational AI OS for everyone and everywhere



MORE MACHINE UNDERSTANDABLE



建立数据链接
预先抽取语义

更加规范的数据表达

可计算的数据
粗糙数据



文本

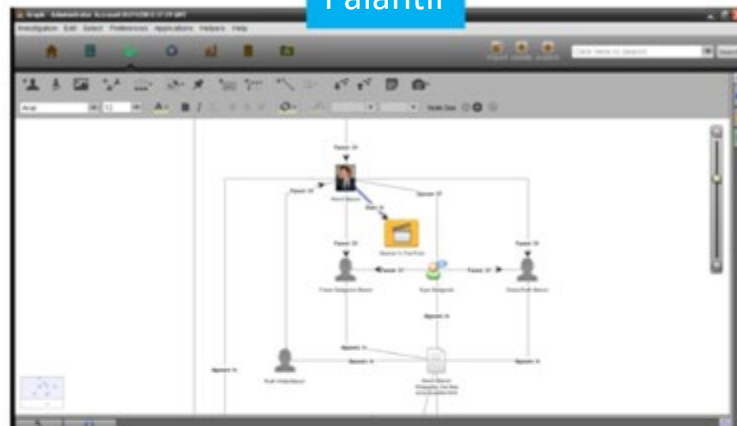


多媒体



传感器

Palantir



Kensho



Web视角 像建立文本之间的超链接一样，建立数据之间的语义链接，并支持语义搜索

NLP视角 怎样从文本中抽取语义和结构化数据

KR视角 怎样利用计算机符号来表示和处理知识

AI视角 怎样利用知识库来辅助理解人的语言

DB视角 用图的方式去存储知识，图数据库（ Neo4j、 TigerGraph、 Amazon Neptune、 JanusGraph、 ArangoDB ）

做好KG要兼容并蓄，综合利用好KR、NLP、Web、ML、DB等多方面的方法和技术

聪明的AI vs 有学识的AI

人的大脑依赖所学的知识进行思考、逻辑推理、理解语言



目录



CONTENTS

02

典型的知识库项目介绍

各种知识图谱项目

Computer Science & Technology

Freebase

LINKING OPEN DATA
W3C SWEO Community Project

WordNet
A lexical database for English

2HISH me

PKUBASE

NELL

schema.org

....the new SEO?

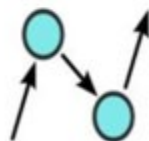
DBpedia

XLORE

WIKIDATA

WEB CHILD

CN-DBpedia



ConceptNet

An open, multilingual knowledge graph

Herbnet

YAGO
select knowledge

LinkedGeoData.org

WEBKB

linked life data

Cyc是在1984年由Douglas Lenat开始创建。最初的目标是要建立人类最大的常识知识库。典型的常识知识如“Every tree is a plant”，“Plants die eventually”等

Cyc知识库主要由术语Terms和断言Assertions组成。Terms包含概念、关系和实体的定义。Assertions用来建立Terms之间的关系，这既包括事实Fact描述，也包含规则Rule的描述

最新的Cyc知识库包含有50万条Terms和700万条Assertions。Cyc的主要特点是基于形式化的知识表示方法来刻画知识。形式化的优势是可以支持复杂的推理。但过于形式化也导致知识库的扩展性和应用的灵活性不够。Cyc提供开放版本OpenCyc。

<http://www.cyc.com/>

WordNet是最著名的词典知识库，主要用于词义消歧。WordNet由普林斯顿大学认识科学实验室从1985年开始开发

WordNet主要定义了名词、动词、形容词和副词之间的语义关系。例如名词之间的上下位关系（如：“猫科动物”是“猫”的上位词），动词之间的蕴含关系（如：“打鼾”蕴含着“睡眠”）等

WordNet3.0已经包含超过15万个词和20万个语义关系

WordNet: a lexical database for English

GA Miller - Communications of the ACM, 1995 - dl.acm.org

Abstract Because meaningful sentences are composed of meaningful words, any system that hopes to process natural languages as people do must have information about words and their meanings. This information is traditionally provided through dictionaries, and

被引用次数: 9323 相关文章 所有 34 个版本 引用 保存 更多

<http://wordnet.princeton.com/>

ConceptNet是常识知识库。最早源于MIT媒体实验室的Open Mind Common Sense (OMCS)项目。OMCS项目是由著名人工智能专家Marvin Minsky于1999年建议创立

