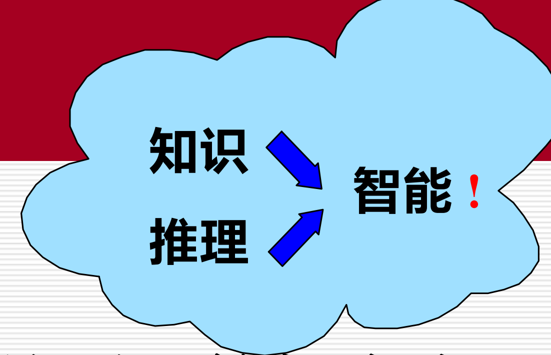
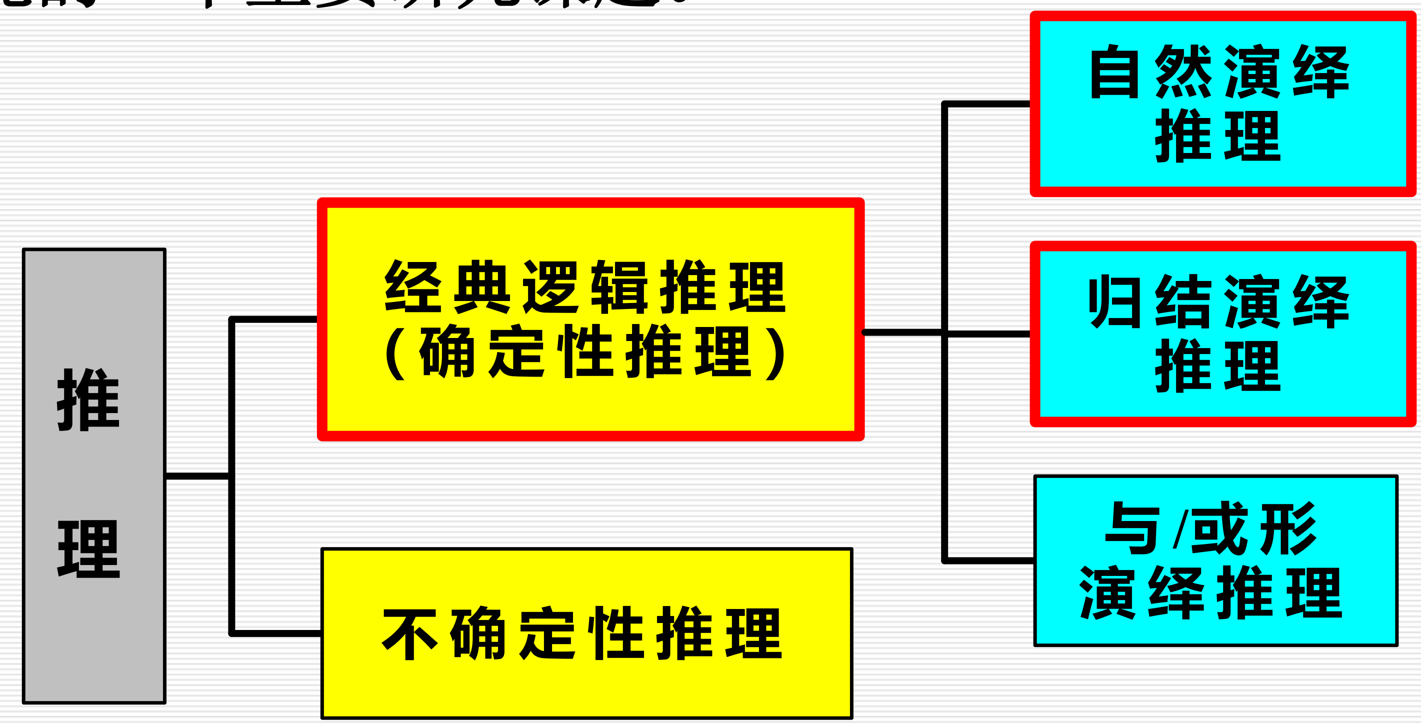
人工智能

