第三章
符号学派

知识表达、知识推理

符号学派

- 符号学派的代表，人工智能的创始人之一约翰•麦卡锡对人工智能符号学派理解：

> （人工智能）是关于如何制造智能机器，特别是智能的计算机程序的科学和工程。它与使用机器来理解人类智能密切相关，但人工智能的研究并不需要局限于生物学上可观察到的那些方法。

- 思想和观点直接继承自图灵，从功能的角度来理解智能
- 把智能看成为黑箱，只关心输入输出，假设知识已先验地存储于黑箱中
- 用知识表示和搜索来实现智能
- 擅长利用现有知识做推理、规划、逻辑运算和判断

符号主义

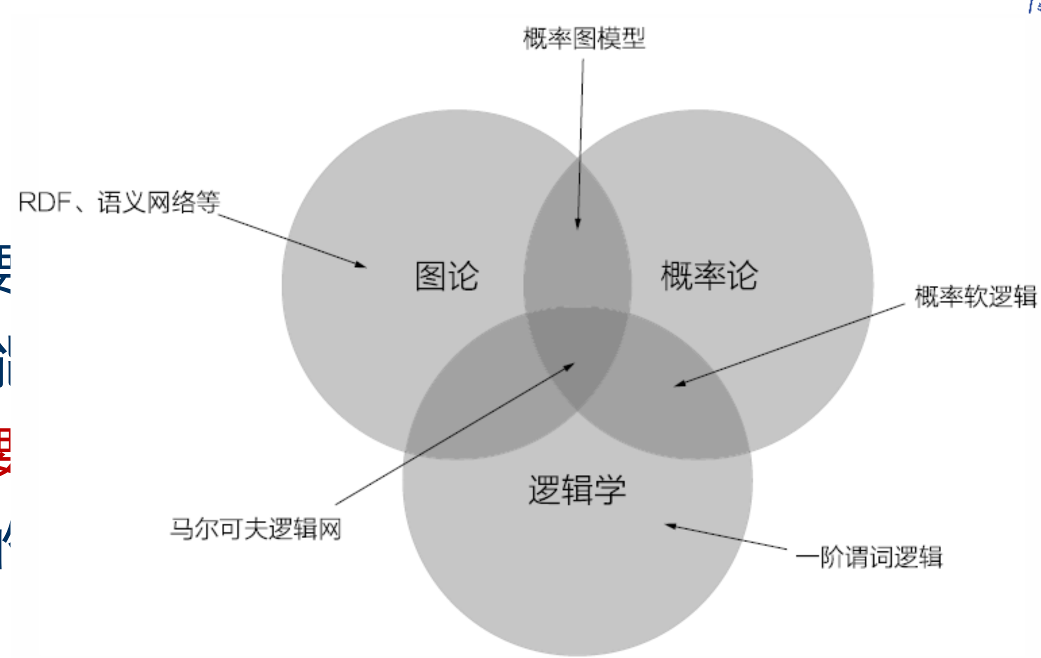
- 认知即计算
- 知识是信息的一种形式，是构成智能的基础
- 知识表示、知识推理、知识运用是人工智能的核心

Newell

Simon

AI System=Knowledge + Reasoning

知识表达



- 人类的智能活动主要
为了使计算机具有智
具有知识。但知识需要
机中去，因此，知识的
究课题。
- 本章首先介绍知识与知识表示的概念，然后介绍一阶谓词逻辑、产生式、框架、语义网络、知识图谱等当前人工智能中应用广泛的知识表示方法，为后面介绍推理方法、专家系统等奠定基础。

知识表达

- 人类的智能活动主要是获得并运用知识。知识是智能的基础。为了使计算机具有智能，能模拟人类的智能行为，就必须使它具有知识。但知识需要用适当的模式表示出来才能存储到计算机中去，因此，知识的表示成为人工智能中一个十分重要的研究课题。
- 本章首先介绍知识与知识表示的概念，然后介绍一阶谓词逻辑、产生式、框架、语义网络、知识图谱等当前人工智能中应用广泛的知识表示方法，为后面介绍推理方法、专家系统等奠定基础。