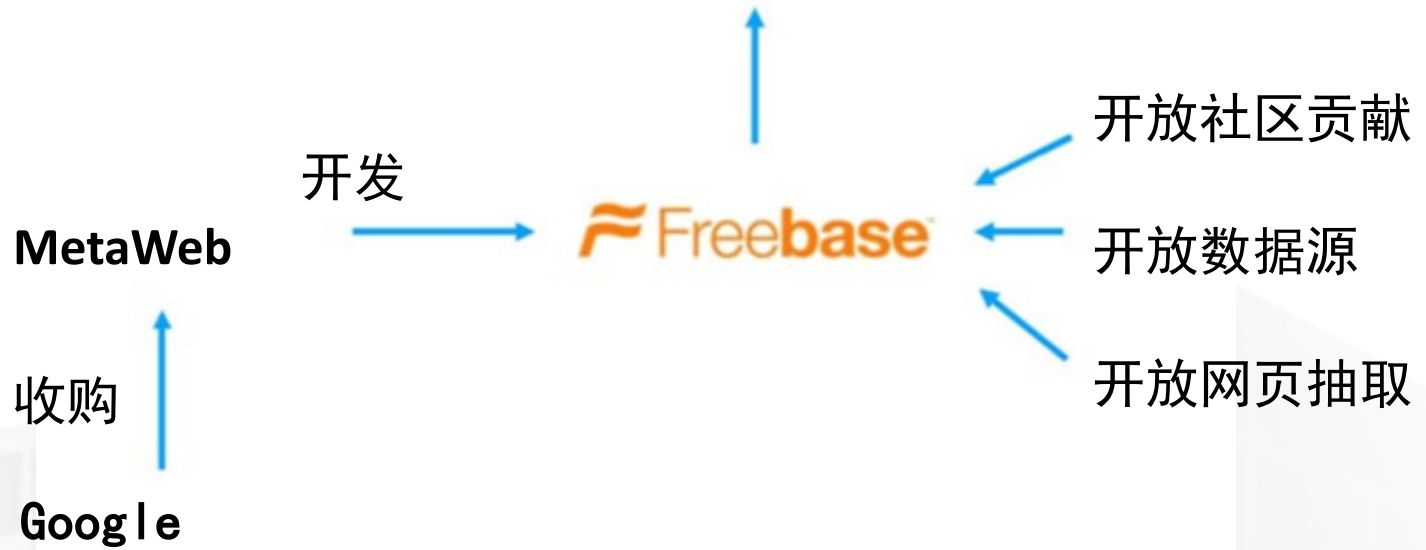
ConceptNet主要依靠互联网众包、专家创建和游戏三种方法来构建。新版本导入大量开放的结构化数据，如DBPedia、Wikinary, Wordnet等

ConceptNet知识库以三元组形式的关系型知识构成。ConceptNet5版本已经包含有2800万关系描述。与Cyc相比，ConceptNet采用了非形式化、更加接近自然语言的描述，而不是像Cyc那样采用形式化的谓词逻辑

与链接数据和谷歌知识图谱相比，ConceptNet比较侧重于词与词之间的关系。从这个角度看，ConceptNet更加接近于WordNet，但是又比WordNet包含的关系类型多。此外，ConceptNet完全免费开放，并支持多种语言