第3章　确定性推理方法

第3章 确定性推理方法

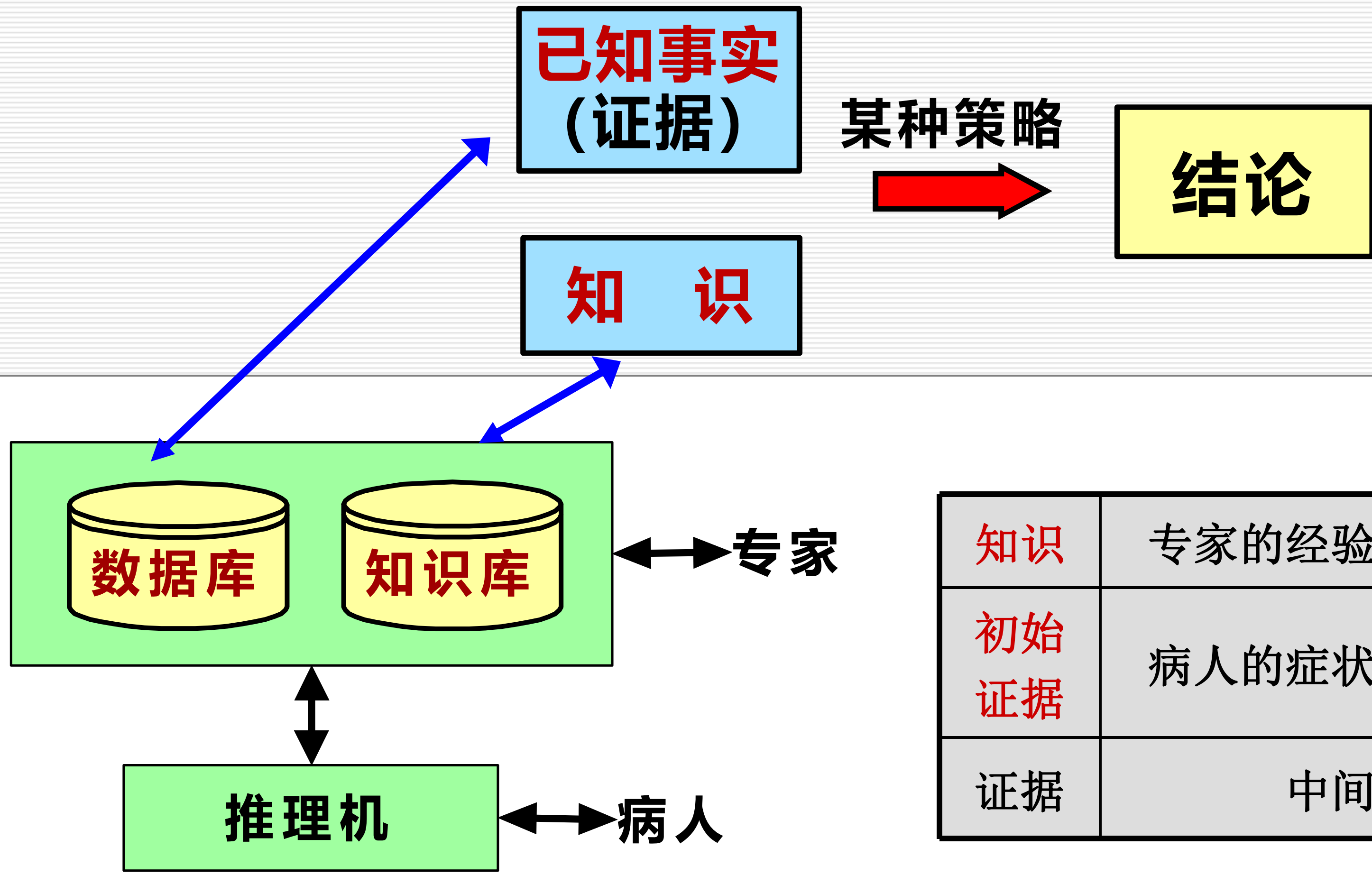


前面讨论了把知识用某种模式表示出来存储到计算机中去。但是，为使计算机具有智能，还必须使它具有思维能力。推理是求解问题的一种重要方法。因此，推理方法成为人工智能的一个重要研究课题。



3.1 推理的定义（reasoning）

所谓推理，就是从已知事实（证据）出发，按照某种策略，不断运用知识库中已有的知识，逐步匹配，直到推出结论的过程。



医疗专家系统

知识	专家的经验、医学常识
初始证据	病人的症状、化验结果
证据	中间结论

3.1 推理方式及其分类

演绎推理、归纳推理

（1）归纳推理 **(inductive reasoning):** 个别 → 一般

- 完全归纳推理（必然性推理）
- 不完全归纳推理（非必然性推理）

检查全部产品合格 —完全归纳推理→ 该厂产品合格

检查全部样品合格 —不完全归纳推理→ 该厂产品合格

3.1 推理方式及其分类

演绎推理、归纳推理

（2）演绎推理 (deductive reasoning)： 一般 → 个别

- 三段论式（三段论法）

① 足球运动员的身体都是强壮的； （大前提）

② 高波是一名足球运动员； （小前提）

---

③ 所以，高波的身体是强壮的。 （结　论）

3.1 推理方式及其分类

确定性推理、不确定性推理

（1）**确定性推理**：推理时所用的知识与证据都是确定的，推出的结论也是确定的，其真值或者为真或者为假。

（2）**不确定性推理**：推理时所用的知识与证据不都是确定的，推出的结论也是不确定的。

- 不确定性推理
  - 似然推理（概率论）
  - 近似推理或模糊推理（模糊逻辑）

3.1 推理方式及其分类

正向推理、逆向推理

- 正向推理（事实驱动推理）：已知事实 → 结论

- 正向推理：从已知的事实命题出发，采用逻辑推理规则，与知识库中的知识匹配，产生新的事实，直至结论出现在事实集中。常用的逻辑推理规则包括：
  - 假言推理： P，P→Q，=> Q。已知P为真，且P→Q，则Q为真。
  - 例如：铜是金属；如果x是金属，则x可以导电；铜可以导电
  - 拒式假言推理： P→Q，¬Q，=> ¬P。如果P→Q为真，但Q不成立，则P也不成立。
  - 例如：如果x是金属，则x可以导电；木头不导电；木头不是金属
- 正向推理，每前进一步总能得到“新的知识”。如果我们要证明的知识恰好被推出，则证明完成。

3.1 推理方式及其分类

- 实现正向推理需要解决的问题：
  - 确定匹配（知识与已知事实）的方法。
  - 按什么策略搜索知识库。
  - 冲突消解策略。
- 正向推理简单，易实现，但目的性不强，效率低。

3.1 推理方式及其分类

正向推理、逆向推理

- **逆向推理（目标驱动推理）：** 以某个假设目标作为出发点。

- 与正向推理相反的，就是反向推理。**反向推理**以某个假设目标作为出发点，来展开推理。

- 其基本思路是：首先选定一个假设目标，然后寻找使该目标成立的证据事实，如果所需的证据事实都能找到，则目标成立。否则，向前推进一步，将当前所需的证据事实作为新的假设，继续推理。

