论不完全性定理：从数学基础的视角

程勇（武汉大学哲学学院）

摘要：发表于1931年的哥德尔不完全性定理是数学基础领域最重要深刻的成果之一，它对逻辑、哲学、数学、理论计算机科学等领域的发展都产生了重要深刻的影响。基于文献中的最新研究进展，本文从如下几个主要方面系统阐释哥德尔之后学界对不完全性定理的研究成果：不完全性定理的解读、不完全性定理在不同领域中的影响，不完全性定理的不同证明、不完全性定理的推广形式、不完全性定理成立的限度。[1]

关键词：不完全性定理、不完全性的限度、内涵性问题、解释、具体不完全性、实质不完全

1 导论

哥德尔不完全性定理是20世纪数学基础领域最重要深刻的成果之一，是现代逻辑发展史上的里程牌，对逻辑学、哲学、数学、理论计算机科学等领域的发展产生了广泛、深刻而持久的影响，极大地改变了1931年之后数学基础研究的面貌。关于不完全性定理的经典书籍，参见[18, 42, 46, 57, 59]。关于不完全性定理的研究报告，参见[4, 13, 41, 58]。

下面我们先解释本文用到的一些基本概念。一个形式理论包括如下基本构件：形式语言、公理、推演规则。若未作特殊说明，本文中的理论均指形式理论。形式语言包括符号表和公式的形成规则。符号表包括逻辑符号和非逻辑符号。有限个符号构成符号串。公式是一类根据形成规则递归定义的"有意义"的符号串。公理是一类特殊的公式（包括逻辑公理和非逻辑公理），是理论中演绎推理的出发点。推演规则告诉我们如何由给定的一些公式推出其他公式。给定理论T及其公式A，我们称"A在理论T中可证"，记为$T \vdash A$，若存在一有限长度的公式序列使得此序列的最后一个公式是A，

[1]基金项目：国家社科基金一般项目"不完全性的限度及其内涵性问题研究"（项目编号：18BZX131）阶段性成果。本文写于2021年，谨以本文纪念哥德尔不完全性定理发表90周年。

且对此序列中的任意公式B，或者B是理论T的公理，或者B是由此序列中公式B前面的公式使用某种推演规则而得到的。罗宾逊(Robinson)算术$\mathbf{Q}$ 和皮亚诺(Peano)算术$\mathbf{PA}$是不完全性研究中的两个重要理论。

罗宾逊算术$\mathbf{Q}$是由塔斯基、莫斯托夫斯基(Mostowski) 和罗宾逊在[63]中引入的。

定义 1 *罗宾逊算术$\mathbf{Q}$的非逻辑符号包含一个常元符号$\mathbf{0}$，一个一元函数符号$\mathbf{S}$，两个二元函数符号$+$ 和$\times$。理论$\mathbf{Q}$的非逻辑公理由如下语句组成：*

$\mathbf{Q}_1$ $\forall x\forall y(\mathbf{S}x = \mathbf{S}y \to x = y)$;

$\mathbf{Q}_2$ $\forall x(\mathbf{S}x \neq \mathbf{0})$;

$\mathbf{Q}_3$ $\forall x(x \neq \mathbf{0} \to \exists y(x = \mathbf{S}y))$;

$\mathbf{Q}_4$ $\forall x\forall y(x + \mathbf{0} = x)$;

$\mathbf{Q}_5$ $\forall x\forall y(x + \mathbf{S}y = \mathbf{S}(x + y))$;

$\mathbf{Q}_6$ $\forall x(x \times \mathbf{0} = \mathbf{0})$;

$\mathbf{Q}_7$ $\forall x\forall y(x \times \mathbf{S}y = x \times y + x)$.

皮亚诺算术$\mathbf{PA}$和罗宾逊算术$\mathbf{Q}$具有相同的语言。我们称$\mathbf{PA}$的语言为算术语言，称算术语言下的理论为算术理论。皮亚诺算术$\mathbf{PA}$的非逻辑公理由定义1中的语句$\mathbf{Q}_1$-$\mathbf{Q}_2$，$\mathbf{Q}_4$-$\mathbf{Q}_7$, 及如下公理模式（归纳原则）组成：$(\phi(\mathbf{0}) \wedge \forall x(\phi(x) \to \phi(\mathbf{S}x))) \to \forall x\, \phi(x)$, 其中$\phi$ 是含有至少一个自由变元x 的算术公式。

每个自然数都对应于算术语言中的一个项。例如，自然数0对应于算术语言中的常元符号$\mathbf{0}$，自然数1对应于算术语言中的项$\mathbf{S0}$，依次类推。我们用$\bar{n}$ 表示自然数n 在算术语言中所对应的项。在算术的标准模型中，论域是自然数集，常元符号$\mathbf{0}$被解释为最小的自然数0，函数符号$\mathbf{S}$被解释为自然数上的后续函数$S(n) = n + 1$, 函数符号$+$和$\times$ 分别被解释为自然数上的加法和乘法运算。

算术化是数理逻辑中的一种重要方法，它的思想是给理论的符号表编码。每个符号有个自然数编码，从而有限长度的符号串也可用单个自然数编码。基于编码，符号串上的关系和运算就可转换为自然数上的关系和运算。关于算术化方法的技术细节，参见[46]。通过算术化，任何公式或公式

序列都可被一个自然数编码（称为哥德尔编码）。给定算术语言中的公式ϕ，我们用$\sharp(\phi)$ 表示ϕ的哥德尔编码，用$\ulcorner\phi\urcorner$表示ϕ的哥德尔编码在算术语言中所对应的项。本文中，通过算术化，我们常将公式集等同于其对应的哥德尔编码集。

直观地说，我们称一公式集Σ 是递归集，若存在一个能行的算法使得任给Σ语言中的公式A，此算法可判定A是否是Σ 中的元素。我们称理论T是递归公理化的，若T的公理集是递归集。我们称理论T是一致的，若不存在公式A使得A 和$\neg A$都在T中可证。我们称理论T是完全的，若对理论T语言中的任意公式A，或者A在T中可证，或者$\neg A$ 在T中可证。我们称公式A是理论T的不可判定命题（或A是独立于理论T的），若A 和$\neg A$都在T中不可证。故理论T是不完全的当且仅当存在T中的不可判定命题。我们称理论T是实质不完全的，若T的相同语言中的递归公理化的一致扩充都是不完全的。