<http://www.conceptnet.io/>

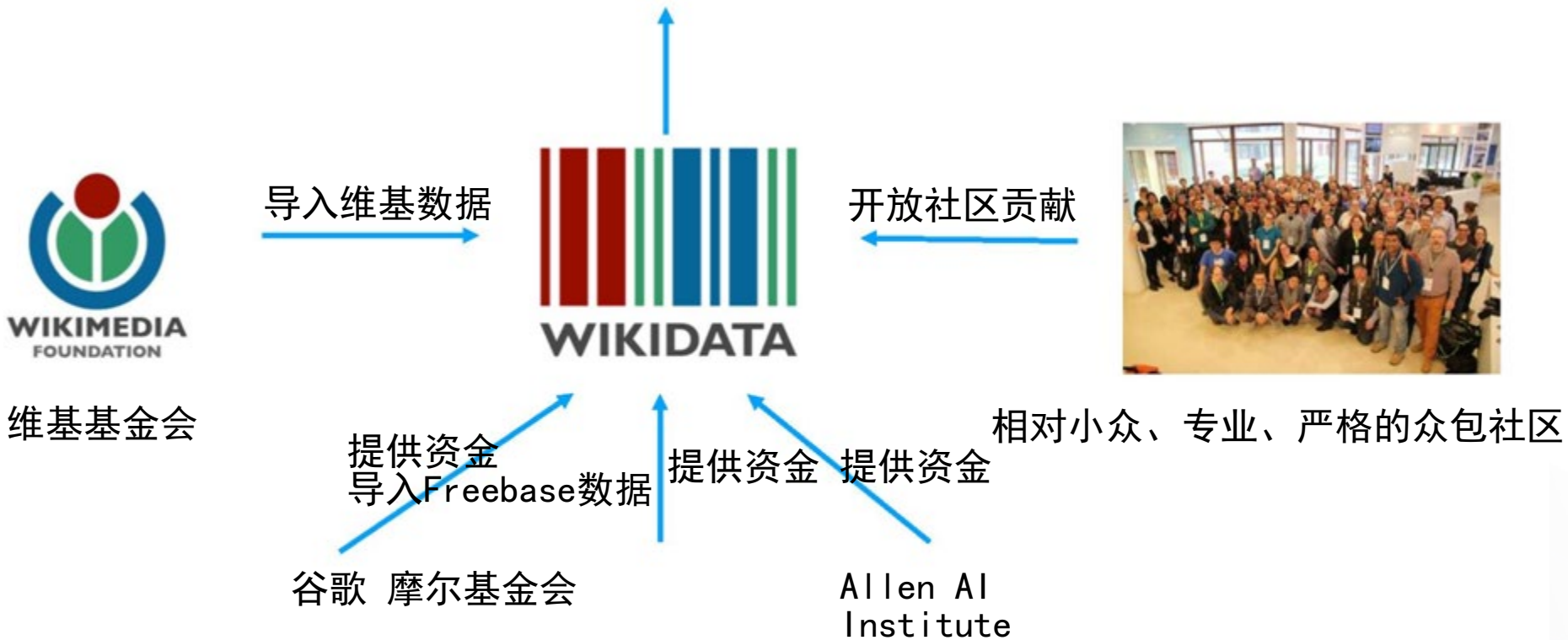
完全免费并允许商业化的开放许可协议



通过开源免费吸引用户贡献数据，增值的应用及技术服务收费

<http://www.freebase.com/>

目标是构建全世界最大的免费知识库，采用CC0完全自由许可协议



仍然面临知识缺失严重的问题

<http://www.wikidata.org/>

DBpedia是早期的语义网项目。DBpedia意指数据库版本的Wikipedia，是从Wikipedia抽取出来的链接数据集。DBpedia采用了一个较为严格的本体，包含人、地点、音乐、电影、组织机构、物种、疾病等类定义

此外，DBpedia还与Freebase，OpenCYC、Bio2RDF等多个数据集建立了数据链接。DBpedia采用RDF语义数据模型，总共包含30亿RDF三元组

<http://dbpedia.org>

YAGO是由德国马普研究所研制的链接数据库。YAGO主要集成了Wikipedia、WordNet和GeoNames三个来源的数据。YAGO将WordNet的词汇定义与Wikipedia的分类体系进行了融合集成，使得YAGO具有更加丰富的实体分类体系

YAGO还考虑了时间和空间知识，为很多知识条目增加了时间和空间维度的属性描述。目前，YAGO包含1.2亿条三元组知识。YAGO是IBM Watson的后端知识库之一

<http://mpii.de/yago>

微软 Concept Graph

Concept Graph是以概念层次体系为中心的知识图谱。与Freebase等知识图谱不同，Concept Graph以概念定义和概念之间的IsA关系为主

例：给定一个概念如“Microsoft”，Concept Graph返回一组与“微软”有IsA关系概念组，如：“Company”，“Software Company”，“Largest OS Vender”等。这被称为概念化“Conceptualization”

