信念修改的理论与方法

黄智生 （荷兰阿姆斯特丹大学计算机科学系留学生）

摘 要

目前非单调推理、知识与信念逻辑(亦称关于知识的推理)和信念修改理论已成为人工智能理论研究中的三大新的热门课题。近半年来，笔者参加了国际上先后召开的四次重要的人工智能理论学术会议：JELIA'90(人工智能逻辑，荷兰阿姆斯特丹，1990年9月)、CSL'90(德国海德堡，1990年10月)、MEDLAR'91(英国伦敦，1991年3月)和KR'91(美国波士顿，1991年4月)。本文将简要介绍有关信念修改的基本概念、理论和方法，也介绍了这一理论研究的进一步发展梗概。

在人工智能理论研究中，目前非单调推理，知识与信念逻辑(又称关于知识的推理)和信念修改理论已成为三大新的热门课题，而且各有自己对应的专门化国际会议以促进其研究。对于非单调推理，有每两年一度的非单调推理国际会议(Workshop on Nonmonotonic Reasoning)。对于知识和信念逻辑，有关于知识的推理理论问题国际会议(Theoretical Aspects of Reasoning about Knowledge)。对于信念修改理论，有理论变化的逻辑国际会议(Workshop on the Logic of Theory Change)。当然，上述三个领域并非是各自独立的，它们之间具有很强的联系，但各自又有不同的侧重面。本文将简要地介绍有关信念修改的基本概念，理论与方法。

一、知识与信念

在目前的人工智能理论研究中，“知识”和“信念”是极为相近又有一定区别的两个重要概念。对于存贮在数据库和知识库中的信息，有时我们称它们为知识，有时称它们为信念，这取决于不同的情况。一般地说，知识是指那些为多数人所接受，具有较高可信度的又可用于推理的信息式数据；而信念是指那些仅为个人所相信的，为个人推理所用的，相对不稳定的信息式数据。在知识库管理中，当外来的新信息与已有存贮在知识库中的信息相矛盾的时候，我们就需要对原有的知识库进行修改和更新，而对于研究如何进行这种修改和管理的理论，我们现在习惯地称之为信念修改理论(Theory of Belief

---

first principles”，而不看成是异常情况设计”(design by exception)的OO风范。(3)SA记法不利于人机接口模型的构造，而这一模型在GUI(图形用户界面)环境中正变得越来越重要。

2.假若SA和OOA的范围不同，我们如何从正反两方面来划分这些范围，这种划分是否重叠？我认为这是一个文化问题而不是一个技术问题。即使从技术上看OOA比SA应用更广，许多数据处理组织对于这种改变仍有太多的惰性。只有当旧的风范不能解决该组织的问题时，新的风范才会被普遍接受。因此我感到只有在SA不能胜任的环境中，OOA才得到充分的承认，这些环境可能是一些庞大的，可视的工程，也可以是那些从一开始就把着眼点放在重用及图形用户接口的工程项目。

〔刘晓丹译自《SIGPLAN Notices》25(1990).№.10 王振宇校〕

Revision)。对于知识和信念之间的区别，已经有较好的形式化研究工具。用模态逻辑的术语来讲，知识系统是一个S5系统，而信念系统则是一个弱S5系统，又称为KD45系统。[1,2,6]

二、信念修改的基本理论

从理论上讲，知识库中的信息可以用一个逻辑语言系统来表示。我们假定这个逻辑语言系统$\mathscr{L}$由命题逻辑组成。这样，我们也可以相应地定义一个结论操作函数(Consequence Operation Function)，通常用C_n表示。对于任何一个命题集合(或称知识库)A，B，它们都满足下列关系，

(1)**包含性**　$A\subseteq C_n(A)$

即一个知识库的结论集合应包含原来的前题。

(2)**递代性**　$C_n(A)=C_n(C_n(A))$

即对结论集合再进行推论操作还等于原有的结论集合。

(3)**单调性**　若$A\subseteq B$，则$C_n(A)\subseteq C_n(B)$

即在原有的知识库中再加上新的信息不影响原有的结论。

一个命题集合A，若对于结论操作是封闭的（即$A=C_n(A)$），则称为一个封闭理论(Closed Theory)或简称为理论。一个命题集合B，若它能产生一个理论A(即$A=C_n(B)$)，则称之为A的基(Base)。

对于一个知识库进行更新，从理论上讲，主要有下列三种基本操作：

(1)**扩充**(Expansion)：一个新的命题A连同它的逻辑结论被加到原有的知识库K中，这里新的信息与原有的信息没有矛盾。这种操作结果通常被表示为K^+_A。

(2)**修改**(Revision)：一个与原有的知识库K不协调的新命题A被加入，但为了保证所产生的新知识库的协调性，就得除去原有知识库中的某些命题。这种操作的结果通常被表示为K^*_A。