我们定义算术语言下公式的算术层级。有界公式（Σ^0_0 或Π^0_0 或Δ^0_0公式）是由原子公式仅使用命题联结词和有界量词（形如$\forall x \leq y$ 或$\exists x \leq y$）构造而成。Σ^0_{n+1} 公式是等价于形如$\exists x\phi$ 的公式，其中ϕ 是Π^0_n公式。Π^0_{n+1}公式是等价于形如$\forall x\phi$ 的公式，其中ϕ 是Σ^0_n公式。一公式是Δ^0_n公式，若其既是Σ^0_n公式又是Π^0_n公式。我们称理论T是ω-一致的，若不存在理论T语言中的公式$\varphi(x)$使得$T \vdash \exists x\varphi(x)$，且对任意的自然数$n$, $T \vdash \neg\varphi(\bar{n})$；称理论$T$是1-一致的，若不存在这样的$\Delta^0_1$ 公式$\varphi(x)$。

在哥德尔时代，罗素-怀特海《数学原理》中的逻辑系统**PM** (Principia Mathematica)可能是当时最有名的用来形式化经典数学的逻辑系统。哥德尔于1930年8月26日在维也纳向卡尔纳普透露了他的结果，并在1930年9月7日著名的Königsberg 会议上的一次圆桌讨论中宣布了他的结果（第一不完全性定理）（参见[74]）。[2] 哥德尔不完全性定理发表于1931年1月的论文“论**PM**及其相关系统中的形式不可判定命题(On Formally Undecidable Propositions in Principia Mathematica and Related Systems I)”（参见[28]）。之后，哥德尔这一对数学基础研究十分重要的定理开始在学界快速传播（参见[16]）。[3]哥德尔[28]中不完全性定理的证明使用

[2]冯诺依曼参加了此会议，当时他正在作与希尔伯特纲领相关的研究，他立即明白了哥德尔结果的重要性（参见[16]）。1930年11 月20日，冯诺依曼给哥德尔写了一封信告诉哥德尔他发现了一个由哥德尔结果得到的推论：一致性的不可证明性（参见[16]）。然而，哥德尔早于冯诺依曼发现了同样的结论，在他已经提交的关于不完全性定理的论文中已包含第二不完全性定理的陈述。

[3]当时学界对哥德尔不完全性定理的反应也是褒贬不一。逻辑和数学基础领域的一些重要人物很快就接受了哥德尔的成果并理解其意义，但当时也存在相当多的误解和抵制，有关不

的是基于**PM**及戴德金-皮亚诺算术公理的一个简单类型理论，记为**P** (参见[4])。哥德尔1931年发表的第一不完全性定理的原始形式是：对任何在理论**P**基础上增加一递归的公理集而形成的具有与**P**相同语言的理论T，若T是ω-一致的，则T是不完全的（参见[28]中定理VI）。哥德尔指出：任何满足如下两个条件的ω-一致的理论都存在不可判定命题：（1）公理集是递归集；（2）所有递归关系都在此理论中可表示。[4]哥德尔[28]中提出第二不完全性定理：假设理论**P**是一致的，则**P**的一致性在**P**中不可证（参见[28]中定理XI）。[5] 哥德尔并未发表第二不完全性定理的证明细节（关于证明细节可参见希尔伯特-伯奈斯(Hilbert-Bernays)[36], 费弗曼(Feferman)[19] 和布洛斯(Boolos)[6]）。哥德尔在[28]中只是给出第二不完全性定理证明的概要，哥德尔证明的主要思路是在**P**中形式化第一不完全性定理的证明。

哥德尔在1932年的论文[29]中提出了基于皮亚诺算术的扩充形式的不完全性定理。他在1934年的论文[30]中给出1931年论文中结果的另一种形式。哥德尔于1963年在他的1931年论文的英文译文的注解中写道："基于图灵的工作，我们有关于形式理论的一般精确定义，因此我们可有第一和第二不完全性定理的一般形式：我们可严格地证明，在任何包含足够算术理论的递归公理化的一致形式理论中都存在不可判定的算术命题，且此理论的一致性在此理论内不可证明"(参见[26])。

后人的研究极大地改进了不完全性定理的表述，下面是现代版本的哥德尔不完全性定理：

第一不完全性定理 令T 是**Q**的递归公理化的扩充。若T 是ω-一致的，则T是不完全的。

第二不完全性定理 令T 是**Q**的递归公理化的扩充。若T 是一致的，则"T的一致性"在T中不可证。[6]

哥德尔第一不完全性定理的证明假设了基础理论是ω-一致的。罗瑟（Rosser）于1936年改进了哥德尔的结论，罗瑟的证明仅假设了基础理论是一致的：

完全性定理被学界接受情况的详细说明，参见[16, 45]。

[4]以一元关系为例，我们称自然数上的一元关系$R(x)$ 在理论T中可表示，若存在T语言中的公式$\phi(x)$使得，若$R(m)$ 成立，则$T \vdash \phi(\overline{m})$; 若$R(m)$ 不成立，则$T \vdash \neg\phi(\overline{m})$。我们称公式$\phi(x)$是关系$R(x)$的表示公式。

[5]理论**P**的一致性是理论**P**中的一个语句，其表达的意思是存在理论**P**中不可证的语句。

[6]下文中，我们将讨论如何在T中表达T的一致性。

哥德尔-罗瑟第一不完全性定理 若T是$\mathbf{Q}$的递归公理化的一致扩充，则T是不完全的。

下文中，我们用G1表示哥德尔-罗瑟第一不完全性定理及其不同的版本，用G2表示第二不完全性定理及其不同的版本。下文中，我们假定可从上下文中知道G1和G2所代表的具体形式。

基于文献中最新研究进展（包括本人的研究），本文从如下几个主要方面系统阐释哥德尔之后学界关于不完全性定理的研究成果：不完全性定理的解读、不完全性定理在不同领域中的影响，不完全性定理的不同证明、不完全性定理的推广形式、不完全性定理成立的限度。

