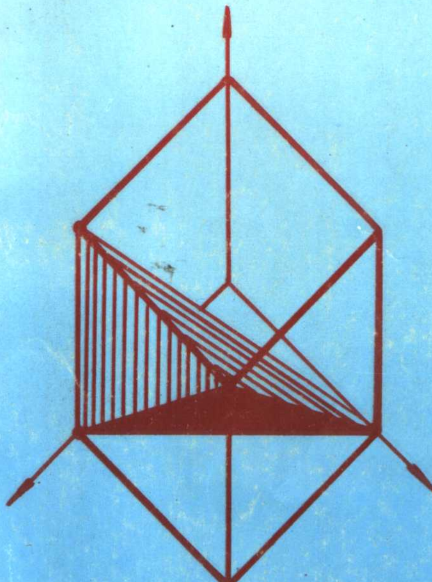
自动推理导论

邱玉辉　张为群　编著



电子科技大学出版社

自动推理导论

邱玉辉　张为群　编著

电子科技大学出版社
·1992·

内 容 简 介

自动推理是八十年代逐步发展形成的一门新兴学科，也是人工智能研究的根本问题。它对专家系统、智能机器人、机器学习和程序正确性证明等领域都有深远的影响。本书主要介绍自动定理证明、非单调推理、概率推理和模糊推理，以及自动推理系统。

本书可作为大学计算机专业高年级学生、研究生的教材或参考书，亦可供计算机科学工作者参考。

自动推理导论

邱玉辉　张为群　编著

*

电子科技大学出版社出版

（中国成都建设北路二段四号）

电子科技大学出版社激光照排中心照排

四川省自然资源研究所印刷厂胶印

*

开本 850×1168 1/32　印张 7　字数 170 千字

版次 1992 年 6 月第一版　　印次 1992 年 6 月第一次印刷

印数 1～1500 册

中国标准书号　ISBN 7-81016-330-2/TP·26

［川］016（15452·152）　定价：4.00 元

序　言

自动推理是八十年代逐步发展形成的一门新兴学科，它是自动定理证明的发展，是人工智能的一个重要研究方向。它对专家系统、智能机器人、机器学习和程序正确性证明等许多领域，产生了深远的影响。

人的很多智能活动都是和推理能力相联系的。按照传统逻辑的观点，推理是从已知事实推出新事实的过程。在这种推理下，新事实仅是那些所提供的确定的事实逻辑地推出的结论。然而，现实世界中事物之间的关系往往是不确定的，由于观察的限制，使得证据本身呈现出不精确性，从而使推理过程表现出不确定性。因此，人类思维涉及非传统的逻辑推理，即常识推理。基于传统逻辑的精确推理，已进行了大量的、卓有成效的研究，而基于非传统逻辑的不精确推理的研究还很不充分，有的理论上还有争议，但所取得的令人注目的成就已广泛用于各类专家系统的开发，以及一般知识工程和推理机器的研究。甚至可以这样说，常识推理研究的突破性进展也许是人工智能的飞跃。因此，在计算机科学系的本科生、研究生以及计算机科学技术人员中加强自动推理知识的普及与应用，无疑是十分重要的。

目前，已有不少数理逻辑的专门教材，但其内容主要是介绍传统逻辑，且偏深偏难，不太适合于普及，同时较少地涉及非传统逻辑系统，后者正好在计算机科学，特别是在人工智能领域中有广泛的应用，本书正是为满足这种需要而编写的。

全书着眼于自动推理及其在人工智能中的应用。

第一章为引言，概括地介绍基于传统逻辑的自动定理证明和基于非传统逻辑的非单调推理、概率推理和模糊推理等概念。

第二、三章分别简明地介绍所谓两个演算，即命题演算和一阶

谓词演算，这是为学习全书提供预备性知识。

第四、五、六章系统地介绍自动定理证明，着重介绍归结原理及其改进方法，它是自动推理的重要内容。对于非归结方法，限于篇幅，本书未作介绍，读者可参考有关文献。

第七章主要介绍非单调推理，包括缺席推理、模态非单调推理、限定推理和 CWA 推理，它们代表了非单调推理的轮廓。

第八、九章介绍概率推理，包括基于 Nilsson 的概率逻辑、Bayes 理论和 Dempster-Shafer 的证据理论下的概率推理。

第十章介绍模糊推理，包括基于 Zadeh 的模糊逻辑和可能性理论下的模糊推理。

最后一章，概略地介绍了自动推理系统。

本书是在为我系人工智能与认知方向研究生授课的讲稿基础上，修改、充实而成，可作为计算机科学系高年级学生选修课和研究生课教材，也可供有关专业人员参考。

本书第一章至第十章由邱玉辉编写，第十一章由张为群编写。在编写本书的过程中，南开大学陈有琪教授，重庆大学的童頫教授，对本教材的编写大纲和书稿都提出了很好的意见，谨向他们表示衷心的感谢。