(3)**压缩**(Contraction)：知识库K中的某些命题A被除去，而没有加入任何新的命题。这种操作的结果通常被表示为K^-_A。

扩充操作比较容易被如下定义：

(Def$^+$)　$K^+_A \underline{\underline{\mathrm{def}}} C_n(K\cup\{A\})$

而修改和压缩操作则比较难处理，因为它涉及到删去原有的某些命题，其中可以有许多种不同的选择和处理方式，很难在理论的层次上把它们确定下来。然而，Levi指出，修改和压缩这两个操作是可以相互定义的，即：

(def*)　$K^*_A \underline{\underline{\mathrm{def}}} (K^-_{\neg A})^+_A$

(即加入A来修改K，相当于先压缩去A的否定式，再扩充A)

(def$^-$)　$K^-_A \underline{\underline{\mathrm{def}}} K\cap K^*_{\neg A}$

(即除去A，可以被认为是先加入¬A来修改K，再取与K的交集)

在这里，值得注意的是：压缩不等同于通常逻辑程序中的删除操作。要压缩掉A，不能简单地在逻辑程序中删去某个Horn子句集，因为通常A不是显式地直接表示成一个子句集，而是隐含地表示在整个知识库K之中。此外，为了保证被压缩后知识库的逻辑协调性，在压缩操作中，可能还需要除去原有知识库K中另外的某些命题。(def*)通常被称为Levi等式(Levi Identity)，而(Def$^-$)称为Harper等式。

由于修改和压缩可以相互定义，我们只要研究其中的一种操作就够了。Alchourron，Gärdenfors，和Makinson提出了关于压缩操作的一系列假设，被称为AGM假设(AGM Postulates)。它们是：

(a)K^-_A是一个封闭理论(封闭性)

(即我们应该定义压缩操作的结果是一个完整的结论集合)。

(b)$K^-_A\subseteq K$(包含性)

(被压缩后的信息不比原来的多)。

(c)如果$A\notin K$，则$K^-_A=K$(无关性，Vacuity)

(若要压缩掉与原有知识库无关的命题，其结果与原有的一样)。

(d)如果$A \notin C_n(\varnothing)$，则$A \notin K^-_A$(成功性)

(如果A不是一个绝对的逻辑真理的话，压缩后必不在某结果中)。

(e)如果$C_n(A)=C_n(B)$，则$K^-_A=K^-_B$(一致性)

(两个结论集合一样的命题集，其压缩后的结果也是一样的)。

(f)$K \subseteq C_n(K^-_A \cup \{A\})$ (恢复性)

(先压缩掉A，则扩展A，其结果仍包含原有的知识库K)。

(g)$K^-_A \cap K^-_B \subseteq K^-_{A \wedge B}$(合并性)

(压缩掉A与压缩掉B之交集包含于压缩掉A与B之逻辑和的结果之中)。

(h)如果$A \notin K^-_{(A \wedge B)}$，则$K^-_{(A \wedge B)} \subseteq K^-_A$

(若A可以与其它一些B一起压缩掉的话，则压掉A∧B包含在压缩掉A之中)。

在上述假设中，假设(a)～(f)是直觉可接受的。包含性假设规定了当一个命题被除去后，原来不知道的任何信息都不会被吸收进来，它实际上规定了压缩操作的上限。恢复性假设说的是被压缩后的知识库应包含足够的信息以待恢复，它实际上规定了压缩操作的下限。一致性假设表示仅仅是表达上的差异不应影响操作的最后结果。假设(f)和(g)不太直观，有时它们仅被称为补充假设。合并性假设表示压缩掉一个逻辑与公式应比它们各自分开压缩去掉的信息更少。而假设(h)则是假设(g)的另外一种相应的情况。

我们还可以相应地定义关于修改操作的基本假设：

(a′)K^*_A是一个封闭理论(封闭性)

(b′)$K^*_A \subseteq K^+_A$(包含性)

(c′)如果$\neg A \notin K$，则$K^+_A \subseteq K^*_A$(无关性)

(d′)$A \in K^*_A$(成功性)

(e′)如果$\perp \in K^*_A$，则$\neg A \in C_n(\varnothing)$ (协调性)

(f′)如果$C_n(A)=C_n(B)$，则$K^*_A=K^*_B$(一致性)

(g′)$K^*_{A \wedge B} \subseteq (K^*_A)^+_B$(交含性，Conjunctive Inclusion)

(h′)如果$\neg B \notin K^*_A$，则$(K^*_A)^+_B \subseteq (K^*_{(A \wedge B)})$(交空性，Conjunctive Vacuity)