本文的结构如下：第2节中，我们说明不完全性定理中的几个核心概念，概述不完全性定理证明的基本思想，并对两个不完全性定理作若干说明。第3节中，我们扼要地讨论不完全性定理对数学基础、经典数学、理论计算机科学、哲学四个领域的主要影响。第4节中，我们概括了文献中不完全性定理不同证明的特征，并对这些特征作若干说明。第5节中，我们总结了G1和G2的不同推广形式，并给出若干典型例子。第6节中，我们分别讨论G1和G2成立的限度，并重点讨论G2对一理论成立取决于哪些因素。第7节中，我们总结全文。

2 不完全性定理的解读

在给出哥德尔不完全性定理的证明概要前，我们先对其中的三个核心概念（递归公理化、完全、ω-一致）作若干说明。

第一、一理论是递归公理化的直观含义是：存在一能行的算法使得对理论语言中的任意公式A，此算法可判定A是否此理论的公理。使得G1成立的一个必要条件是理论T是递归公理化的：理论$\mathbf{Q}$的非递归公理化的扩充可能是完全的。在哲学文献中常见的一种不完全性定理的表述是：任何包含初等算术的一致形式理论都是不完全的。严格地说，这种表述是错误的：并非$\mathbf{PA}$的一致扩充都是不完全的。例如，考虑算术的真理论（其由所有在算术的标准模型中为真的语句组成）：算术的真理论是$\mathbf{PA}$的一致扩充，但它是完全的。事实上，算术的真理论不是递归公理化的，因此不满足G1的条件。

第二、这里我们区分语义上的“完备”概念和语法上的“完全”概念。不完全性定理中的“完全”概念表达的是语法上的完全性。哥德尔在1929年的

博士论文中证明了带等词的一阶逻辑的完备性定理：对任意公式集T和公式A，T逻辑蕴涵A当且仅当T可演绎推出A。由一阶逻辑的完备性定理可知，任何一阶理论都是完备的；但并非任何一阶理论都是完全的。对应于语法上的完全性，文献中也常称语义上的完备概念表达的是语义上的完全性。

第三、哥德尔第一不完全性定理假设了“ω-一致”这一条件。而“一致”是比“ω-一致”更弱的条件。ω-一致的理论都是一致的，但一致的理论未必都是ω-一致的。

哥德尔第一不完全性定理证明的主要思想是：算术化、可表示性、自指构造。下文中，若未作特殊说明，我们假设理论T是$\mathbf{Q}$的递归公理化的一致扩充。通过算术化，T语言中的公式及有限公式序列都可用自然数来编码。基于编码，我们可将关于理论T的元数学命题转换成关于自然数的命题。哥德尔在[28]中给出了递归函数的定义，并证明了递归函数的一些基本性质。他在[28]中证明了一个重要的定理：所有的递归关系都在$\mathbf{PA}$中可表示。通过可表示性，关于理论T的系统性质的递归关系是在T中可表示的，从而在T中有对应的表示公式。这样，我们可以在理论内部“谈论”此理论的系统性质，这是算术化思想的要点。

下面，我们给出现代版本哥德尔不完全性定理的证明概要。通过算术化，我们可定义表达理论T内演绎证明关系的自然数上的关系。例如，我们可定义自然数上的如下二元关系：$Proof_T(m,n)$ 成立当且仅当n 是编码为m的公式在T中的一个证明的编码。我们可证明，关系$Proof_T(m,n)$ 是递归的。因递归关系都在$\mathbf{PA}$中可表示，令$\mathbf{Prf}_T(x,y)$为$Proof_T(m,n)$在$\mathbf{PA}$中的表示公式。[7] 由公式$\mathbf{Prf}_T(x,y)$我们可定义可证谓词$\mathbf{Pr}_T(x) \triangleq \exists y \mathbf{Prf}_T(x,y)$。最后，哥德尔构造了语句$\mathbf{G}$（称为哥德尔语句），其断定自身在$T$中不可证，即$T \vdash \mathbf{G} \leftrightarrow \neg \mathbf{Pr}_T(\ulcorner \mathbf{G} \urcorner)$。哥德尔证明了：若$T$是一致的，则$\mathbf{G}$在$T$中不可证；若$T$是$\omega$-一致的，则$\neg \mathbf{G}$ 也在T中不可证。因此，则若T是ω-一致的，则$\mathbf{G}$是T的不可判定命题，故T是不完全的。

我们称公式$\mathbf{Pr}_T(x)$是标准可证谓词若其满足如下条件：

D1 若$T \vdash \phi$, 则$T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner)$;

D2 $T \vdash \mathbf{Pr}_T(\ulcorner \phi \to \varphi \urcorner) \to (\mathbf{Pr}_T(\ulcorner \phi \urcorner) \to \mathbf{Pr}_T(\ulcorner \varphi \urcorner))$;

D3 $T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner) \to \mathbf{Pr}_T(\ulcorner \mathbf{Pr}_T(\ulcorner \phi \urcorner) \urcorner)$.

[7]通过算术化和可表示性，我们得到表示公式$\mathbf{Prf}_T(x,y)$，其在$\mathbf{PA}$中谈论T的系统性质。

哥德尔构造的可证谓词$\mathbf{Pr}_T(x)$满足条件**D1-D3**，因此是标准可证谓词。[8] 由可证谓词$\mathbf{Pr}_T(x)$，我们可定义一致性命题$\mathbf{Con}(T) \triangleq \neg\mathbf{Pr}_T(\ulcorner \mathbf{0} \neq \mathbf{0} \urcorner)$。[9] 我们称由标准可证谓词$\mathbf{Pr}_T(x)$构造的语句$\mathbf{Con}(T)$为典范一致性命题。本文中，若未作特别说明，我们假定$\mathbf{Pr}_T(x)$是标准可证谓词，且用典范一致性命题$\mathbf{Con}(T)$来表达理论T的一致性。

由标准可证谓词$\mathbf{Pr}_T(x)$的性质**D1-D3**，我们可证明：$T \vdash \mathbf{G} \leftrightarrow \mathbf{Con}(T)$。由此，我们可得G2：若$T$是一致的，则$\mathbf{Con}(T)$在$T$中不可证。