由于作者水平有限，书中缺点错误在所难免，希望读者批评指正。

编　者

1990 年 12 月 4 日于重庆

目　录

第一章 引 言

自动推理是八十年代逐步发展形成的一门新兴学科，它是自动定理证明的发展，是人工智能的一个重要研究方向。

人的很多智能活动都是和推理能力相联系的。人工智能的许多研究领域，例如专家系统，智能机器人、机器学习，程序正确性证明等等，都是围绕自动推理展开的。因此，自动推理的研究是人工智能的根本问题，它的发展将对人工智能的许多领域产生深远的影响。

1.1 什么是自动推理

什么是自动推理？要理解它，首先要了解什么是推理。按照传统逻辑的观点，推理是从已知事实推出新事实的过程。在这种推理下，新事实仅是那些所提供的确定的事实逻辑地推出的结论。然而，现实世界中事物之间的关系往往是不确定的，由于观察手段的限制，使得证据本身也常常是不完全的，模糊的，随机的，从而使整个推理过程呈现出不精确性，因此人类思维涉及非传统的逻辑推理，即常识推理，自动推理就是由计算机实现或人机交互式地实现的推理过程。其研究目的在于探索人类思维活动的各种推理形式的基本规律，研制出能辅助人类进行问题求解的智能系统。自动推理的研究内容主要有以下四个方面。

1. 自动定理证明

自动定理证明，又称为自动演绎，它是把人证明定理的过程变成一系列能在计算机上自动实现的符号演算过程，也就是演绎推理过程机械化。

推理机械化最早是在 17 世纪由 Leibniz 提出来的。19 世纪的 D. Hilbert 创立了数理逻辑，使这一设想有明确的数学形式。本世

纪二、三十年代，不少人在这方面进行了理论探讨，但结果多是否定的。1930 年，J. Herbrand 奠定了定理证明机械化的理论基础，但他的方法不可能在计算机上实现。在 1956 年夏的 Dartmouth 大学的人工智能会议上，A. Newell，J. Shaw 和 H. A. Simon 共同发表了名为“逻辑理论机”(LT)的程序，证明了《数学原理》第二章中的 38 个定理，而且更重要的是引入了若干基本的自动定理证明概念和人工智能技术。此后，P. Gilmore 和王浩各自独立地提出了机器证明谓词演算的定理。他们都没有采用 A. Newell 和 H. A. Simon 的启发式方法，而是使用从传统逻辑证明过程导出的方法。1965 年，J. A. Robinson 根据 Herbrand 定理，提出一阶谓词演算中的归结原理(resolution principle)，使推理机械化达到了实用的程度，从而使定理证明有了迅速的发展，出现了许多改善的归结方法。它们是语义归结(Slagle, 1967; Meltzer, 1966; Robinson, 1965; Kowalski & Hayes, 1969)，锁归结(Boyer, 1971)，线性归结(Loveland, 1970; Luckham, 1970)，SL 归结(Kowalski; Kuehner, 1970)和支持集策略(Wos 等，1965)等。

归结方法已较 Herbrand 定理有较大的进步，但是，在证明过程中产生大量的归结式仍不是今天的计算技术所能实现的。1975 年，Bledsoe 提出自然演绎法，它是保留联词⊃下的人机交互推理的非归结方法。这种方法虽然不是完备的，但效率较高。例如，微积分学中连续函数的和仍连续的定理，用自然演绎法仅 27 步便得到证明，但采用归结方法，归结出 10 万个子句仍没有结果。1976 年，Andrews 提出 Mating 法，它不必将定理前束化和子句化。1979 年，Boyer-Moore 提出了定理证明器，它是把公理和定理作为函数进行处理，而不是采用谓词演算的方法。它已证明大约 400 个定理。

推理机械化的另一个领域是几何定理证明。1950 年，A. Tarski 证明了初等几何的定理是可机械化的，但步骤繁，不实用。1959 年，H. Gelernter 提出了几何定理证明机(GTM)，它在人工智能领

域中具有里程碑的意义，可证明很多中学试题（直线图）。我国数学家吴文俊教授在 GTM 方面有很大的突破，1977 年，他提出了初等几何的机械化方法，1978 年，又把这个方法推广到微分几何。