Concept Graph可以用于短文本理解和语义消歧中。例如，给定一个短文本“the engineer is eating the apple”，可以利用Concept Graph来正确理解其中“apple”的含义是“吃的苹果”还是“苹果公司”。微软发布的第一个版本包含超过540万的概念，1255万的实体，和8760万的关系。Concept Graph主要通过从互联网和网络日志中挖掘来构建



<http://concept.research.microsoft.com>

阿里知识图谱以商品、标准产品、标准品牌、标准条码、标准分类为核心，利用实体识别、实体链指和语义分析技术，整合关联了例如舆情、百科、国家行业标准等9大类一级本体，包含了百亿级别的三元组，形成了巨大的知识网。



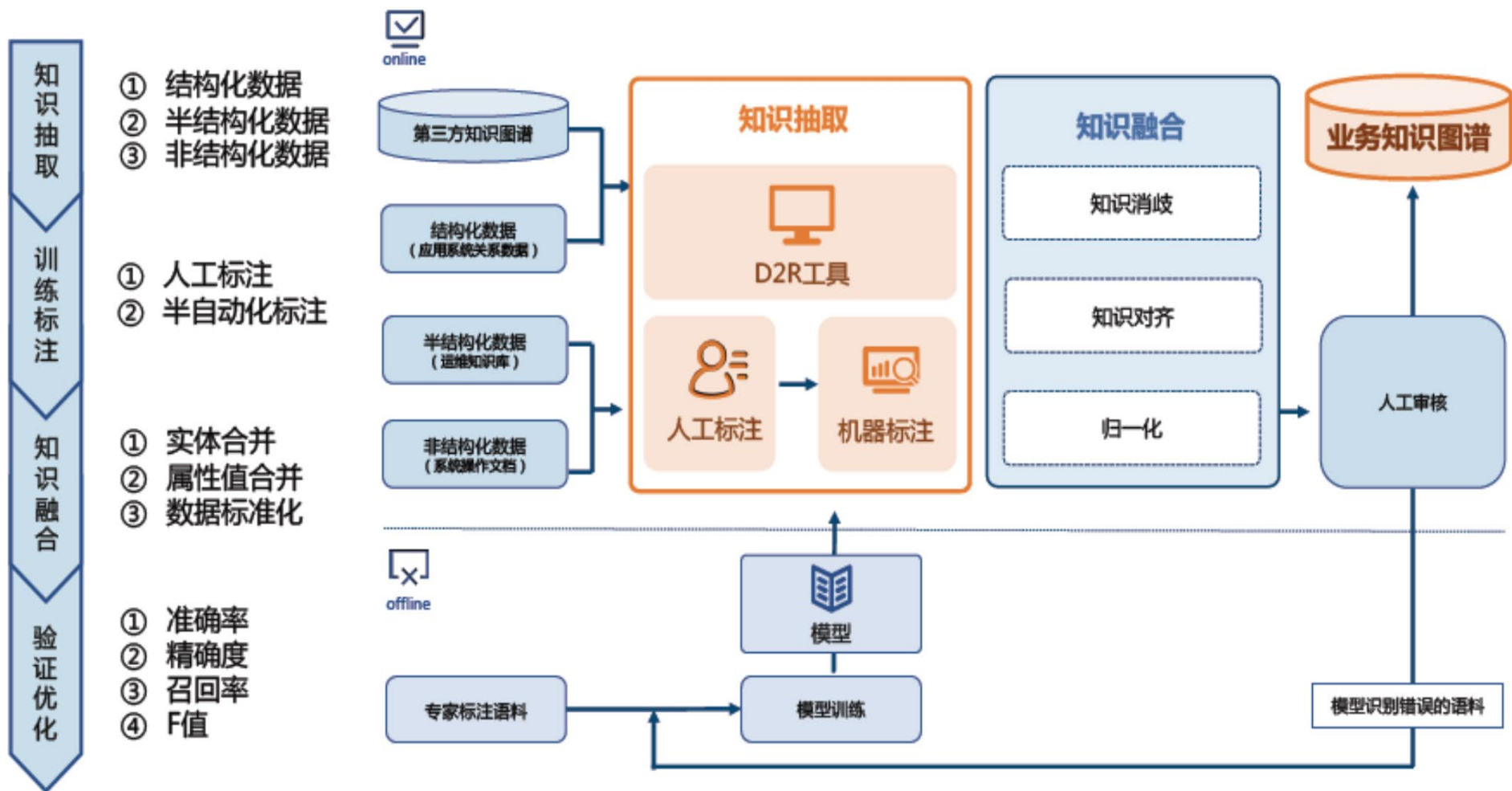
目录



CONTENTS

03

知识图谱构建



知识图谱构建业务流程





F-35战机本身具有良好的隐身性能，其雷达反射截面积约为0.1平方米，在微波雷达的探测下可最大限度实现隐身。但当雷达频率降至米波或升至毫米波时，雷达反射截面积将变大，暴露风险增加。