下面我们先给出关于第一不完全性定理的若干说明。第一、在不完全性定理的原始表述中，哥德尔避免使用类似模型、真、可定义性等语义概念。这一方面是因为在哥德尔论文之前，人们不太清楚可证概念与真概念间的区别；另一方面是因为，由于逻辑中广为人知的语义悖论的存在，语义概念在那时是受到怀疑的（参见[4]）。而塔斯基关于形式语言的真定义是直到1933年才出版，这可解释为何哥德尔在1931年关于不完全性定理的原始表述中使用像ω-一致这样的语法概念，而没有使用语义概念。在1931年，人们还没有一个理论在另一理论中可解释的精确定义，因此，不完全性定理的原始表述受囿于很复杂的理论**P**及其语言。

第二、哥德尔在第一不完全性定理的证明中假设了T是ω-一致的，这个假设比“T是一致的”更强。在哥德尔的证明中，仅假设T是一致的，我们可证明哥德尔语句$\mathbf{G}$在T中不可证；但仅假设T是一致的不足以证明$\neg\mathbf{G}$也在T中不可证。要证明哥德尔语句是T的不可判定命题需假设更强的条件。特别地，在“T是ω-一致的”的假设下哥德尔证明了哥德尔语句在T中不可判定。

第三、完全的理论去掉若干公理后可能是不完全的。不完全的理论增加若干公理可能变为完全的。然而，不完全性定理表达的是满足一定条件的理论的实质不完全性，即不仅此理论是不完全的，而且此理论的任意递归公理化的一致扩充都是不完全的：即无论我们怎么能行地扩充此理论的公理，一致扩充后的理论总存在不可判定命题。

第四、不完全性定理和一阶逻辑完备性定理也是紧密相关的。由一阶逻辑完备性定理，“理论T可演绎推出公式A”是个很强的概念：它意味着A在T的所有模型中都为真。由此可见，若理论T是完全的，则对T语言中

[8] 罗瑟在证明G1时使用的是Rosser可证谓词，其定义如下：$\mathbf{Pr}_T^R(x) \triangleq \exists y(\mathbf{Prf}_T(x,y) \wedge \forall z \leq y \neg\mathbf{Prf}_T(\dot{\neg}x, z))$, 其中$\mathbf{Prf}_T(x,y)$是证明谓词。

[9] $\mathbf{Con}(T)$是表达T的一致性的算术语句，其表达的意思是：形如$\mathbf{0} \neq \mathbf{0}$的矛盾式在$T$中不可证。

的任意公式A，或者A在T 的所有模型中都为真，或者A在T的所有模型中都为假。但完全有可能存在某公式A使得A在T的某些模型中为真，且在T的某些模型中为假，因此T是不完全的。从这角度看，有些一阶理论是不完全的结论并非是很出乎意料的。

第五、不完全性定理中使用的是相对可证性概念：即相对于某一形式系统而言的可证性概念，而不是绝对可证性概念（在任意的一致系统中都可证）。当我们说某一命题可证或不可证时都是相对于某一形式系统而言的。不完全性定理告诉我们，$\mathbf{Q}$的任意递归公理化的一致扩充T_0都存在不可判定命题。在T_0中不可判定的命题A_0可在T_0的某一致递归扩充T_1中可证，但根据不完全性定理，T_1也是不完全的。在T_1中不可判定的命题A_1可在T_1的某一致递归扩充T_2中可证，但T_2中也存在不可判定命题A_2。这个过程可类似地继续下去。

第六、G1并非告诉我们一致的理论都是不完全的。事实上有很多一致的完全理论。例如无端点的稠密线性序理论（$\mathbf{DLO}$）是完全的。即使是算术理论，也并非所有的算术理论都是不完全的。例如，给定特征p的代数闭域理论($\mathbf{ACF}_p$)，实闭域理论($\mathbf{RCF}$)等算术理论都是完全的。

第七、虽然$\mathbf{PA}$是不完全的（存在在$\mathbf{PA}$中不可证的算术真命题），但在标准模型下为真的Σ^0_1 算术语句都在$\mathbf{PA}$中可证。然而，任何包含足够算术信息（例如$\mathbf{Q}$）的一致递归公理化理论总存在不可证的Π^0_1算术真语句。例如，哥德尔语句是形如$\forall x\phi(x)$ 的Π^0_1 算术真语句，其在$\mathbf{PA}$中不可证，但对任意的自然数n, $\phi(\bar{n})$在$\mathbf{PA}$中可证。然而，图灵在[64]中证明了任意的Π^0_1算术真语句都在$\mathbf{PA}$的某一超限迭代系统中可证。费弗曼在[20] 中推广了图灵的工作，并证明了任意算术真语句都在$\mathbf{PA}$ 的某一超限迭代系统中可证。

下面我们给出关于第二不完全性定理的几点说明。第一、G2告诉我们$\mathbf{Con}(\mathbf{PA})$在$\mathbf{PA}$中不可证，但对$\mathbf{PA}$的任意有限子系统$S$而言，$\mathbf{Con}(S)$在$\mathbf{PA}$中可证。

第二、对G2的一个常见错误解读是$\mathbf{PA}$的一致性只能在$\mathbf{PA}$的扩充理论中可证。根岑（Gentzen）在1936年提出一理论，其是原始递归算术加上长度为ϵ_0的无量词的超穷归纳原则，并在此理论中证明了$\mathbf{PA}$的一致性。[10] 根岑的理论既非$\mathbf{PA}$的子理论也非$\mathbf{PA}$的扩充理论，但在解释度意义上，根岑理论比$\mathbf{PA}$更强（参见[73]）。

第三、G2仅告诉我们，若T是一致的，则$\mathbf{Con}(T)$在T中不可证。G2并

[10] ϵ_0 是最小的序数α 使得$\omega^\alpha = \alpha$.