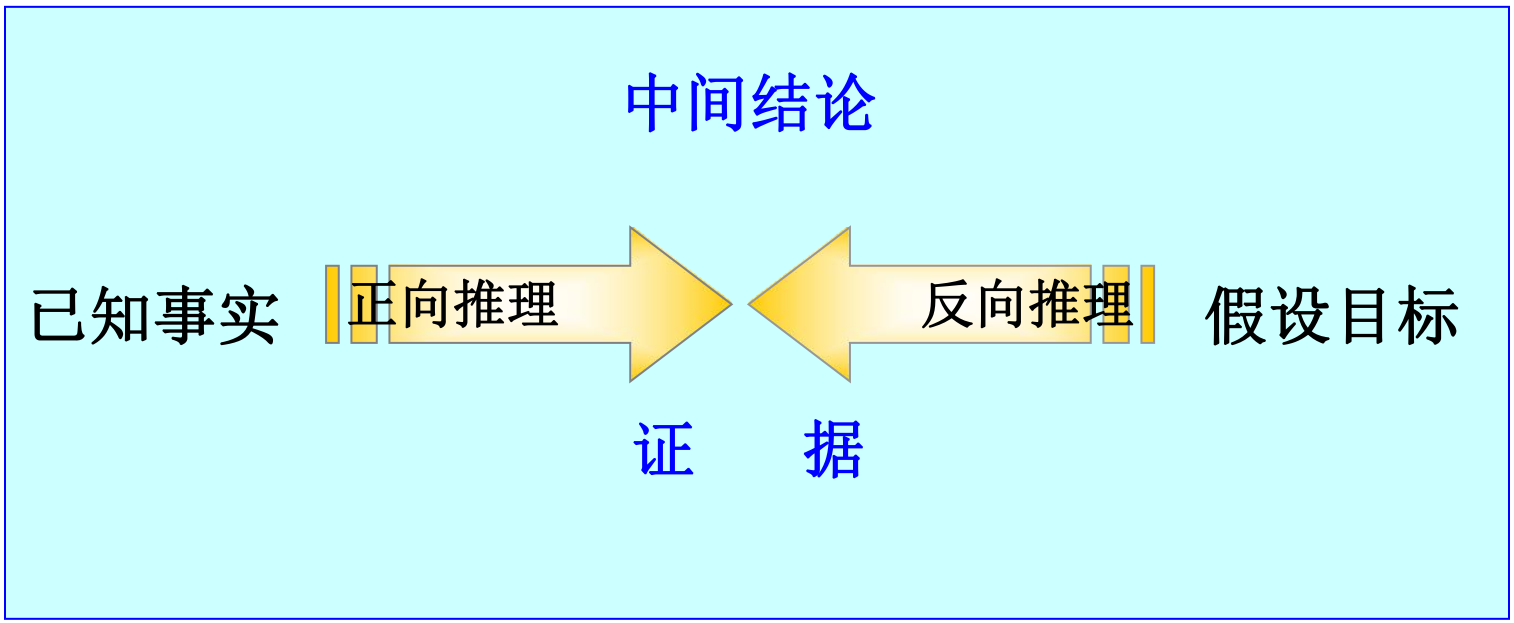
3.1 推理方式及其分类

- 逆向推理需要解决的问题：
  - 如何判断一个假设是否是证据？
  - 当导出假设的知识有多条时，如何确定先选哪一条？
  - 一条知识的运用条件一般都有多个，当其中的一个经验证成立后，如何自动地换为对另一个的验证？
  - ……..
- 逆向推理：目的性强，利于向用户提供解释，但选择初始目标时具有盲目性，比正向推理复杂。

3.1 推理方式及其分类

正向推理、逆向推理

- **双向推理：** 正向推理与逆向推理同时进行，且在推理过程中的某一步骤上“**碰头**”的一种推理。



3.1 如何将推理自动化

- 我们从几个维度来考察了人类智能推理的特征。
  - 推理的归纳和演绎；
  - 推理的正向和逆向；
  - 推理的确定和不确定。
- 可见，计算机如果要实现自动推理，最可行的方法就是采用反向、演绎的方法。
- 人工智能要模拟人的推理，也需要从上述几个角度来讨论可行性。
- 比如，第一个维度中，归纳和演绎，两种方法中，归纳方法做推理很难，但在规律发现、学习方面有优势。因此归纳方法适合于学习规律，而不适用于推理，实际上，机器学习可以看作一种归纳。
- 相比而言，演绎更适合描述人类推理。
- 其次，从推理的方向来看，
  - 正向推理，如果知识库中的规则和知识很多，则每推理一步都会产生大量的无用推理，因此不适合计算机。
  - 反向推理则可以在每一步都有具体的目标，目标单一，适合计算机处理。
- 最后，从确定性考虑，
  - 人工智能早期只涉及了确定性推理，如定理证明。
  - 随着产生式系统的发展，逐渐具备了不确定性推理的能力。

3.4 归结原理（消解原理）

- **Resolution Principle**
- **1965**年，美国数学家鲁滨逊提出归结原理。
- 基本出发点：要证明一个命题为真都可以通过证明其否命题为假来得到。
- **Resolution**：将多样的推理规则简化为一个。

什么叫消解

- 例1：
  - 小王说他下午或者去图书馆或者在家休息
  - 小王没去图书馆

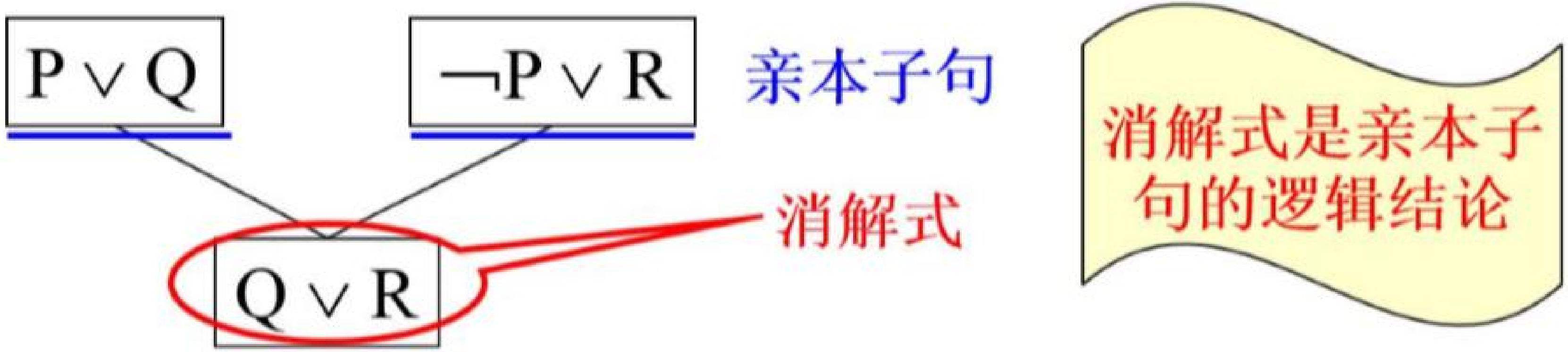
R——小王下午去图书馆

S——小王下午在家休息

$$\left.\begin{array}{l} R \vee S \\ \\ \neg R \end{array}\right\} \Rightarrow S$$

- 例2：
  - 如果今天不下雨，我就去你家 $\neg P \rightarrow Q \Leftrightarrow P \vee Q$
  - 今天没有下雨 $\neg P$

什么叫消解



- 消解只能在仅含否定和析取联接词的公式（子句）间进行
- 必须先把公式化成规范的形式（范式，子句集）

含变量的消解

- 例：苏格拉底论断

  凡人都会死.　　$\forall x\,(\mathrm{Man}\,(x) \to \mathrm{Mortal}\,(x))$

  苏格拉底是人.　　Man (Socrates)

  如何得到结论：苏格拉底会死.　　Mortal(Socrates)

- 要完成消解还面临几个问题
  - “∀”和“→”必须去掉
    - $\mathrm{Man}\,(x) \to \mathrm{Mortal}\,(x) \Leftrightarrow \neg \mathrm{Man}\,(x) \vee \mathrm{Mortal}(x)$
    - “∀”怎么办？

    化为子句集

  - 如果能去掉“∀”，$\neg$Man (x) 和Man (Socrates)也不能构成互补对，形式不一样，怎么办？

    置换与合一

一阶谓词的标准化问题

- 用归结原理实现定理证明是基于一阶谓词知识表示形式；
- “标准谓词子句”格式：
  - ✓ 每条知识之间均为合取关系
  - ✓ 每条知识内部的若干文字均为析取式
- 显然，要求专家在设计知识时按照标准格式很难，因此在1973年，提出一个变换序列，可以将任意形式的一阶谓词语句等价化简为“析取式的合取”，即标准谓词子句。