上述两组假设具有很强的相互联系。以下是Gärdenfors/Makinson关于修改操作与压缩操作之间的基本假设集的定理：

定理一 (Gärdenfors/Makinson定理)

如果修改操作“*”通过Levi等式来定义，而且其对应的压缩操作满足基本假设(a)～(e)以及(g)～(h)的话，则其修改操作*满足基本假设(a′)～(h′)。

如果压缩操作“－”通过Harper等式来定义，而且其对应的修改操作满足基本假设(a′)～(h′)，则其压缩操作满足基本假设(a)～(h)。

在上述定理中，值得注意的是，为了验证修改操作基本假设，不必使用压缩操作的恢复性假设。上述两组基本假设仅仅提供有关操作的一般性规律，尚未提供任何具体的处理方法。但是，信念修改的总原则应该是，修改应是最小化的，即它不应该去掉不应该去掉的信息。从理论上讲，一个信念修改模型，又称一个BRM，是一个二元组(K, *)，这里K是一个知识库集。(但根据知识与信念之间的约定的区别，K更确切地应该被称为一个信念集(belief sets))，而*：$K \times Sent(\mathscr{L}) \to K$是一个从信念集和句子集到信念集的修改函数*(K, A)。我们习惯于把*(K, A)记为K^*_A。

三、信念修改的方法

具体的信念修改方法可以通过不同的方式来定义。而且，由于修改操作可以通过压缩操作来定义，人们习惯于先研究有关压缩操作的方法，然后相应地推出有关修改的方法。目前，从理论上讲，主要有三种压缩方法：完全满足压缩(full meet contraction)，最大选择压缩(maxichoice contraction)和部分满足压缩(partial meet contraction)。

1.完全满足压缩

要构造一个具体的压缩函数，人们很自然地首先去考虑去掉一个命题后所产生的所有可能的结果集合。如果没有其它标准来确定这些可能的结果集合中哪些是更合适的结果的话，我们就简单地选择所有这些可能的结果集合的交集作为压缩的最后结果。这样，我们有下列定义：

我们用$K\downarrow A$来表示信念集K中没有蕴含命题A的最大子集族。即$K\downarrow A \underline{\underline{def}} \{K'\subseteq K \mid A\notin C_n(K')$而且$\forall K''$：$K'\subset K''\subseteq K$则$A\subseteq C_n(K'')\}$

这样，我们可以定义如下一个完全满足压缩操作：

$$K^-_A \underline{\underline{def}} \begin{cases} \cap(K\downarrow A) & \text{如果} A\notin C_n(\varnothing) \\ K & \text{否则} \end{cases} \quad (3.1)$$

上式表明，如果命题A是一个绝对的逻辑真理(如：$\varphi\rightarrow\varphi$)，则不作任何处理，否则取其最大子集族$K\downarrow A$的交集。

定理二(完全满足压缩定理)　如果压缩操作是一个完全满足压缩操作的话，则对所有的命题$A\in K$而且$A\notin C_n(\varnothing)$，有：

$K^-_A = K\cap C_n(\{\neg A\})$。

上述定理表明，完全满足压缩的结果只剩下信念集K中$\neg A$的逻辑结论。这不是一个很理想的结果。

2.最大选择压缩

基于完全满足压缩所存在的问题，Alchourron和Makinson于1982年提出了最大选择压缩操作。其基本思想是，我们可以定义一个选择函数r，它能够从$K\downarrow A$的所有元素中选择一个元素作为压缩操作的结果。即：

$$K^-_A \underline{\underline{def}} \begin{cases} r(K\downarrow A) & \text{如果} A\notin C_n(\varnothing) \\ K & \text{否则} \end{cases} \quad (3.2)$$

Gärdenfors于1988年指出最大选择压缩满足压缩操作假设(a)～(f)，但不一定满足假设(g)和(h)。但如果选择函数r是有序化的话，则其最大选择压缩操作满足基本假设(g)和(h)。所谓选择函数r是有序化的，是指存在一个K中所有子集上的偏序关系⊑使得选择函数r总能选择出一个相对于这个偏序关系的最大元素来。

定理三(最大选择压缩定理)　如果压缩操作是一个最大选择压缩操作，则对所有的命题$A\in K$，和任何命题B，有：

$(A\vee B)\in K^-_A$或者$(A\vee\neg B)\in K^-_A$

上述定理表明，最大选择压缩会导致被压缩掉的命题A与其它任何命题B或其否定式$\neg B$的逻辑或进入其结果。完全满足压缩的结果集太小，而最大选择压缩的结果集又太大。这是其不足之处。