本章主要内容

1. 知识与知识表示的概念
2. 一阶谓词逻辑表示法
3. 产生式表示法
4. 框架表示法
5. 知识图谱

知识的概念

- 知识：在长期的生活及社会实践中、在科学研究及实验中积累起来的对客观世界的认识与经验。
- 知识：把有关信息关联在一起所形成的信息结构。
- 知识反映了客观世界中事物之间的关系，不同事物或者相同事物间的不同关系形成了不同的知识。

例如：

"雪是白色的"。—— 事实

"如果头痛且流涕，则有可能患了感冒"。—— 规则

知识的特性

- **1. 相对正确性**

　任何知识都是在一定的条件及环境下产生的，在这种条件及环境下才是正确的。

1+1=2　（十进制）
1+1=10（二进制）

人工智能中知识的相对正确性更加突出：

除了人类知识本身的相对正确性外，通常还将知识限制在所求解问题的范围内，所涉及的知识对求解问题是正确的就行。

知识的特性

- **2. 不确定性**

由于现实世界的复杂性，信息可能精确/模糊，关联可能确定/不确定，知识于是存在“真”的程度的问题。

➢ 随机性引起的不确定性	“如果头痛且流涕，则有可能患了感冒”
➢ 模糊性引起的不确定性	“今天天气很热。”到底多少度是热？
➢ 经验引起的不确定性	经验性的专家知识很难精确表述出来。
➢ 不完全性引起的不确定性	“不存在外星人”。

知识的特性

- **3. 可表示性与可利用性**

知识的可表示性：知识可以用适当形式表示出来，如用语言、文字、图形、向量等。

知识的可利用性：知识可以被利用。

知识表示就是将人类知识形式化或者模型化，把知识描述或约定成一种计算机可以接受并处理的数据结构。

知识表示方法与利用方法相关，不存在万能的知识表示模式。

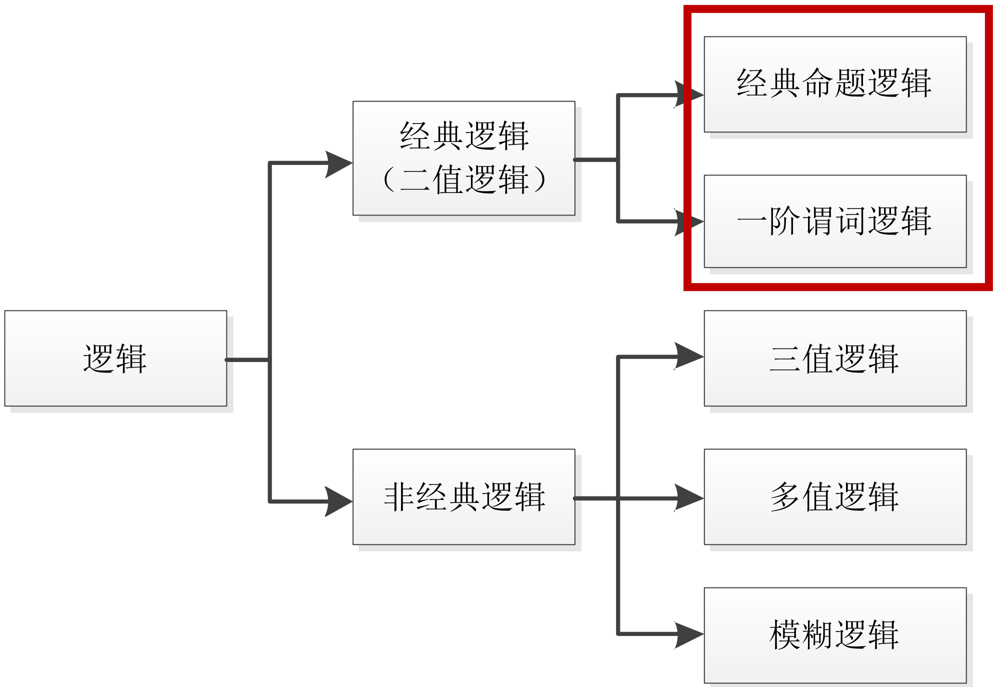
选择知识表示方法的原则：（1）充分表示领域知识。（2）有利于对知识的利用。（3）便于对知识的组织、维护与管理。（4）便于理解与实现。