谓词公式化为子句集的方法

- **原子（atom）谓词公式：** 一个不能再分解的命题。
- **文字（literal）：** 原子谓词公式及其否定。
- P：正文字， $\neg P$：负文字。
- **子句（clause）：** 任何文字的**析取式**。任何文字本身也都是子句。

$$P(x)\vee Q(x),\qquad \neg P(x,f(x))\vee Q(x,g(x))$$

- **空子句**
- **子句集：** 由子句构成的集合

空子句是永假的，不可满足的。

谓词公式化为子句集的方法

- 例3.2： 将下列谓词公式化为子句集。

$$(\forall x)((\forall y)P(x,y) \rightarrow \neg(\forall y)(Q(x,y) \rightarrow R(x,y)))$$

- 解：（1）消去谓词公式中的“ $\rightarrow$ ” 符号

$$\Longleftrightarrow (\forall x)(\neg(\forall y)P(x,y) \vee \neg(\forall y)(\neg Q(x,y) \vee R(x,y)))$$

$P \rightarrow Q \Leftrightarrow \neg P \vee Q$, $P \leftrightarrow Q \Leftrightarrow (P \wedge Q) \vee (\neg P \wedge \neg Q)$

（2）把否定符号 $\neg$ 移到紧靠谓词的位置上

$$\Longleftrightarrow (\forall x)((\exists y)\neg P(x,y) \vee (\exists y)(Q(x,y) \wedge \neg R(x,y)))$$

双重否定律 $\neg(\neg P) \Leftrightarrow P$

狄.摩根律 $\neg(P \wedge Q) \Leftrightarrow \neg P \vee \neg Q, \neg(P \vee Q) \Leftrightarrow \neg P \wedge \neg Q$

量词转换律 $\neg(\exists x)P \Leftrightarrow (\forall x)\neg P$, $\neg(\forall x)P \Leftrightarrow (\exists x)\neg P$

谓词公式化为子句集的方法

$$(\forall x)((\exists y)\neg P(x,y) \vee (\exists y)(Q(x,y) \wedge \neg R(x,y)))$$

（3）变量标准化：保证每个量词有自己唯一的变量名

$$\Leftrightarrow \quad (\forall x)((\exists y)\neg P(x,y) \vee (\exists z)(Q(x,z) \wedge \neg R(x,z)))$$

$(\exists x)P(x) \equiv (\exists y)P(y), \quad (\forall x)P(x) \equiv (\forall y)P(y)$

（4）消去存在量词

a.

b.

$y = f(x),$

$z = g(x)$

$\Leftrightarrow$ $(\neg P(x, f(x)) \vee (Q(x, g(x)) \wedge \neg R(x, g(x))))$

对于一般情况

$$(\forall x_1)((\forall x_2)\cdots((\forall x_n)((\exists y)P(x_1, x_2, \cdots, x_n, y)))\ldots)$$

存在量词 y的Skolem 函数为 $y = f(x_1, x_2, \cdots, x_n)$

Skolem化：用Skolem函数代替每个存在量词量化的变量的过程。

谓词公式化为子句集的方法

- 例3.2 将下列谓词公式化为子句集。

$$(\forall x)((\forall y)P(x,y) \to \neg(\forall y)(Q(x,y) \to R(x,y)))$$

- 解：

$y = f(x),$

$z = g(x)$

$$\Leftrightarrow (\forall x)(\neg P(x, f(x)) \vee (Q(x, g(x)) \wedge \neg R(x, g(x))))$$

（5）化为前束形

前束形=（前缀）{母式}

（前缀）：全称量词串。
{母式}：不含量词的谓词公式。

此时，只有一个全称量词，且已经位于公式的最左边，所以，这一步不需要做任何工作。

$$(\forall x)(\neg P(x, f(x)) \vee (Q(x, g(x)) \wedge \neg R(x, g(x))))$$

（6）化为 **Skolem 标准形**

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

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

$(\forall x_1)(\forall x_2)\cdots(\forall x_n)M$，称为Skolem标准形的母式。

$\Longleftrightarrow (\forall x)((\neg P(x, f(x)) \vee Q(x, g(x))) \wedge (\neg P(x, f(x)) \vee \neg R(x, g(x))))$

（7）略去全称量词

（8）消去合取词

（9）子句变量标准化

谓词公式化为子句集的方法

（6）化为 **Skolem 标准形**

$\Longleftrightarrow (\forall x)((\neg P(x,f(x)) \vee Q(x,g(x))) \wedge (\neg P(x,f(x)) \vee \neg R(x,g(x))))$

（7）略去全称量词

$\Longleftrightarrow (\neg P(x,f(x)) \vee Q(x,g(x))) \wedge (\neg P(x,f(x)) \vee \neg R(x,g(x)))$

（8）消去合取词

$\Longleftrightarrow \{\neg P(x,f(x)) \vee Q(x,g(x)),\quad \neg P(x,f(x)) \vee \neg R(x,g(x))\}$

（9）子句变量标准化

$\Longleftrightarrow \{\neg P(x,f(x)) \vee Q(x,g(x)),\quad \neg P(y,f(y)) \vee \neg R(y,g(y))\}$

谓词公式化为子句集的方法

- **例3.3 将下列谓词公式化为子句集。**

$$(\forall x)\{[\neg P(x)\vee\neg Q(x)]\to(\exists y)[S(x,y)\wedge Q(x)]\}\wedge(\forall x)[P(x)\vee B(x)]$$

- **（1）消去蕴含符号**

$$(\forall x)\{\neg[\neg P(x)\vee\neg Q(x)]\vee(\exists y)[S(x,y)\wedge Q(x)]\}\wedge(\forall x)[P(x)\vee B(x)]$$

- **（2）把否定符号移到每个谓词前面**

$$(\forall x)\{[P(x)\wedge Q(x)]\vee(\exists y)[S(x,y)\wedge Q(x)]\}\wedge(\forall x)[P(x)\vee B(x)]$$

- **（3）变量标准化**

$$(\forall x)\{[P(x)\wedge Q(x)]\vee(\exists y)[S(x,y)\wedge Q(x)]\}\wedge(\forall w)[P(w)\vee B(w)]$$

- **（4）消去存在量词，设*y*的函数是*f(x)*，则**

$$(\forall x)\{[P(x)\wedge Q(x)]\vee[S(x,f(x))\wedge Q(x)]\}\wedge(\forall w)[P(w)\vee B(w)]$$

谓词公式化为子句集的方法

- 例3.3 将下列谓词公式化为子句集。（续）

（5）化为前束形

$$(\forall x)(\forall w)\{\{[P(x)\wedge Q(x)]\vee[S(x,f(x))\wedge Q(x)]\}\wedge[P(w)\vee B(w)]\}$$

（6）化为标准形

$$(\forall x)(\forall w)\{\{[Q(x)\wedge P(x)]\vee[Q(x)\wedge S(x,f(x))]\}\wedge[P(w)\vee B(w)]\}$$

$$(\forall x)(\forall w)\{Q(x)\wedge[P(x)\vee S(x,f(x))]\wedge[P(w)\vee B(w)]\}$$