没有告诉我们$\mathbf{Con}(T)$是独立于T的。仅假设T是一致的不足以推出$\neg\mathbf{Con}(T)$也在T中不可证。要证明$\mathbf{Con}(T)$是独立于T的，我们需假设更强的条件。事实上，假设T是1-一致的，我们可证明$\mathbf{Con}(T)$是独立于T的（条件"T 是1-一致的" 比条件"T是一致的"更强）。令1-$\mathbf{Con}(T)$为在$\mathbf{PA}$语言中表达T是1-一致的语句。不动点引理告诉我们：若T是$\mathbf{Q}$的一致递归扩充，则对任意仅含一个自由变元的公式$\phi(x)$, 存在语句θ 使得$T \vdash \theta \leftrightarrow \phi(\ulcorner\theta\urcorner)$. 我们已知如下结论：

事实 1 *令T 为$\mathbf{PA}$的递归公理化的一致扩充。*

- $T \vdash \mathbf{Con}(T) \to \mathbf{Con}(T + \neg\mathbf{Con}(T))$;
- $T \nvdash \mathbf{Con}(T) \to \mathbf{Con}(T + \mathbf{Con}(T))$;
- $T \vdash \mathbf{Con}(T) \to \mathbf{Con}(T + \theta)$, *其中$\theta$ 是罗瑟语句;*[11]
- $T \vdash 1\text{-}\mathbf{Con}(T) \to \mathbf{Con}(T + \mathbf{Con}(T))$.

关于不完全性定理的应用的一个误解是我们可不断地迭代增加一致性命题：由$\mathbf{Con}(T)$ 可得$\mathbf{Con}(T + \mathbf{Con}(T))$，再可得$\mathbf{Con}(T + \mathbf{Con}(T + \mathbf{Con}(T)))$，由此类推下去。然而，由事实1可知，这并不成立。为了能迭代地增加一致性命题，我们需假设比"T是一致的"更强的条件：T是1-一致的。

如下事实揭示了$\mathbf{Con}(T)$ 和1-$\mathbf{Con}(T)$之间的本质区别。理论T的反演原则，记为$\mathbf{Rfn}_T$，是形如$\mathbf{Pr}_T(\ulcorner\phi\urcorner) \to \phi$的语句模式，其中$\phi$ 为任意T语言中的语句。令Γ-$\mathbf{Rfn}_T$ 表示限制到语句类Γ 上的理论T的反演原则。

事实 2 ([58]) *令T 为$\mathbf{PA}$的递归公理化的一致扩充。则在T 中，$\mathbf{Con}(T)$与Π^0_1-$\mathbf{Rfn}_T$ 等价，且1-$\mathbf{Con}(T)$ 与Σ^0_1-$\mathbf{Rfn}_T$ 等价。*

作为事实2的一个推论，1-$\mathbf{Con}(T) \vdash$ 1-$\mathbf{Con}(T + \mathbf{Con}(T))$ (参见[49] 中命题3). 因此，若我们假设1-$\mathbf{Con}(T)$成立, 则我们可得$\mathbf{Con}(T)$, $\mathbf{Con}(T + \mathbf{Con}(T))$, $\mathbf{Con}(T+\mathbf{Con}(T+\mathbf{Con}(T)))$，由此类推下去（事实上，1-$\mathbf{Con}(T)$比这些语句都强）。

[11] 令$\mathbf{Pr}^R_T(x)$ 是罗瑟可证谓词。理论T的罗瑟语句θ是我们应用不动点引理到谓词$\neg\mathbf{Pr}^R_T(x)$上所得的不动点语句：即$T \vdash \theta \leftrightarrow \neg\mathbf{Pr}^R_T(\ulcorner\theta\urcorner)$. 我们可证罗瑟语句$\theta$是独立于理论$T$的。

由上可知，哥德尔语句与罗瑟语句的区别，以及$\mathbf{Con}(T)$ 与1-$\mathbf{Con}(T)$间的区别是十分重要的。然而，这些区别在关于不完全性定理的非形式的哲学讨论中常被忽略。

第四、无论是数学上还是哲学上，G2与G1 是本质上不同的。G1的涵义并非取决于算术化方法及可证谓词的使用；在这个意义上，我们可以说G1是外延性的。而G2的涵义是取决于我们如何表达理论的一致性；在这个意义上，我们可说G2是内涵性的。G2对一理论是否成立取决于很多因素。关于G2的内涵性问题，参见本文第6节的讨论。

3 不完全性定理在不同领域中的影响

不完全性定理自其发表以来对逻辑、经典数学、理论计算机科学、哲学等领域都产生了广泛、深刻而持久的影响。本节中，我们基于文献中关于不完全性定理的研究成果分析其对逻辑学（特别是数学基础）、经典数学、理论计算机科学、哲学的影响。深入分析不完全性定理对这里任何一个领域的影响都需要专著的篇幅。事实上，不完全性定理的影响不限于这些领域。限于篇幅，本文仅讨论不完全性定理对这几个主要领域的影响的主要方面。

不完全性定理对数学基础的影响主要体现在如下五个方面。第一、不完全性定理在一定意义上揭示了形式化方法本质上的局限性。不完全性定理揭示了包含$\mathbf{Q}$的递归公理化的一致扩充理论的证明能力的限度。不管理论T有多强，只要T是递归公理化的一致理论，且包含关于算术的足够信息（如$\mathbf{Q}$），那么总存在理论T中的不可判定命题。罗宾逊算术$\mathbf{Q}$ 是很弱的算术理论，因此我们可以说不完全性定理揭示了大多数形式理论的证明能力的限度。

第二、不完全性定理揭示了“真”和“可证”这两个概念间的本质区别。在不完全性定理之前，人们普遍认为“一个数学命题是真的”和“一个数学命题是可证的”是两个相同的概念。不完全性定理揭示了“在$\mathbf{PA}$中可证”和“在标准模型下为真”这两个概念间的本质区别：在$\mathbf{PA}$中可证的算术命题在标准模型下都为真，但存在独立于$\mathbf{PA}$的算术真命题。令$\mathbf{Truth}$为所有在算术的标准模型中为真的算术语句的集合，$\mathbf{Prov}$为所有在$\mathbf{PA}$ 中可证的算术语句的集合。不完全性定理揭示了$\mathbf{Prov}$是$\mathbf{Truth}$的真子集。作为不动点引理的应用，塔斯基真不可定义定理告诉我们：$\mathbf{Truth}$在算术语言中是不可定义的。而$\mathbf{Prov}$在算术语言中是可定义的，但不是递归集。