2. 非单调推理

演绎推理是人类思维的重要形式，它的特点是，在现有知识的基础上，通过严密的逻辑推理所获得的新知识与已有知识是相容的，但是它是以给定完全信息，且信息是处于静止状态，不会发生变更为前提，这样的推理称为单调推理（monotonic reasoning）。

事实上，人们掌握的信息往往是不完全的，且处于不断变化中，因而人们对世界上各种事物的认识和确信，总是处于不断调整之中。今天获得的新知识，到了明天可能就不再有效，原有的认识就要修改，甚至废弃。也就是说，由于增加新知识后，有时要否定以前的推理结果，这样的推理称为非单调推理（nonmonotonic reasoning）。它的特点是，一个理论 $\mathscr{T}$ 的公理增加时，$\mathscr{T}$ 中的定理也可能增加，也可能减少。

非单调推理是自动推理中非常活跃的领域，提出了许多推理方法。1978 年，R. Reiter 提出封闭世界假设（CWA），K. Clark 讨论了谓词完备化，1979 年，J. Doyle 建立了非单调推理系统 TMS。1980 年，J. McCarthy 提出限定（circumscription）。1982 年，D. McDermott 和 J. Doyle 定义了一种逻辑模型。1986 年，J. McCarthy研究了一类非单调推理的应用。1987 年，R. Reiter 对非单调推理给出了很好的综述，以后又出现了许多改进。

非单调推理是人们在日常生活中常遇到的推理现象。人类在认识世界，改造世界的过程中所拥有的知识，随着实践-认识-再实践-再认识而不断深化和增长，导致这一现象的根本原因就是人们推理时所根据的知识具有不完全性。非单调推理是处理不完全知识的有力工具。人工智能的一切有趣的应用几乎都涉及非单调推理形式，非单调推理在逻辑程序设计、设备诊断和行为推断等方面得到了有效的应用。

3. 概率推理

人们所掌握的知识不仅常常呈现出不完全性，而且表现出随机性，概率论则是处理具有偶然因素时能得出结论的有力工具。在人工智能研究中，概率是最早地广泛用于不确定性表示的形式化。

为了把一阶逻辑的机制扩充到概率论的使用，使之适用于不确定的陈述中，我们必须把逻辑中的一个命题的含义同概率论中随机变量的含义联系起来，并使一阶逻辑的普通理论语义同每个命题的值"真"或"假"相关联。为此，修改语义使得每个陈述同二值随机变量的概率分布相联系，这个分布就是命题的一个解释。例如，对于命题 A，我们指定分布$\{(1-p),p\}$。按照这个解释，A 为"真"的概率是 p。当然，如同在传统逻辑中那样，我们不可能随意地给定命题的解释。例如，分布$\{(1-p),p\}$指派 A 为真的概率是 p，意味着$\neg A$ 的概率必须是 $1-p$。

1976 年，Duda 等人在地质勘探专家系统 PROSPECTOR 中，采用 Bayes 理论处理不精确信息的推理。其中，用 A 和$\neg A$ 表示变量，并把推理规则表示成"若 A 则 B(信度 $\lambda,\bar{\lambda}$)"，λ 是似然比，表示本规则的充分性度量，$\bar{\lambda}$ 表示本规则的必要性度量，它们有明确的概率意义，并测定一个断言的概率改变是如何影响其它断言的。PROSPECTOR 的似然推理过程，就是根据证据 A 的概率 $p(A)$，利用规则$(\lambda,\bar{\lambda})$，把结论 B 的先验概率 $p(B)$更新为后验概率 $p(B|A)$ 的过程。

1986 年，Nilsson 发表了题为"概率逻辑"一文，1987 年，Guggenheimer 和 Freedman 在 Nilsson 工作的基础上，提出了"概率逻辑基础"，从而为概率推理奠定了理论基础。

4. 模糊推理

人们所掌握的知识不仅常常呈现出不完全性，而且表现出模糊性。1965 年，Zadeh 首先提出模糊集合的概念，并指出模糊集合更适合于图象分类和信息加工。以后，他创立了模糊逻辑，并系统地提出以字或句子为值的语言变量和一种不精确的近似推理。