（7）略去全称量词

$$Q(x)\wedge[P(x)\vee S(x,f(x))]\wedge[P(w)\vee B(w)]$$

（8）消去合取词，把母式用子句集表示

$$\{Q(x),\ P(x)\vee S(x,f(x)),\ P(w)\vee B(w)\}$$

（9）子句变量标准化 $\{Q(x),\ P(y)\vee S(y,f(y)),\ P(w)\vee B(w)\}$

谓词公式化为子句集的方法

- **例3.5** 将下列谓词公式化为不含存在量词的前束形。

$$(\exists x)(\forall y)((\forall z)(P(z) \wedge \neg Q(x,z)) \to R(x,y,f(a)))$$

- **（1）消去存在量词**

$$(\forall y)((\forall z)(P(z) \wedge \neg Q(b,z)) \to R(b,y,f(a)))$$

- **（2）消去蕴含符号**

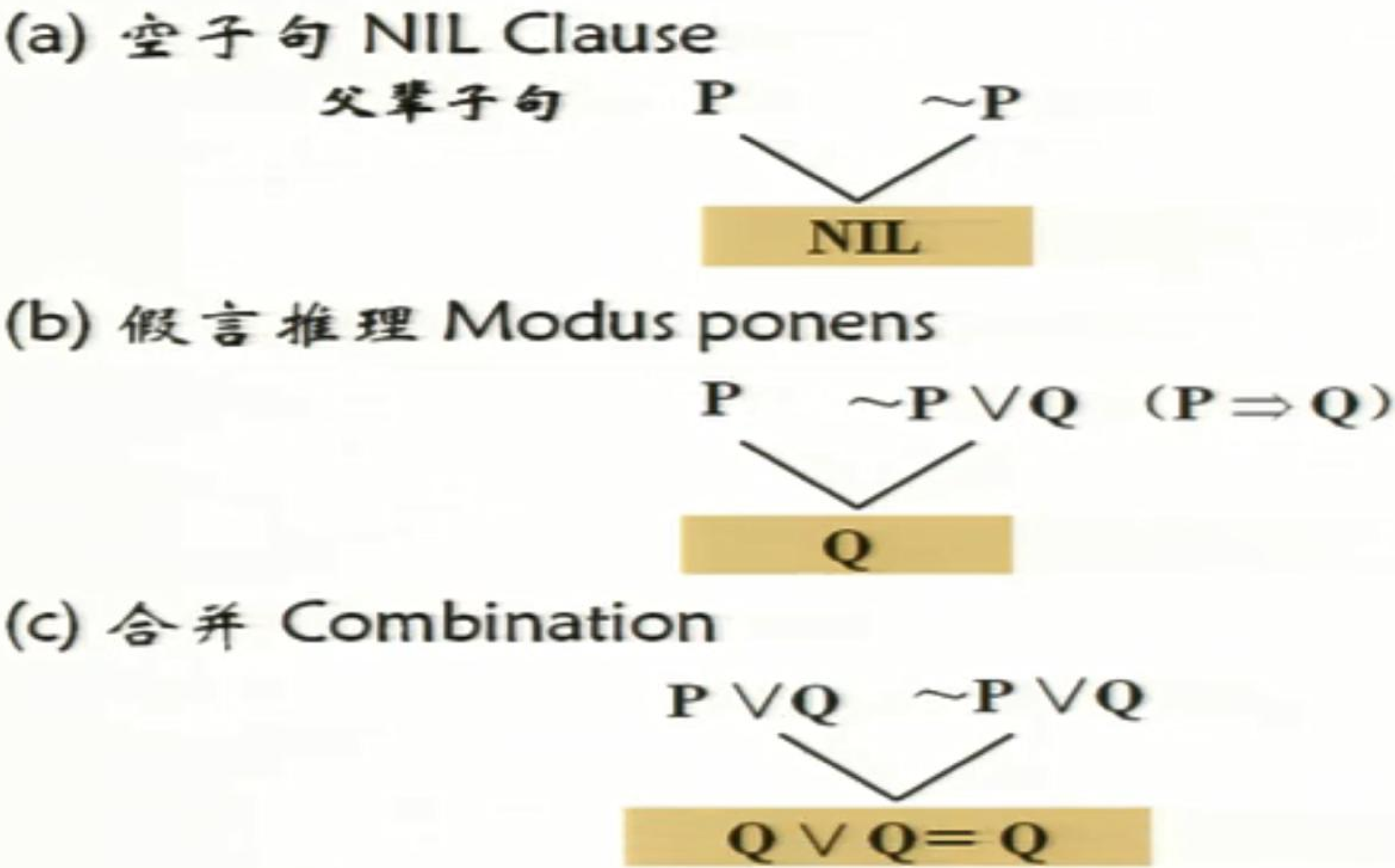
$$(\forall y)(\neg(\forall z)(P(z) \wedge \neg Q(b,z)) \vee R(b,y,f(a)))$$