第三、不完全性定理否定了怀特海-罗素的一种强版本的逻辑主义纲领：期望建构一种形式系统使得在其中可以形式化全部数学理论，且可证明所有的数学定理，从而表明数学在本质上可还原为逻辑。G1直接否定了这种强版本的逻辑主义纲领：不存在一种理论在其中可以证明所有的数学真理。

第四、不完全性定理对希尔伯特纲领的发展产生了深刻的影响。希尔伯特纲领有两个主要目标：（1）证明$\mathbf{PA}$是完全的（所有关于算术的真命题都在$\mathbf{PA}$中可证）；（2）用有限性方法证明算术理论的一致性。G1直接否定了第一个目标。人们通常认为G2也直接否定了希尔伯特纲领，对此学者们有不同的看法（参见[17]）。我们没有关于“有限性方法”这个概念的精确定义。对于何谓“有限性方法”，人们有不同的看法。若“用有限性方法证明算术命题A”是指A在$\mathbf{PA}$中可证，那么在这种意义上，我们可以说G2直接否定了希尔伯特纲领的第二个主要目标。若我们将可数序数长度的超穷归纳法也看作“有限性方法”，则根岑的结果表明，我们可在某种比$\mathbf{PA}$强的理论中使用有限性方法证明$\mathbf{PA}$的一致性。因此，G2并非直接否定了希尔伯特纲领，而是表明了希尔伯特纲领应用范围的局限性（其全局的实现是不可能的）。在不完全性定理之后，希尔伯特纲领在序数分析、反推数学等领域中得到很好的局部实现。[12]

第五、不完全性定理是数学基础领域重要的否定性结果。在其影响下，人们之后发现了很多逻辑中的其他重要否定性结果，如塔斯基真不可定义定理、丘奇一阶理论不可判定性定理、图灵停机不可判定性定理、连续统假设独立性定理等。这些否定性结果揭示了逻辑中的一些核心概念（如可证性、可定义性、可计算性，可判定性、独立性）的限度，是数理逻辑与数学基础领域最重要深刻的成果之一。

下面我们讨论不完全性定理对经典数学的影响。不完全性定理是原创性的数学成果，其证明使用了数学方法。例如，在哥德尔的证明中，他使用了如中国剩余定理和素数唯一表示定理之类的数论方法，以及如算术化、可表示性、自指构造之类的元数学（证明论）方法。不完全性定理揭示了在逻辑和数学中普遍存在的不完全性现象。不完全性定理告诉我们，任何包含足够算术信息（如$\mathbf{Q}$）的一致递归公理化理论都存在不可判定命题（即“遗漏”一些算术真理）。一种流行的看法是：不完全性定理的证明使用了算术化、自指构造等纯逻辑方法，与经典数学没有什么关联；哥德尔语句是

[12]关于不完全性定理之后希尔伯特纲领的发展及其对数理逻辑（特别是证明论）和数学哲学的影响，文献中有大量的讨论。参见代表性的文献：[21, 61, 72]。

纯粹的逻辑构造，没有实在的数学含义。的确，哥德尔语句表达的不是关于自然数的性质，而是关于算术理论自身的系统性质。在这个意义上，我们可以说哥德尔的证明揭示的是逻辑上的不完全性。在哥德尔之后，一个自然的重要问题是：算术理论**PA**是否数学上完全的？即经典数学中的算术语句是否都在**PA**中可证或可否证？[13] 是否存在具有实在数学含义的不可判定的算术命题（如数论、组合中的命题）？

我们称由具有实在数学含义的不可判定命题所揭示的不完全性现象为具体不完全性。事实上，具体不完全性在数学中是普遍存在的。在哥德尔之后，人们在经典数学中发现了很多具有实在数学含义的独立于**PA**的算术语句。[14] 这些经典数学中的不可判定语句的构造没有使用算术化方法和可证谓词，具有实在的数学含义。事实上，这些具有实在数学含义的独立性语句比哥德尔语句更复杂。哥德尔语句在**PA**中等价于典范一致性命题$\mathbf{Con}(\mathbf{PA})$，而这些独立性语句不仅独立于**PA**，而且独立于$\mathbf{PA}+\mathbf{Con}(\mathbf{PA})$（参见[4]）。一个有趣而令人惊讶的事实是，很多经典数学中的具有实在数学含义的不可判定语句实际上是等价于某种可通过算术化方法来表示的元数学命题。这一事实揭示了经典数学中的不可判定命题与元数学（证明论）中的算术命题间的关联。

哥德尔之后的研究揭示了比**PA**更强的理论（如高阶算术和集合论）的数学上的不完全性。例如，哥德尔和柯恩证明了连续统假设是独立于公理集合论**ZFC**的。事实上，人们在很多数学领域中（如分析、代数、拓扑、数理逻辑等）发现了大量具有实在数学含义的在**ZFC**中不可判定的数学命题。

哈维-弗里德曼（Harvey Friedman）是哥德尔之后数学基础领域的国际知名专家。他在具体不完全性领域作出了很多重要的贡献。他指出："数学基础研究的长期广泛影响和重要性在很大程度上取决于不完全性现象在经典数学中的普遍程度"（参见[27]）。哈维-弗里德曼的工作将具体不完全性的研究从一阶算术扩充到高阶算术和公理集合论。哈维-弗里德曼的专著"布尔关系理论与不完全性（Boolean Relation Theory and Incompleteness）"是关于具体不完全性研究前沿进展的最新文献。在[27] 中，哈维-弗里德曼系统地研究并给出很多在一阶算术**PA**，高阶算术及**ZFC**的不同强度的系统

[13]公式A在**PA**中可否证是指$\neg A$在**PA**中可证。

[14]例如，Kanamori-McAloon principle, the Kirby-Paris sentence, the Hercules-Hydra game, the Worm principle, the flipping principle, the arboreal statement, P.Pudlák's Principle, the kiralic and regal principles。关于这些语句的内容参见[4, 7, 10]。

中来自经典数学的不可判定命题。文献[15, 10] 中给出一个在二阶算术中可表达的具体数学定理，并证明了此定理在二阶算术和三阶算术中都不可证，但在四阶算术中可证。[15]