1978 年，Zadeh 又建立了可能性理论，为进一步研究模糊语言和模糊推理提供了数学工具。

模糊推理在专家系统中得到了有效的应用。不精确推理最早是在 MYCIN 中引入的，表示启发式规则强度的确信度因子引入近似应用的概念。在 PROSPECTOR 中也采用一些模糊集理论的形式，将特定的矿床知识用推理网络的形式进行存储，模糊集理论主要用在推理网络的逻辑关系中。1982 年，Kohout & Bandler 综合模糊多值逻辑在专家系统中的应用，为设计模糊专家系统提供了新的框架。1976 年，Shafer 提出证据理论，以处理由不知道引起的不确定性。他采用确信函数作为度量，通过对一些事件的概率加以限制来建立确信函数，而不必说明精确的难于获得的概率。1983 年，Prade 综述了不同的近似推理技术，包括 Shafer 的确信理论，Zadeh 的可能性理论，Bayesian 推理以及 MYCIN 和 PROSPECTOR 中可找到的富于经验的建议，而且包含很多扩充的文献内容，为近似推理的新探索提供了新的思路和方法。

1.2 形式理论

形式理论是这样一种数学方法，它用数学公式的形式表示断言，并把演绎作为从公式推出其它公式的过程。

定义 1.1 一个形式理论 $\mathscr{FT}$ 是一个四元组

$$\mathscr{FT} = (\mathscr{S}, \mathscr{W}, \mathscr{A}, \mathscr{R})$$

其中

$\mathscr{S}$ 是一个可数的符号集，称为符号表，其元素称为符号。一个有穷符号串称为 $\mathscr{FT}$ 的表达式。

$\mathscr{W}$ 是 $\mathscr{FT}$ 的表达式集的子集，称为合式公式集，其元素为合式公式（wff）。

$\mathscr{A}$ 是 $\mathscr{W}$ 的一个子集，称为 $\mathscr{FT}$ 的公理集，其元素称为 $\mathscr{FT}$ 的公理。

$\mathscr{R}$ 是合式公式间关系的有限集 $\{R_1, R_2, \cdots, R_n\}$，其元素称为推

理规则。

定义 1.2 设 $\mathscr{R}=\{R_1,R_2,\cdots,R_n\}$，若对于每个 R_i，$R_i(W_1,W_2,\cdots,W_k,W)$ 对某些 $W_1,W_2,\cdots,W_k,W$ 成立，则称 W 是集合 $\{W_1,W_2,\cdots,W_k\}$ 依据 R_i 而得的直接结果。

定义 1.3 设 Γ 是 $\mathscr{F}\mathscr{T}$ 中公式的集合，称 $wff\ V$ 是 Γ 的演绎结果，记为 $\Gamma\vdash_{\mathscr{F}\mathscr{T}}V$，如果存在一个有穷合式公式序列 $W_1,W_2,\cdots,W_n$，使得 $V=W_n$，且对于每个 i，W_i 或者是公理，或者 $W_i\in\Gamma$，或者 W_i 是 $W_j,W_k(j,k<i)$ 应用某些推理规则 R_i 直接推出的结果。这个序列称为 V 的演绎或一个证明，Γ 的元素称为演绎的前提。

当不致造成混乱时，下标 $\mathscr{F}\mathscr{T}$ 可省略，写成 $\Gamma\vdash V$。

如果 Γ 是有穷集 $\{W_1,W_2,\cdots,W_n\}$，我们可用 $W_1,W_2,\cdots,W_n\vdash V$ 代替 $\{W_1,W_2,\cdots,W_n\}\vdash V$。若 Γ 为空集，则称 V 为 $\mathscr{F}\mathscr{T}$ 的一个定理，记作 $\vdash_{\mathscr{F}\mathscr{T}}V$。全体定理的集合称为 $\mathscr{F}\mathscr{T}$ 的理论。

为便于记忆，我们把推理规则表示为

$$\frac{W_1,W_2,\cdots,W_n}{V}$$

其中，$W_1,W_2,\cdots,W_n$ 称为规则的前提或前件，V 称为结论或后件。为使读者对形式理论的特性有初步的了解，我们对其结构特性作简单的介绍。

定义 1.4 形式理论 $\mathscr{F}\mathscr{T}$ 称为是独立的，如果它的每一公理都是独立的，即把任一公理 A 从 $\mathscr{F}\mathscr{T}$ 中删去后，所得理论 $\mathscr{F}\mathscr{T}'$ 不满足 $\vdash_{\mathscr{F}\mathscr{T}'}A$。

定义 1.5 形式理论 $\mathscr{F}\mathscr{T}$ 称为是一致的或相容的，如果不存在 $\mathscr{T}$ 的公式 V，使得 $\vdash_{\mathscr{F}\mathscr{T}}V$ 和 $\vdash_{\mathscr{F}\mathscr{T}}\neg V$ 同时成立。