$$(\forall y)((\exists z)(\neg P(z) \vee Q(b,z)) \vee R(b,y,f(a)))$$

- **（3）设z的函数是$g(y)$，则**

$$(\forall y)(\neg P(g(y)) \vee Q(b,g(y)) \vee R(b,y,f(a)))$$

消解推理规则



- 假言推理： P，P→Q， => Q。已知P为真，且P→Q，则Q为真。
- 例如：铜是金属 ； 如果x是金属，则x可以导电； 铜可以导电

消解推理规则

(d) 重言式 Tautologies

重言式又称为**永真式**

$$P \vee Q \qquad \sim P \vee \sim Q$$

$$\boxed{Q \vee \sim Q}$$

$$P \vee Q \qquad \sim P \vee \sim Q$$

$$\boxed{P \vee \sim P}$$

(e) 链式（三段论） Chain

$$\sim P \vee Q \qquad \sim Q \vee R$$

$$\boxed{\sim P \vee R}$$

含有变量的消解式

- 含有变量的子句的消解
  - 要把消解推理规则推广到含有变量的子句，必须找到一个作用于父辈子句的置换，使父辈子句含有互补文字。
- Example

$P[x,f(y)] \vee Q(x) \vee R[f(a),y]$ $\sim P[f(f(a)),z] \vee R(z,w)$

置换：用另外的变量置换该变量，用函数置换该变量，用常量置换该变量。

含有变量的消解式

- 含有变量的子句的消解
  - 要把消解推理规则推广到含有变量的子句，必须找到一个作用于父辈子句的置换，使父辈子句含有互补文字。
- Example

$P[x,f(y)] \vee Q(x) \vee R[f(a),y]$ $\qquad$ $\sim P[f(f(a)),z] \vee R(z,w)$

$\sigma = \{f(f(a))/x,\ f(y)/z\}$

$Q[f(f(a))] \vee R(f(a),\ y) \vee R(f(y),\ w)$

消解反演求解过程

可以把要解决的问题作为一个要证明的命题。消解通过反演产生证明。

消解反演证明

给出公式集{S}和目标公式L

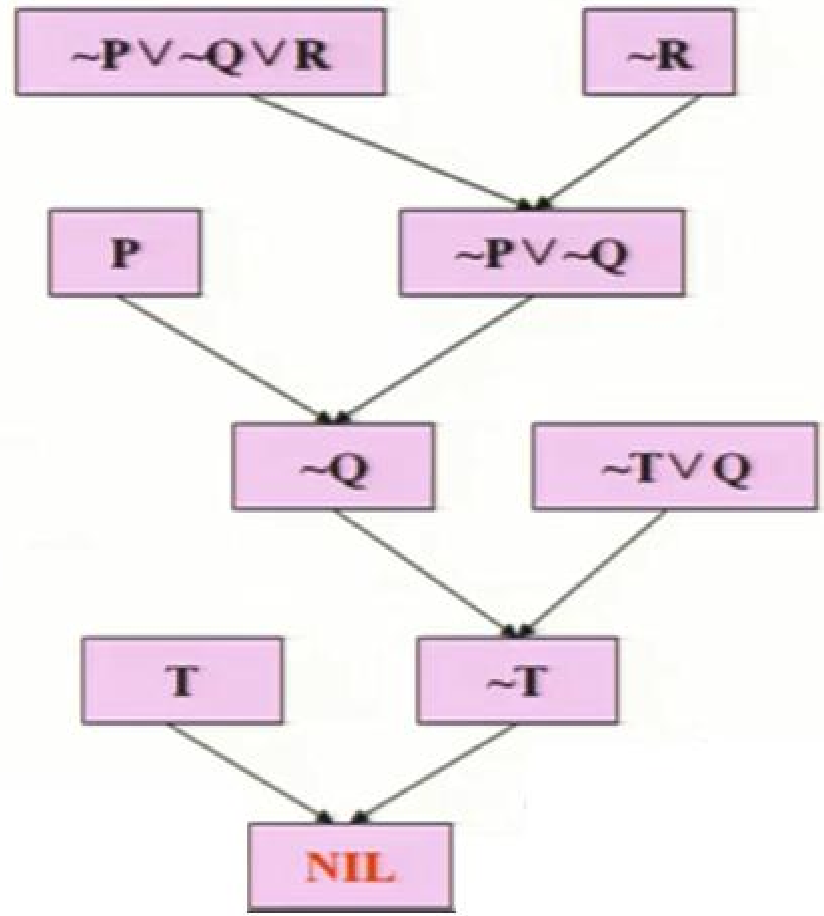
- 否定L，得$\sim L$；
- 把$\sim L$添加到S中去；
- 把新产生的集合$T=\{\sim L,S\}$化成子句集；
- 应用消解原理，力图推导出一个表示矛盾的空子句。

Example 1

- 设事实的公式集合
  {P，(P∧Q)⇒R，
  (S∨T)⇒Q，T}，
  证明：R

- 否定结论，将公式化为子句，得子句集：
- {P，~P∨~Q∨R，
  ~S∨Q，~T∨Q，T，
  ~R}

消解反演树



消解反演证明三段论

"Fido是狗"，

"所有狗都是动物"，

"所有动物都会死"，

求证"Fido会死"。

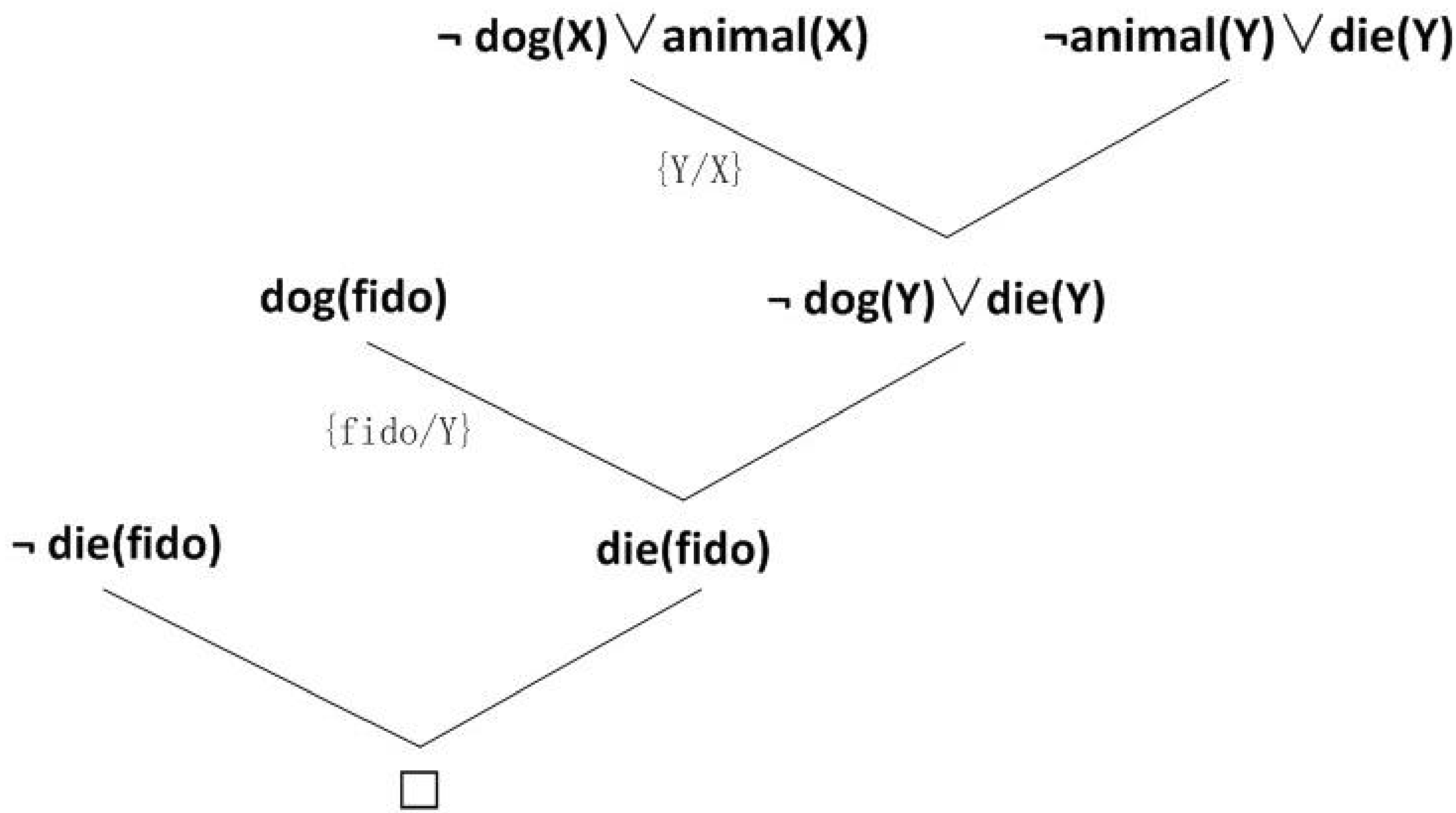
我们将已知条件表示为一阶谓词，并且将其中的蕴含符转换，得到知识库：

Fido 是狗	dog(fido)	dog(fido)
所有狗都是动物	$(\forall X)(dog(X) \rightarrow animal(X))$	$\neg dog(X) \lor animal(X)$
所有动物都会死	$(\forall Y)(animal(Y) \rightarrow die(Y))$	$\neg animal(Y) \lor die(Y)$