关于具体不完全性的研究是重要而深刻的，其揭示了经典数学中的不完全性现象，及不完全性定理对经典数学的影响。哈维-弗里德曼关于具体不完全性的研究纲领期望表明我们可在经典数学的几乎任何一个领域中找到有实在数学含义的在**ZFC**中不可判定的数学命题。哈维-弗里德曼的研究纲领是重要且有前景的。此纲领若能得到成功实现，这表明不完全性在数学中是无处不在的，这将是哥德尔之后数学基础领域最重要的发现之一。

下面我们简要讨论不完全性定理对理论计算机科学的影响。理论计算机科学的一个重要研究主题是计算的能力和限度。哥德尔的工作与可计算性理论紧密相关。哥德尔的证明中包含了可计算性理论中一些重要思想的萌芽：如算术化方法、递归函数等。哥德尔在证明中使用的递归函数和算术化方法是理论计算机科学研究中两种很重要的工具。递归函数是在哥德尔之后发展起来的递归论（可计算性理论）这一数理逻辑分支的基础概念和理论基石之一。算术化方法的思想是用自然数来给理论的一些语法对象如项、公式、证明等编码。算术化方法在理论计算机科学研究中是一种关键且有用的技术，在递归论的发展中扮演了重要的角色。

不完全性定理和不可判定性理论间也紧密相关。我们称一致理论T是实质不可判定的，若T的相同语言中的任意递归公理化的一致扩充都是不可判定的。我们可证明，理论T是实质不可判定的当且仅当T是实质不完全的。这表明，完全性/不完全性与可判定性/不可判定性之间是紧密关联的。

在理论计算机科学中，否定性结果构成一种重要而独特的传统，其中的一个典型例子是图灵停机不可判定性定理。图灵的停机问题与不完全性定理间也有紧密的联系。我们可通过停机问题不可判定性定理证明G1和G2。哥德尔于1934年给出一般可计算函数的精确数学定义，之后丘奇和图灵独立地给出可计算函数不同的精确数学定义，后来人们证明了可计算函数的这些不同的定义是等价的。更多关于不完全性定理对理论计算机科学的影响的讨论，参见研究文集[2]“哥德尔与数学基础：真的视野”。

和其他数学定理不同的是，不完全性定理的影响力并非仅限于数学家和逻辑学家的圈子，它在哲学中也有广泛的影响。不完全性定理在哲学领域的影响表现在，自其发表以来，它引发了与其相关的一些哲学问题的广

[15]更多关于具体不完全性的研究，参见[7, 10, 27].

泛而持久的讨论：如逻辑与数学的本质，心智与机器的本质，人类智能与机器智能间的区别，机器智能的限度，不完全性定理与逻辑主义、希尔伯特纲领及直觉主义间的关联等。学界已有不少关于不完全性定理的哲学意义及其在哲学中的应用的讨论，但人们对于这些哲学中的应用的合理性是有争议的。关于不完全性定理在哲学上的相关性，参见[50]。关于哥德尔的哲学思想及其对于发现不完全性定理的影响，参见王浩的著作[68, 69]。关于不完全性定理的正确使用与误用，参见[25]。下文中我们重点讨论两个与不完全性定理紧密相关的哲学论题：反机械主义论题和哥德尔析取论题。

学界已有不少基于不完全性定理的对“人类心智是否可机械化”这一哲学问题的讨论。[16] 机械主义论题认为人类心智是可机械化的；而反机械主义论题认为人类心智是不可机械化的。[17] 图灵提出了一种计算模型：图灵机，并通过图灵机模型给出“可机械化”这一概念精确的数学定义。反机械主义论题可更精确地表述为：人类心智的数学输出超越任何图灵机的数学输出。这里，我们讨论的重点不是“人类心智是否可机械化”这个一般的哲学问题，而是不完全性定理与反机械主义论题间的关联。对G1的一种流行解读是：G1逻辑上蕴涵反机械主义论题。文献中有不少基于不完全性定理的支持反机械主义论题的论证，其中最著名的是卢卡斯（Lucas）论证和彭罗斯（Penrose）论证。[18] 卢卡斯论证在文献中受到广泛的批评。[19] 彭罗斯在[48]中提出一个比卢卡斯论证更复杂精细的支持反机械主义论题的论证。彭罗斯论证是到目前为止支持反机械主义论题的最复杂最有前景的一个论证，其在文献中被广泛讨论、仔细分析。[20] 最新研究表明，基于不完全性定理的支持反机械主义论题的论证（即认为G1逻辑蕴涵反机械主义论题）源于对不完全性定理的某些误解。[21] 现今绝大多数哲学家和逻辑学家认为卢卡斯论证和彭罗斯论证及其变体都不是令人信服的。然而，人们在对卢卡斯论证和彭罗斯论证的错误根源的认识上存在一定的分歧。克拉耶夫斯基（Krajewski）在最近的文章中详细讨论了基于不完全性定理的支持反机械主义论题的论证的历史，仔细分析了卢卡斯论证和彭罗斯论证的问

[16]参见[9, 11, 23, 39, 40, 55]。

[17]单独个体的心智能力是很有限的，这里我们不是考虑单独个体的心智能力，而是考虑理想化的人类心智本质上可以做什么。人类心智包括很多维度，这里我们仅考虑人类心智的数学能力：即人类心智所能发现证明的数学真理。

[18]参见[44, 47]。

[19]关于对卢卡斯论证的讨论的历史，参见[23]。关于对卢卡斯论证的有影响的批评，参见[5]。

[20]参见[43, 56, 39]。

[21]参见[60, 11, 23]。

题所在，并得出“G1并非逻辑蕴涵反机械主义论题”这一结论（参见[60]）。[22]