本章主要内容

1. 知识与知识表示的概念
2. 一阶谓词逻辑表示法
3. 产生式表示法
4. 框架表示法
5. 知识图谱

一阶谓词逻辑表示法

- 逻辑（思维的规律和规则）
- 用谓词逻辑来理解语文/表示知识



任何一个命题的的真值必为“真”或“假”其中之一

最先应用于人工智能；

在知识的形式化表示、自动定理证明方面有重要作用

命题

- 命题逻辑是最基础的，是谓词逻辑的特殊形式
- 命题是一个非真即假的陈述句

> 若命题的意义为真，称它的真值为真，记为 T。
> 若命题的意义为假，称它的真值为假，记为 F。
> 一个命题可在一种条件下为真，在另一种条件下为假。

例如：3<5

例如：1+1=10

例如：太阳从西边升起

命题逻辑

- 命题逻辑：研究命题及命题之间关系的符号逻辑系统
  - 简单命题（原子命题）：简单陈述句表达的命题
  - 复合命题：通过否定、合取、析取、条件等连接词，将原子命题构成复合命题

$\neg P$

P：北京是中华人民共和国的首都

- 命题逻辑表示法的局限性：无法反映所描述事物的结构及逻辑特征，也不能将不同事物间的共同特征表述出来

A：老李是小李的父亲

P：李白是诗人
Q：杜甫也是诗人

用命题描述事物太粗糙了→谓词

词的分类

- “词”：语义的最小单位
- 词语的不同类型：
  - 语言学上的分类：
    - （词性）名词、动词、形容词、副词、介词等等
    - （句子中的成分）主、谓、宾、定、状、补等等
  - 谓词逻辑的构成：
    - 个体词、谓词、量词、逻辑连接词（连词）、标点符号、元语言词汇

词的意思 -> 词的关系 -> 句子的意思

个体词

- 一阶谓词逻辑中，将原子命题分成主语和谓语，于是有了个体词与谓词的概念
- 个体：独立存在的客体，具体事物，抽象事物（概念）
- 个体词≈给个体起的名字
  - 一般会用小写字母来指代确定的个体
- 个体可以是常量、变元、函数
  - 常量：表示一个或一组指定的个体
  - 变元：不指定的一个或一组个体
  - 函数：一个个体到另一个个体的映射

谓词

- 谓词：描述个体性质或关系的词
- 谓词的一般形式：　$P(x_1, x_2,\ldots, x_n)$

个体 $x_1, x_2,\ldots, x_n$：某个独立存在的事物或者某个抽象的概念；

谓词名 P：刻画个体的性质、状态或个体间的关系。

"老张是一个教师"："老张"→个体词 z，"是一个教师"→谓词名 T

$T(z)$ 描述了个体老张是一个教师这种性质，是一个一元谓词，其中个体词是常量。

"$x>y$"：二元谓词 $Greater(x, y)$，其中个体词是变元。

谓词

"老张的儿子作为一个教师为华科工作"：*Work* (*son(z)*, *hust*, *teacher*)，是三元谓词，其中*son(z)*这个个体是一个函数。

谓词	函数
有真/假	无真/假，是个体域中一个个体到另一个个体的映射

知识，非结构化的文字表达的语句 -> 谓词（一定的定义和解释） -> 结构化的表达 -> 便于计算机进行判断和推理

- 一阶谓词：谓词 $P(x_1, x_2,\ldots, x_n)$ 中，若x_i都是个体常量、变元或函数
- 二阶谓词：x_i又是一个一阶谓词。我们只讨论一阶谓词

量词

- 刻画谓词与个体间的关系，引入量词
  - 全称量词 $\forall x$： 对个体域中的所有（或任一个）个体 x；
  - 存在量词 $\exists x$： 在个体域中存在个体 x 。