3.部分满足压缩

部分满足压缩是完全满足压缩和最大选择压缩的一个折衷方案，即我们可以定义一个选择函数r，它能够从$K\downarrow A$的元素集中选择出一组元素。然后，我们就可以取这些被选元素的交集作为部分满足的压缩结果。

$$K^-_A \underline{\underline{def}} \begin{cases} \cap r(K\downarrow A) & \text{如果} A\notin C_n(\varnothing) \\ K & \text{否则} \end{cases} \quad (3.3)$$

部分满足压缩同样满足压缩基本假设(a)～(f)。但为了使其满足补充假设(g)和(h)，我们还需要引入一个在信念集K的所有子集族2^K上的一个关系≤，但这个关系≤不一定是一个偏序关系。我们说一个压缩操作是关系的(relational)，是指存在在2^K上的一个二元关系⊑，使得被选子集相对于⊑是最大子集，即，它满足下式：

$$r(K\downarrow A) = \{K'\in K\downarrow A \mid \forall(K''\in K\downarrow A)\ K''\sqsubseteq K'\}$$

定理四(部分压缩定理，Gärdenfors，1988)　如果2^K上的二元关系⊑是可传递的，则其相应的部分满足压缩操作满足AGM假设(a)～(h)。

四、信念修改的理论研究的进一步发展

上面我们仅仅介绍了关于信念修改理论研究的最基本的一些概念、假设和定理。近几年，信念修改已成为人工智能理论研究的热门课题之一，已经取得许多新的进展。限于篇幅，在这里我们不能逐一详细介绍，但是，其主要进展有下列几项：

1. Gärdenfors的认识牢靠性理论（1988年）

AGM假设仅提供信念修改的一般性法则，Gärdenfors提出了认识牢靠性（epistemic entrenchment）的概念，可用于描述更为具体的操作法则。所谓的认识牢靠性是引入了命题集上的一个完全的偏序关系≤。对于两个命题A和B，A≤B表明B至少在认识上比A更牢靠。换言之，如果命题B被修改掉，则比B更不牢靠的A必先被除去。根据认识牢靠性的概念，Gärdenfors提出了一系列公理系统和假设。

2. Nebel的信念基修改理论（1989年）

一般的信念修改理论不能直接用于人工智能的实际应用，因为它要求修改结果是一个演绎封闭的无限的命题集合。为了使信念修改理论更实用化，Nebel对AGM理论进行修改使之能针对信念基的有限集情况。他定义了一个信念基压缩操作⊖，即，

$$K \ominus A \overset{\text{def}}{=} \begin{cases} (K \vee \neg A) \wedge V_{c \in (K \downarrow A)} C & \text{如果} A \notin C_n(\varnothing) \\ K & \text{否则} \end{cases}$$

Nebel证明了信念基压缩同样满足相应的AGM理论。

3. Doyle的理性信念修改理论（1990年，1991年）

Jon Doyle认为在AGM理论和Nebel的理论中，对进行修改的可用信息有过多过强的不必要标准。他提出了理性修改(rational revision)的概念，这里所谓的理性取的是决策论上常用概念，即能够选择出最偏爱的结果。理性修改可以在各种可能的修改中选择最偏爱的修改结果。Doyle相应地定义了理性压缩和理性修改等操作，并提出了一系列的定理和结果。理性信念修改克服了以前一些理论的不足之处，同时也为信念修改理论的实用化提供了较好的应用前景。

参考文献

〔1〕黄智生，关于知识的推理，《计算机科学》，1991年第1期

〔2〕黄智生，基于可能世界语义的知识逻辑及其问题，《哲学研究》，1990年第5期

〔3〕Peter Gärdenfors, Knowledge in Flux: Modeling the Dynamics of Epistemic States. Bradford Books, MIT Press, Cambridge, MA, 1988.

〔4〕Peter Gärdenfors, Makinson, D., Revisions of knowledge systems using epistemic entrenchment, Procedings of the second conference on Theoretical Aspects of Reasoning about Knowledge, M. Vardi(ed.), Morgan Kaufmann Pulb., Los Altos, CA, 1988.

〔5〕Bernhard Nebel, Reasoning and Revision in Hybrid Representation Systems, Springer Verlag, 1990.

〔6〕Zhisheng Huang, Karen Kwast, Awareness, Negation, and Logical Omniscience, Logics in Artificial Intelligence, Springer Verlag, 1991.

〔7〕A Fuhrmann, M. Morreau(Eds.), The Logic of Theory Change, Springer Verlag, 1991.

下期主要内容预告

软件技术漫谈

Fortran新标准的主要特征及面向过程语言的生命力问题；

类型理论和程序设计；面向对象的可视知识研究；

并发模型的比较研究；伯克利大学的数据库研究；

编译型PROLOG数据库的结构模型及其分析；

一个以Ada为核心的面向对象的开发系统。