要证明的结论为die(fido)。将其取反，得到¬die(fido)，也加入到已知条件中。

消解反演证明三段论

(1)dog(fido)
(2)¬ dog(X) ∨ animal(X)
(3)¬animal(Y) ∨ die(Y)
(4)¬die(fido)



消解反演证明

例3.9 储蓄问题--P86

前提：每个储蓄钱的人都获得利息。

结论：如果没有利息，那么就没有人去储蓄钱

证明：

(1) 规定原子公式：

$S(x, y)$ 表示 "x储蓄y"

$M(x)$ 表示 "x是钱"

$I(x)$ 表示 "x是利息"

$E(x, y)$ 表示 "x获得y"

(2) 用谓词公式表示前提和结论：

反演求解过程

- 消解反演求解
  - 把由目标公式的否定产生的每个子句添加到目标公式否定之否定的子句中去。
  - 按照反演树，执行和以前相同的消解，直至在根部得到某个子句止。
  - 用根部的子句作为一个回答语句

求解：**Fido**在哪里--P86

"如果无论John到哪里去，Fido也就去那里，那么如果John在学校里，Fido在哪里呢？"

应用归结原理求解问题

- **例** 已知：

F_1：王（Wang）先生是小李（Li）的老师。

F_2：小李与小张（Zhang）是同班同学。

F_3：如果x与y是同班同学，则 x 的老师也是 y 的老师。

求：小张的老师是谁？

应用归结

F_1：王（Wang）先生是小李（Li）的老师。

F_2：小李与小张（Zhang）是同班同学。

F_3：如果 x 与 y 是同班同学，则 x 的老师也是 y 的老师。

求：小张的老师是谁？

- 解：
  - 定义谓词：

    $T(x, y)$： x 是 y 的老师。

    $C(x, y)$： x 与 y 是同班同学。
  - 把已知前提表示成谓词公式：

    F_1： $T(Wang, Li)$

    F_2： $C(Li, Zhang)$

    F_3： $(\forall x)(\forall y)(\forall z)(C(x, y) \wedge T(z, x) \rightarrow T(z, y))$
  - 把目标表示成谓词公式，并把它否定后与 *ANSWER* 析取：