All

Exist

$(\forall x)\ A(x)$: 对于所有的x来说，都有性质A

$(\exists x)\ A(x)$: 存在个体x，其有性质A。

- 量词非常重要，缺失了量词的句子，大多难判真假。

例如：把A解释为个头超过1米8， x变元在人群中选取个体，则：

$(\forall x)\ A(x)$ 为假， $(\exists x)\ A(x)$为真。

量词

- 全称量词和存在量词出现在同一个命题中，量词的次序也很重要

"男人比女人高" → $H(man, woman)$　　T/F？

给谓词逻辑前面加上量词：

1. $(\exists\, man)(\forall\, woman)$：世界上最高的那个人是男人。
2. $(\exists\, man)(\exists\, woman)$：世界上某一个男人比某一个女人要高。
3. $(\forall\, man)(\forall\, woman)$：世界上最矮的那个男人比最高的女人还高。
4. $(\forall\, man)(\exists\, woman)$：世界上最矮的那个人是女人。

逻辑连接词（连词）

- 无论命题逻辑还是谓词逻辑，均可用连接词把简单命题连接起来构成复合命题，表示更复杂的含义。
- 连词表示某种运算过程，所以又叫作逻辑运算符

（1）$\neg$： “否定” （ negation ）或“非”。

（2）$\vee$： “析取” （disjunction）—— “或”。

（3）$\wedge$： “合取” （conjunction）—— “与”。

（4）$\rightarrow$： “蕴含”（implication）或“条件”（condition）。

（5）$\leftrightarrow$ ： “等价” （equivalence）或“双条件” （bicondition）

逻辑连接词（连词）

（1）¬：“否定”（ negation ）/ “非”

- 否定位于它后面的命题。P为真，¬P为假；P为假，¬P为真。

“机器人不在2号房间”：$\neg$ *Inroom* (*robot*, *r*2)

（2）∨：“析取”（disjunction）—— “或”。

- 表示连接的两个命题具有“或”的关系。两个命题只要有一个为真，析取的结果就为真，两个命题同为假时，结果才为假。

“李明打篮球或踢足球”：

Plays (*liming*, *basketball*) ∨ *Plays* (*liming*, *football*)

逻辑连接词（连词）

（3）∧： “合取”（conjunction）——“与”。

- 表示连接的两个命题具有“与”的关系
- 两个命题只要有一个为假，合取的结果就为假，两个命题同为真时，结果才为真。

“我喜欢音乐和绘画”：

$$Like\ (I, music) \wedge Like\ (I, painting)$$

“小李住在一栋黄色的房子里”：

$$Live\ (li, house) \wedge Color\ (house, yellow)$$

逻辑连接词（连词）

（4）→：“**蕴含**”(implication) 或“**条件**”(condition)。

- P→Q: 如果P，则Q。P称为条件的**前件**，Q称为条件的**后件**。

“如果刘华跑得最快，那么他取得冠军。”：

RUNS (*liuhua*，*faster*)→*WINS* (*liuhua*，*champion*)

蕴含和汉语中的“如果……则……”有区别！→前后的命题可以没有意思上的关联。

太阳从西边出来→雪是白的：

Rise (*sun*，*west*) → *Color* (*snow*，*white*)　　**T / F ?**

- 只有前件为真、后件为假时，蕴含的结果才为假，其余均为真。

逻辑连接词（连词）

- **只有前件为真、后件为假时，蕴含的结果才为假，其余均为真。**

太阳从x边出来→雪是y的：

$Rise(sun, east) \rightarrow Color(snow, white)$	T→T	T
$Rise(sun, east) \rightarrow Color(snow, red)$	T→F	F
$Rise(sun, west) \rightarrow Color(snow, white)$	F→T	T
$Rise(sun, west) \rightarrow Color(snow, red)$	F→F	T

形式逻辑，定义出来的： $A \rightarrow B \leftrightarrow \neg A \vee B$

吹牛例子：如果你能赢我，我就跟你姓。

不接受例子：如果你考上清华，就给你买最新手机。

逻辑连接词（连词）

- **只有前件为真、后件为假时，蕴含的结果才为假，其余均为真。**

T → T	T
T → F	F
F → T	T

（1）你赢了（A真），我也跟你姓了（B真），A → B为真，我敢吹敢当，响当当。
（2）你赢了（A真），我耍赖不跟你姓（B假），A → B为假，我言而无信被人笑话。
（3）你输了（A假），这时候我无论怎么做（B真/假），牛皮都没吹破，A → B总是真！
掌握了这个逻辑，我们就可以到处吹牛了……