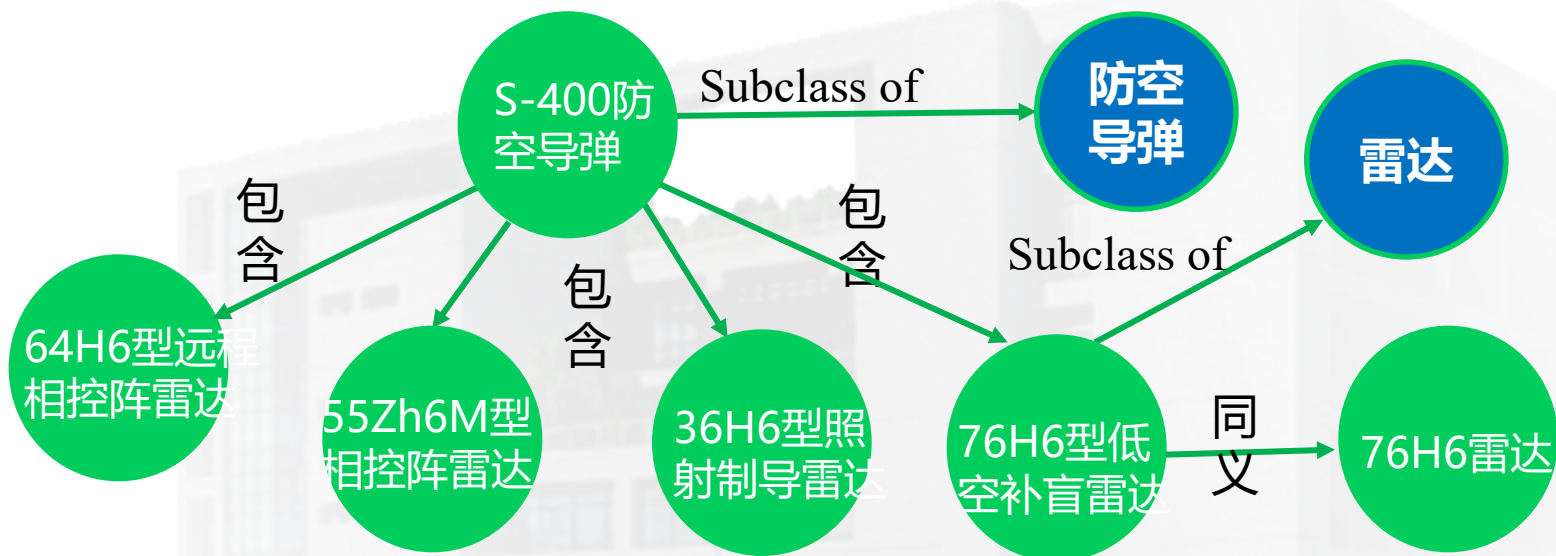
S-400防空导弹团级作战单位拥有64H6型远程相控阵雷达、55Zh6M (Nebo-M) 型相控阵雷达、36H6型照射制导雷达、76H6型低空补盲雷达。在环形扫描模式下，该雷达系统能够在300千米和600千米高度范围内追踪多达200个空中目标

难点和挑战:

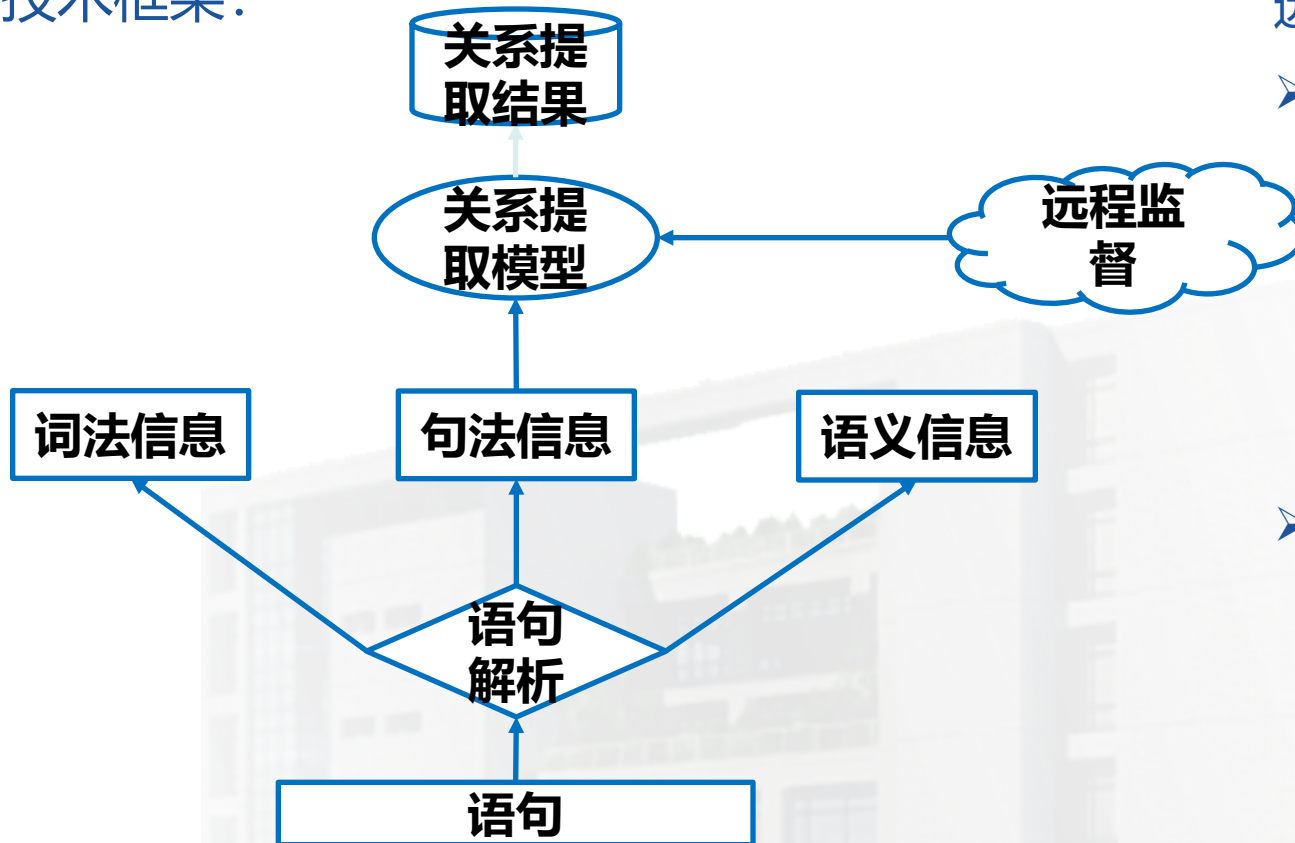
- 命名实体类型多样,数量众多,不断有新的命名实体涌现,如新的人名、地名等。
- 命名实体构成结构比较复杂,某些类型的命名实体词的长度没有一定的限制。
- 在不同领域、场景下,命名实体的外延有差异,可能产生歧义。