    G： $\neg (\exists x) T(x, Zhang) \vee ANSWER(x)$

应用归结原理求解问题

- 把上述公式化为子句集：

（1） $T(Wang, Li)$

（2） $C(Li, Zhang)$

（3） $\neg C(x, y) \vee \neg T(z, x) \vee T(z, y)$

（4） $\neg T(u, Zhang) \vee ANSWER(u)$

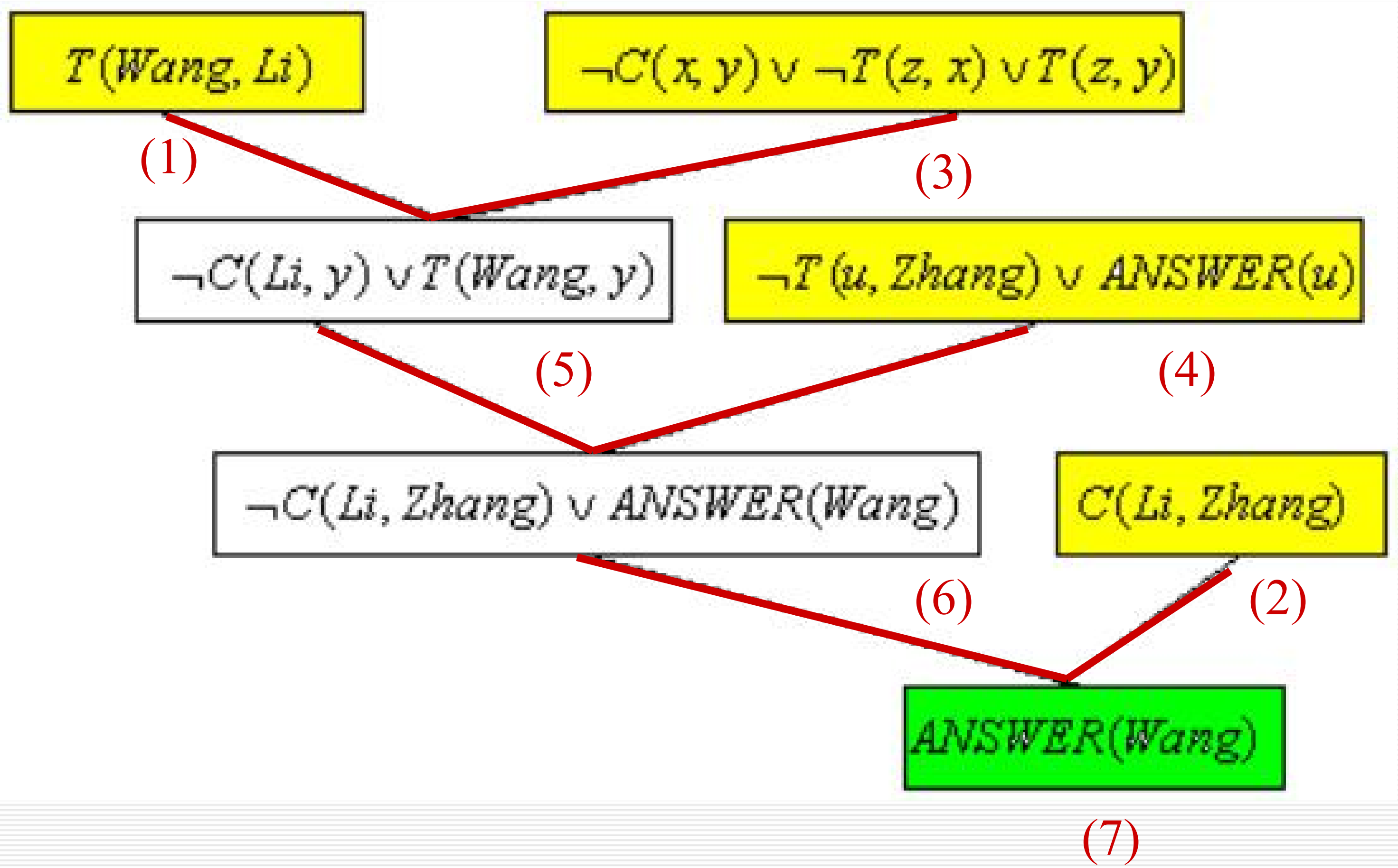
- 应用归结原理进行归结：

（5） $\neg C(Li, y) \vee T(Wang, y)$ （1）与（3）归结

（6） $\neg C(Li, Zhang) \vee ANSWER(Wang)$ （4）与（5）归结

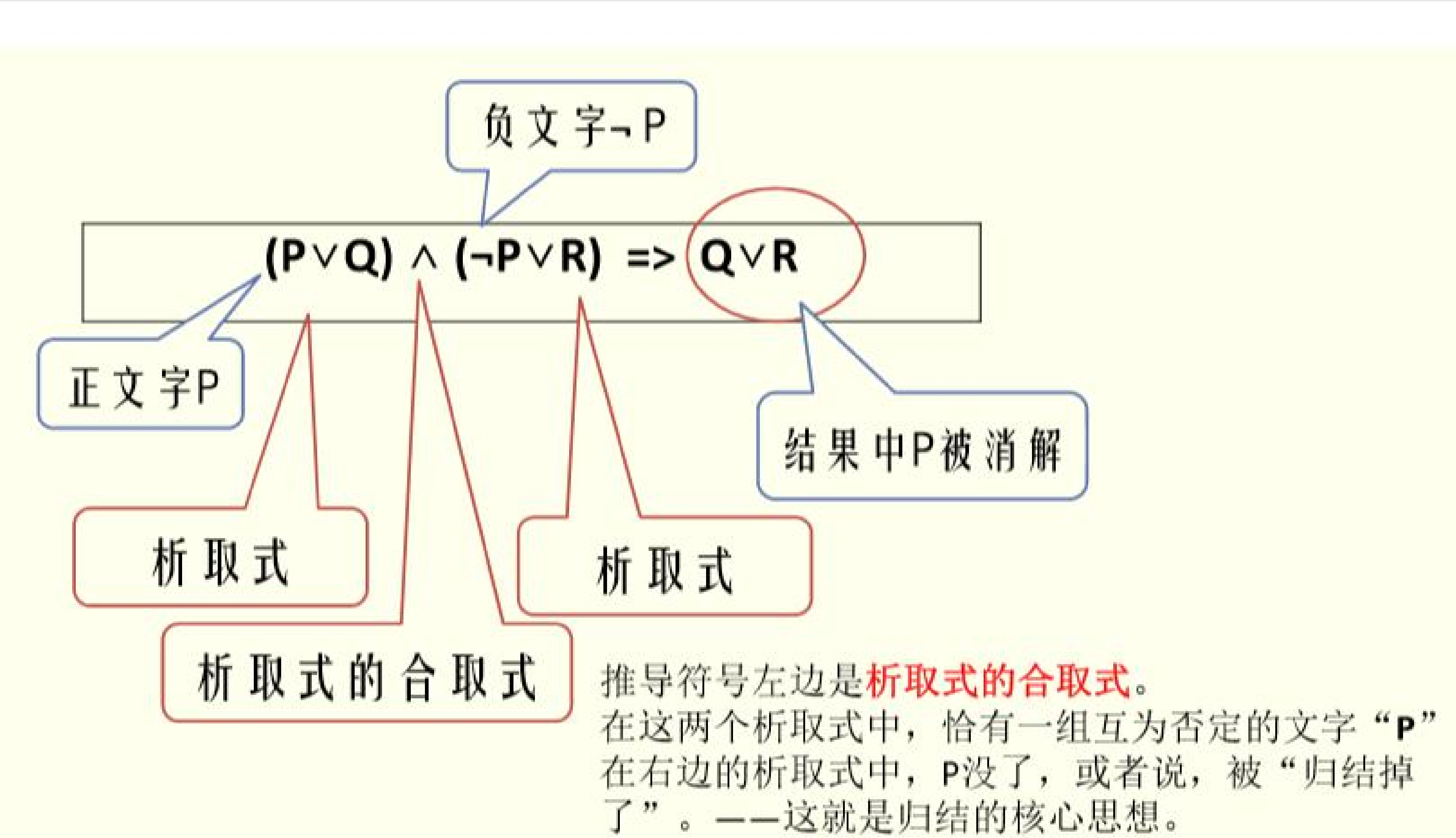
（7） $ANSWER(Wang)$ （2）与（6）归结

应用归结原理求解问题


$T(Wang, Li)$
$\neg C(x, y) \vee \neg T(z, x) \vee T(z, y)$
(1)
(3)
$\neg C(Li, y) \vee T(Wang, y)$
$\neg T(u, Zhang) \vee ANSWER(u)$
(5)
(4)
$\neg C(Li, Zhang) \vee ANSWER(Wang)$
$C(Li, Zhang)$
(6)
(2)
$ANSWER(Wang)$
(7)

重点

归结原理



负文字¬P
(P∨Q) ∧ (¬P∨R) => Q∨R
正文字P
结果中P被消解
析取式
析取式
析取式的合取式
推导符号左边是析取式的合取式。
在这两个析取式中，恰有一组互为否定的文字“P”。
在右边的析取式中，P没了，或者说，被“归结掉了”。——这就是归结的核心思想。

重点

- 归结原理用于自动推理
- 反证法
- 其前提假设：

– （1）证明一个命题 P 成立，等价于证明其逆命题 ¬P 不成立。

– （2）一个已知为真的事实集，其内部不包含矛盾。

– （3）如果将一个命题加入事实集，并且导致事实集出现矛盾，则原命题不成立。

– （4）最常见的矛盾形式，是 ：$P \land \neg P \Rightarrow \square$ 一旦事实集中出现空符号□，则认为事实集中存在矛盾。

□ 归结原理用于自动推理

- 有了上面几条假设，我们可以设想，从目标的待证命题出发：
  - 将待证命题 P 取逆命题，
  - 将逆命题 ¬P 作为“事实”加入已知为真的事实集中，
  - 如果我们在这样的事实集中推导出 □，则该事实集存在矛盾，
  - 则证明逆命题 ¬P不成立，命题P成立
- 如何推导出□呢？这里利用了归结原理。
  - 归结原理是在一个事实集中导出矛盾的最快方式，每归结一步，事实集都在简化而不是变复杂。这就有利于□的推导。
  - 另外，归结原理的目标单一，只要在事实集中推导出□，问题就解决，而与具体证明问题无关，因此是一个通用方法。