式逻

吹牛例子：如果

不接受例子：如

（1）我考上了清华（A真），爸爸也给我买了新手机（B真），A → B为真，皆大欢喜。
（2）我考上了清华（A真），爸爸耍赖不给买（B假），A → B为假，因为这个真的接受不了！
（3）我没考上清华（A假），爸爸给买当然好，不给买也没话说，都行吧。A → B都是真。

逻辑连接词（连词）

（5）↔：“**等价**”(equivalence)或“**双条件**”(bicondition)

- P ↔ Q表示：P当且仅当Q

谓词逻辑真值表

P　Q	¬P	P∨Q	P∧Q	P→Q	P↔Q
T　T	F	T	T	T	T
T　F	F	T	F	F	F
F　T	T	T	F	T	F
F　F	T	F	F	T	T
	否定	析取	合取	蕴含	等价

谓词公式

- 可按下述规则得到谓词演算的谓词公式

(1) 单个谓词是谓词公式，称为原子谓词公式。

(2) 若A是谓词公式，则$\neg$A也是谓词公式。

(3) 若A，B都是谓词公式，则A$\wedge$B，A$\vee$B，A$\rightarrow$B，A$\leftrightarrow$B也都是谓词公式。

(4) 若A是谓词公式，则$(\forall x)A$，$(\exists x)A$也是谓词公式。

(5) 有限步应用(1)-(4)生成的公式也是谓词公式。

连接词的优先级别从高到低排列：

$\neg$， $\wedge$， $\vee$， $\rightarrow$， $\longleftrightarrow$

谓词公式

- 量词的辖域
  - 量词的辖域：位于量词后面的单个谓词或者用括号括起来的谓词公式。
  - 约束变元与自由变元：辖域内与量词中同名的变元称为约束变元，不同名的变元称为自由变元。

例如：

$$\exists x\,(P(x, y) \rightarrow Q\,(x, y)) \vee R(x, y)$$

$(P(x, y) \rightarrow Q\,(x, y))$：$\exists x$ 的辖域，辖域内的变元 x 是受$(\exists\, x)$约束的变元，$R(x, y)$中的 x 是自由变元。

公式中的所有 y 都是自由变元。

谓词公式的性质

1. 谓词公式的解释

- 谓词公式在个体域上的解释：个体域中的实体对谓词演算表达式的每个常量、变量、谓词和函数符号的指派。
- 对于每一个解释，谓词公式都可求出一个真值（T或F）。

命题逻辑：各个命题变元指派真值（解释），然后通过逻辑运算可求命题公式的真值。

谓词逻辑：含有个体变元和函数，因此首先要考虑它们在个体域中的取值，然后才能对谓词指派真值。

谓词公式的性质

2. 谓词公式的永真性、可满足性、不可满足性

- 定义1：如果谓词公式 P 对个体域 D 上的任何一个解释都取得真值T，则称 P 在 D 上是永真的；如果 P 在每个非空个体域上均永真，则称 P 永真。

- 定义2：如果谓词公式 P 对个体域 D 上的任何一个解释都取得真值F，则称 P 在 D 上是永假的；如果 P 在每个非空个体域上均永假，则称P永假。

- 定义3：对于谓词公式 P，如果至少存在一个解释使得 P 在此解释下的真值为T，则称 P 是可满足的，否则，则称 P 是不可满足的。

谓词公式的性质

2. 谓词公式的永真性、可满足性、不可满足性

- 定义1：如果谓词公式 P 对个体域 D
值T，则称 P 在 D 上是永真的；如果
则称 P 永真。

- 定义2：如果谓词公式 P 对个体域 真
值F，则称 P 在 D 上是永假的；如果 P 在 上均永假，
则称P永假。

- 定义3：对于谓词公式 P，如果至少存在一个解释使得 P 在此解释下的真值为T，则称 P 是可满足的，否则，则称 P 是不可满足的。

判断永真永假，必须对每个个体域上的所有解释逐一判定，当解释的个数无限时，公式的永真永假性就很难判定了。

谓词公式的性质

3. 谓词公式的等价性

- 定义4：设P与Q是两个谓词公式，D是它们共同的个体域，若对D上的任何一个解释，P与Q都有相同的真值，则称公式P和Q在D上是等价的。如果D是任意个体域，则称P和Q是等价的，记为$P \leftrightarrow Q$ 。