一致性是形式理论应当具有的性质。

定义 1.6 形式理论 $\mathscr{F}\mathscr{T}$ 称为是完全的，如果对 $\mathscr{F}\mathscr{T}$ 的任一公式 V，或者 $\vdash_{\mathscr{F}\mathscr{T}}V$，或者 $\vdash_{\mathscr{F}\mathscr{T}}\neg V$。

定义 1.7 形式理论 $\mathscr{F}\mathscr{T}$ 称为可判定的，如果存在一个算法，对 $\mathscr{F}\mathscr{T}$ 的任一公式 V，可确定 $\vdash_{\mathscr{F}\mathscr{T}}V$ 是否成立。否则，称

$\mathscr{F}\mathscr{T}$ 是不可判定的。如果上述算法虽不一定存在，却有一个过程，可对该理论作出肯定的判断，但对非定理的公式过程未必终止，因而未必能作出判断，这时称 $\mathscr{F}\mathscr{T}$ 为半可判定的。

下一章里，我们将介绍一个最简单而又是重要的形式理论的例子，即命题演算。在第三章中，我们将介绍另一个重要的形式理论，即一阶谓词演算。这些内容是为读者学习本书其余部分内容而引入的，并未对数理逻辑的内容进行详细的讨论。

第二章 命题演算($\mathscr{PC}$)

数理逻辑又名符号逻辑，它研究推理中前提和结论之间的形式关系。在数理逻辑中，引入了一套形式语言，组成了一个形式理论，从而把数理逻辑的研究归结为对形式理论的研究。命题演算($\mathscr{PC}$)又称为命题逻辑，一阶谓词演算($\mathscr{FPC}$)又称为一阶谓词逻辑，它们是数理逻辑的基础。

数理逻辑是现代数学的重要基础，也是计算机科学的重要基础之一。

2.1 命题和联结词

定义 2.1 凡可确定真假的陈述句称为命题。当命题为真时，称命题有真值 T；当命题为假时，称命题有真值 F。

例 2.1 判断下列语句是否为命题。

(1)你好吗？

(2)雪是白的。

(3)3 是偶数。

(4)请勿吸烟！

(5)如果王英出差了，那么他就不出席大会。

由命题的定义，可知(1)、(4)不是命题，而(2)、(3)、(5)是命题，其中(2)、(5)是真命题，(3)是假命题。

我们用大写英文字母 A,B,C 等表示命题，这些字母叫做命题词。一个命题词可表示真命题，也可表示假命题，因而又称为命题变元或命题变项。我们还规定符号 T 和 F 为命题常元或命题常项。

复合命题的符号化涉及到联结词。下面介绍 $\mathscr{PC}$ 的几种重要的联结词。

定义 2.2 逻辑联结词$\neg$定义为：$\neg A$ 的真值为 T 当且仅当 A 的真值为 F。称$\neg A$ 为 A 的否定，读作“非 A”。

$\neg$的真值表如表 2.1 所示。

表 2.1 $\neg$的真值表

A	$\neg A$
F	T
T	F

定义 2.3 逻辑联结词$\wedge$定义为：$A\wedge B$的真值为 T 当且仅当 A 和 B 的真值均为 T。称 $A\wedge B$ 为 A 和 B 的合取，读作“A 且 B”。

$\wedge$的真值表如表 2.2 所示。

定义 2.4 逻辑联结词$\vee$定义为：$A\vee B$ 的真值为 F 当且仅当 A 和 B 的真值均为 F。称 $A\vee B$ 为 A 和 B 的析取，读作“A 或 B”。

$\vee$的真值表如表 2.3 所示。

定义 2.5 逻辑联结词$\supset$定义为：$A\supset B$ 的真值为 F 当且仅当 A 的真值为 T 且 B 的真值为 F。称 $A\supset B$ 为 A 蕴涵 B，读作“若 A 则 B”。

表 2.2 $\wedge$的真值表

A	B	$A\wedge B$
F	F	F
F	T	F
T	F	F
T	T	T

表 2.3 $\vee$的真值表

A	B	$A\vee B$
F	F	F
F	T	T
T	F	T
T	T	T

$\supset$的真值表如表 2.4 所示。

$A\supset B$ 的真值为 T 的含义是：前件 A 的真值为 T 是后件 B 的真值为 T 的充分条件，当 A 的真值为 T 时 B 的真值必为 T；当 A 的真值为 F 时，B 的真值可为 T、也可为 F。

定义 2.6 逻辑联结词$\equiv$定义为：$A\equiv B$ 的真值为 T 当且仅当 A 和 B 的真值同时为 T 或同时为 F。称 $A\equiv B$ 为 A 等值 B，读作“A 当且仅当 B”。