- 关系提取是从语句中获取**实体之间关系**的一种技术
- 实体对应知识图谱中的**结点**，关系则对应知识图谱中的**边**

S-400防空导弹团级作战单位拥有64H6型远程相控阵雷达、55Zh6M (Nebo-M)型相控阵雷达、36H6型照射制导雷达、76H6型低空补盲雷达



技术框架：



远程监督：

- 主要思想：将已有的知识对应到丰富的非结构化语料中从而生成大量的训练数据
- 知识来源：人工标注、现有的知识库、特定的语句结构

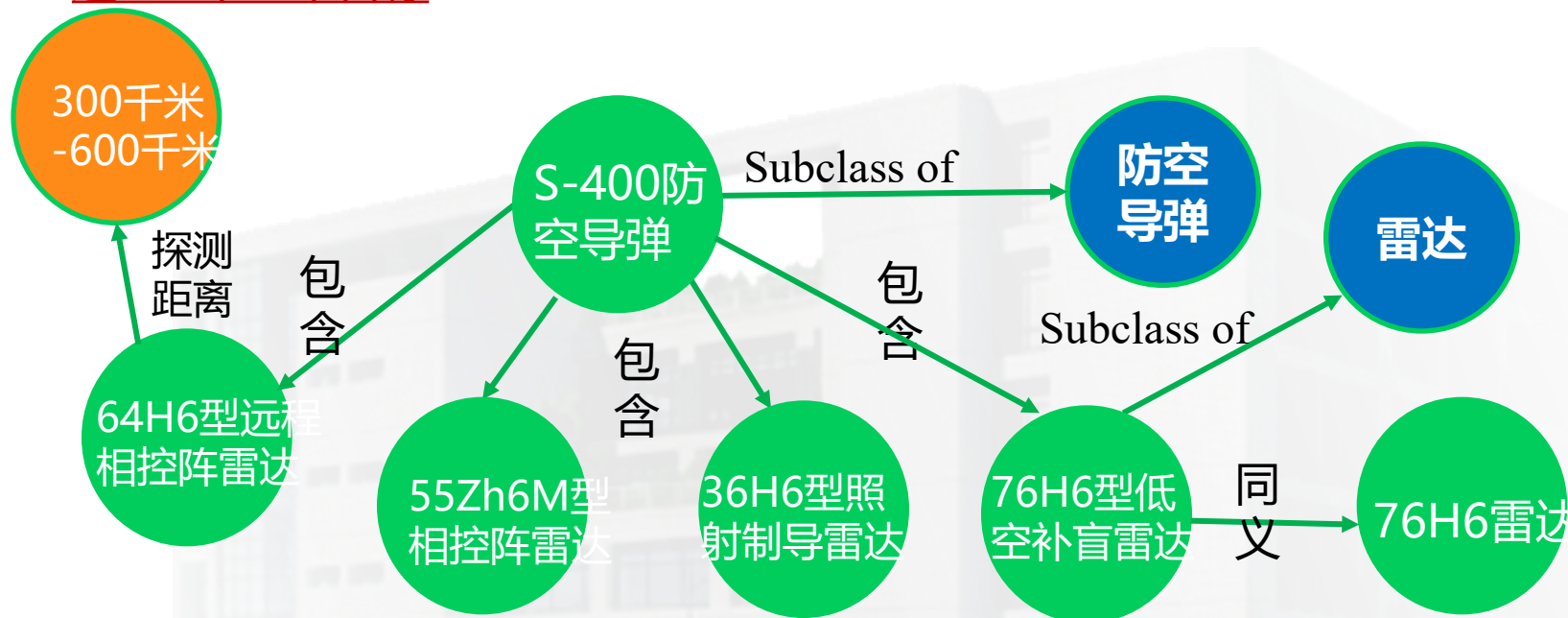


难点和挑战:

- 文本语句的结构复杂，人工制定规则工作量巨大
- 不同语言、不同领域之间存在差异，人工制定规则难以迁移
- 使用有监督学习的方法进行关系提取，人工标注训练样本成本巨大，如何获得足够多的高质量训练样本是一大难题

- 属性提取是从语句中获取实体属性的一种技术
- 实体对应知识图谱中的结点，属性则对应知识图谱中的属性

64H6型远程相控阵雷达能够在300千米和600千米高度范围内追踪多达200个空中目标



主要有两种存储方式：

- 一种是基于RDF（资源描述框架）的存储
- 另一种是基于图数据库的存储

- 存储三元组 (Triple)
- 标准的推理引擎
- W3C标准
- 易于发布数据
- 多数为学术界场景

RDF

- 节点和关系可以带有属性
- 没有标准的推理引擎
- 图的遍历效率高
- 事务管理
- 基本为工业界场景

图数据库

回顾 知识表示

- 如何表示知识是人工智能研究的一个重要议题
- 知识表示
 - 以形式化方式表示知识；
 - 供计算机自动处理；
- **【本章内容】**
 - 2.1 概述
 - 基本的知识表示方式
 - 2.2 谓词逻辑表示法★
 - 2.3 产生式表示法★
 - 2.4 语义网络表示法★
 - 2.5 框架表示法
 - 补充：知识图谱

The End