（1）交换律： $P \vee Q \leftrightarrow Q \vee P$　　$P \wedge Q \leftrightarrow Q \wedge P$

（2）结合律： $(P \vee Q) \vee R \leftrightarrow P \vee (Q \vee R)$　　$(P \wedge Q) \wedge R \leftrightarrow P \wedge (Q \wedge R)$

（3）分配律： $P \vee (Q \wedge R) \leftrightarrow (P \vee Q) \wedge (P \vee R)$

$P \wedge (Q \vee R) \leftrightarrow (P \wedge Q) \vee (P \wedge R)$

谓词公式的性质

（4）德·摩根(De Morgen)定律： $\neg(P \vee Q) \leftrightarrow \neg P \wedge \neg Q$

$\neg(P \wedge Q) \leftrightarrow \neg P \vee \neg Q$

（5）双重否定律（对合律）： $\neg\neg P \leftrightarrow P$

（6）吸收律： $P \vee (P \wedge Q) \leftrightarrow P$ $P \wedge (P \vee Q) \leftrightarrow P$

（7）补余律（否定律）： $P \vee \neg P \leftrightarrow T$ $P \wedge \neg P \leftrightarrow F$

（8）连接词化规律（蕴含、等价等值式）： $P \rightarrow Q \leftrightarrow \neg P \vee Q$

（9）逆否律： $P \rightarrow Q \leftrightarrow \neg Q \rightarrow \neg P$

（10）量词转换律： $\neg(\exists x) P \leftrightarrow (\forall x)(\neg P)$ $\neg(\forall x) P \leftrightarrow (\exists x)(\neg P)$

（11）量词分配律： $(\forall x)(P \wedge Q) \leftrightarrow (\forall x)P \wedge (\forall x)Q$ 全称对应与，存在对应或

$(\exists x)(P \vee Q) \leftrightarrow (\exists x)P \vee (\exists x)Q$

谓词公式的性质

非(P且Q) = (非P)或(非Q)
非(P或Q) = (非P)且(非Q)

（4）德·摩根(De Morgen)定律： $\neg(P \vee Q) \leftrightarrow \neg P \wedge \neg Q$

$\neg(P \wedge Q) \leftrightarrow \neg P \vee \neg Q$

（5）双重否定律（对合律）：

（6）吸收律： $P \vee (P \wedge$

（7）补余律（否定律）：P

（8）连接词化规律（蕴含、等价

否定符移到量词后面时，全称量词变为存在量词，存在量词变为全称量词。
例： P：个头超过1米8.
$(\exists x)$ P：存在一个人个头超过1米8。
$\neg(\exists x) P \leftrightarrow (\forall x)(\neg P)$：
所有人个头都不超过1米8.

（9）逆否律： $P \rightarrow Q \leftrightarrow \neg Q \rightarrow \neg P$

（10）量词转换律： $\neg(\exists x) P \leftrightarrow (\forall x)(\neg P)$ $\neg(\forall x) P \leftrightarrow (\exists x)(\neg P)$

（11）量词分配律： $(\forall x)(P \wedge Q) \leftrightarrow (\forall x)P \wedge (\forall x)Q$ 全称对应与，存在对应或

$(\exists x)(P \vee Q) \leftrightarrow (\exists x)P \vee (\exists x)Q$

谓词公式的性质

- 4. 谓词公式的永真蕴含

▪定义5：对于谓词公式P与Q，如果P→Q永真，则称公式P永真蕴含Q，记作$P \Rightarrow Q$，且称Q为P的逻辑结论，称P为Q的前提。

一些永真蕴含式是进行演绎推理的重要规则

➢ **假言推理** $P, P \to Q \Rightarrow Q$ 如果P为真以及P→Q为真，则可推出Q为真。

> 例如：
>
> P：今天下雨了。Q：地面湿了。如果今天确实下雨了，并且因为下雨地面会湿，所以可推出地面湿了。

谓词公式的性质

- **拒取式推理** $\neg Q, P\rightarrow Q \Rightarrow \neg P$ 如果Q为假以及P→Q为真，则可推出P为假。

例如：

P：今天下雨了。Q：地面湿了。地面没湿，因为下雨地面会湿，所以可推出今天没下雨。

- **假言三段论** $P\rightarrow Q, Q\rightarrow R \Rightarrow P\rightarrow R$ 如果P→Q及Q→R为真，则可推出P→R为真。

例如：

P→Q 燕子是一种鸟　　　Q→R 鸟都有羽毛

P→R 所以燕子是有羽毛的

谓词公式的性质

- **全称固化**　$(\forall x)P(x) \Rightarrow P(y)$　y为个体域中的任一个体。

例如：

$(\forall x)P(x)$：所有人都会die…　y当然也会die …

- **存在固化**　$(\exists x)P(x) \Rightarrow P(y)$　y为个体域中某一个可使$P(y)$为真的个体。

例如：

$(\exists x)P(x)$：有人长得很好看…　y代表刘亦菲长得很好看…

谓词公式的性质

- **反证法**：$P \Rightarrow Q$，当且仅当 $P \wedge \neg Q \leftrightarrow \boldsymbol{F}$。
- 即Q为P的逻辑结论，当且仅当 $P \wedge \neg Q$是不可满足的。

例如： P：路边树上有果子　　Q：果子难吃

（果子好吃早被摘光了）$P \wedge \neg Q$：树上有果子且果子好吃 不可能 $\boldsymbol{F}$

$P \Rightarrow Q$：所以果子难吃。

定理：

Q为P_1，P_2，…，P_n的逻辑结论，当且仅当 $(P_1 \wedge P_2 \wedge \ldots \wedge P_n) \wedge \neg Q$ 是不可满足的。

一阶谓词逻辑知识表示方法

谓词公式表示知识的步骤：

（1）定义谓词及个体；

（2）变元赋值；

（3）用连接词连接各个谓词，形成谓词公式。

例如，用一阶谓词逻辑表示“每个存钱的人都得到利息”：

（1）定义谓词及个体：

个体x表示人，谓词$S(x)$表示存钱，谓词$I(x)$表示获得利息。

（2）变元赋值： $\forall x$

（3）连接谓词：$(\forall x)(S(x) \rightarrow I(x))$

一阶谓词逻辑知识表示方法

谓词公式表示知识的步骤：

（1）定义谓词及个体；

（2）变元赋值；

（3）用连接词连接各个谓词，形成谓词公式。

一阶谓词表示方法并不唯一：

（1）定义个体 x 表示人，y 表示钱，$S(x,y)$表示 x 存了 y 钱，$I(u)$表示 u 是利息，$O(x,u)$表示 x 获得了u 钱

$$(\forall x)((\exists y)(S(x,y)) \rightarrow (\exists u)(I(u) \wedge O(x,u)))$$

一阶谓词逻辑知识表示方法

- 将自然语言翻译为谓词逻辑语言：

➢ 本课程的所有学生都很聪明。

$(\forall x)(C(x)\rightarrow S(x))$。对所有x来说，如果x是本课程学生，x就很聪明。

➢ 本课程的有些学生是女生。

$(\exists x)(C(x)\wedge G(x))$。存在个体x，x是本课程学生，且是女生。

➢ 这个世界上不存在龙。

$\neg(\exists x)L(x)$，不存在个体x，x是龙。

➢ 所有的父母生气时就会发脾气。

$(\forall x)(P(x)\wedge A(x)\rightarrow T(x))$。对于所有的x来说，如果x是父母并且生气，那么x就会发脾气。

華中科技大學
HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

一阶谓词逻辑知识表示方法

- 将自然语言翻译为谓词逻辑语言：

➢ 有些泳池要么就是不干净，要么就是很拥挤。

➢ 玫瑰花和梅花都是花。

➢ 有一些老师会很高兴，当且仅当有一些同学学习很好。

一阶谓词逻辑知识表示的特点

优点：

① 自然性
② 精确性
③ 严密性
④ 容易实现

局限性：

① 不能表示不确定的知识
② 组合爆炸
③ 效率低

应用：

- 自动问答系统（Green等人研制的QA3系统）
- 机器人行动规划系统（Fikes等人研制的STRIPS系统）
- 机器博弈系统（Filman等人研制的FOL系统）
- 问题求解系统（Kowalski等设计的PS系统）