哥德尔本人并非认为G1 逻辑蕴涵反机械主义论题，尽管他相信人类心智本质上不能被机械化且可认识所有的数学真理。哥德尔相信，与图灵机相比，人类心智的独特之处在于它能提出新公理，建构新的数学理论。基于他的理性乐观主义，哥德尔相信人类心智是算术上全能的：即人类心智能够认识所有的算术真理。然而，哥德尔承认，他既不能给出“人类心智不可机械化”的令人信服的论证，也不能给出“不存在绝对不可判定命题”的令人信服的论证。哥德尔认为，从他的不完全性定理他至多可得出如下一个更弱的结论（称为“哥德尔析取论题(GD)”）：若人类心智是可机械化的，则存在绝对不可判定命题（人类心智无法认识的数学真理）。[23] 哥德尔析取论题等价于一析取式命题，我们称“人类心智不可机械化”为GD的第一析取支命题，称“存在绝对不可判定命题”为GD的第二析取支命题。哥德尔析取论题关乎人类智能与机器智能的界限及人类心智认识能力的限度。在哥德尔看来，GD是可由他的不完全性定理推导出来的有哲学意义的数学定理。[24]

下面我们讨论不完全性定理是否蕴涵GD。在我们分析GD前，我们先分析GD涉及的两个核心概念：相对可证性、绝对可证性。我们对相对于某一形式系统而言的相对可证性概念有严格精确的定义。而绝对可证性概念相对而言是更模糊的，我们目前没有关于绝对可证性概念的严格形式定义。绝对可证性概念不是相对于某一形式系统而定义的，其与相对可证性概念间有本质的区别。学界对绝对可证性概念应满足的性质尚未达成很多共识。

令$\langle M_n : n \in \omega\rangle$ 为所有图灵机的列举。我们称$Th(M_n)$ 为图灵机M_n的输出理论，即所有输出语句的集合。下文中，我们将在理论S中可证的语句集等同于理论S对应的图灵机的输出理论。[25]令$\mathbf{K}$为人类心智能知道（认识）的数学命题的集合。令$\mathbf{T}$为所有数学真命题的集合。[26] 我们称理论S是可靠的，若$S \subseteq \mathbf{T}$。下文讨论中，我们假设$\mathbf{K}$是可靠的。我们将绝对可证概念等

[22]更多关于不完全性定理与反机械主义论题间的关系的讨论，参见[39, 40]。

[23]哥德尔最初在1951年的讲演中提出他的析取论题，他写道：“如下的析取命题是不可避免的：或者人类心智超越任何图灵机，或者存在绝对不可判定问题（这两个析取支命题都成立也是可能的，因此严格意义上有三种可能性）”（参见[32]）。

[24]更多关于哥德尔析取论题的讨论，参见最新研究文集[37]。

[25]任一形式理论都可看作某一图灵机的输出理论。

[26]哥德尔称$\mathbf{T}$为客观数学，称$\mathbf{K}$为主观数学。

同于人类心智的可知性概念。在此假设下，$\mathbf{K}$即是所有绝对可证的数学命题的集合。

若未作特别说明，我们假定$Th(M_n)$是$\mathbf{Q}$的扩充，这样G1 和G2 都可应用到$Th(M_n)$ 上。我们称数学命题ϕ是绝对不可判定的，若$\phi \notin \mathbf{K}$ 且$\neg\phi \notin \mathbf{K}$。关于$Th(M_e), \mathbf{K}, \mathbf{T}$ 三者间的关系，哥德尔提出如下两个论点：

(1) 对任意的$e \in \mathbb{N}$, $\mathbf{K}(Th(M_e) \subseteq \mathbf{T}) \to Th(M_e) \subsetneq \mathbf{K}$。

(2) 析取论题：或者$\neg\exists e(Th(M_e) = \mathbf{K})$ 成立，或者$\exists\phi(\phi \in \mathbf{T} \wedge \phi \notin \mathbf{K} \wedge \neg\phi \notin \mathbf{K})$ 成立。

莱因哈特(Reinhardt)在[53]中提出系统$\mathbf{EA_T}$，并证明了若相对可证性、绝对可证性和真这三个概念满足某些原则，则我们可给出GD的严格证明；这证实了哥德尔的观点：GD是个数学定理（参见[39]）。系统$\mathbf{EA_T}$是一种基于认知算术的类型真理论，其公理包括算术公理，关于绝对可证性的公理，和类型真理论的公理。[27] 哥德尔如上的两个论点在$\mathbf{EA_T}$中都可形式化。莱因哈特在[53]中证明了GD在$\mathbf{EA_T}$中可证。

基于$\mathbf{EA_T}$，我们可更精细地分析机械主义论题。首先，我们可区分三种不同程度的机械主义论题：

- 弱机械主义论题(WMT): $\exists e(\mathbf{K} = Th(M_e))$;
- 强机械主义论题(SMT): $\mathbf{K}\exists e(\mathbf{K} = Th(M_e))$;
- 超强机械主义论题(SSMT): $\exists e\mathbf{K}(\mathbf{K} = Th(M_e))$.

本文中我们所讨论的机械主义论题是指弱机械主义论题：存在一图灵机使得人类心智和此图灵机具有相同的数学输出。强机械主义论题是指人类心智知道：存在一图灵机使得人类心智和此图灵机具有相同的数学输出。而超强机械主义论题是指有一个特定的图灵机M 使得：人类心智知道其自身和图灵机M具有相同的数学输出。

莱因哈特[51] 中证明了$\mathbf{EA_T}$+SSMT 是不一致的。这表明在$\mathbf{EA_T}$中，超强机械主义论题不成立：不存在一个特定的图灵机M使得人类心智知道其自身和图灵机M具有相同的数学输出。莱因哈特[52]中证明了$\mathbf{EA_T}$ + WMT 是一致的。这表明GD的第一析取支命题($\neg\exists e(\mathbf{K} = Th(M_e))$)在$\mathbf{EA_T}$中不可证。即从$\mathbf{EA_T}$的角度看，如下是可能的：人类心智和某台图灵机具有相同

[27]关于系统$\mathbf{EA_T}$的细节，参见[39]。

的数学输出。卡尔森(Carlson) [8] 中证明了$\mathbf{EA_T}+\mathsf{SMT}$是一致的。这表明，从$\mathbf{EA_T}$的角度看，如下是可能的：人类心智知道其自身是某台图灵机；只是人类心智不知道其自身是具体的哪台图灵机。哥德尔认为我们未来可能可以证明GD的第一析取支命题：人类心智超越任何图灵机的数学输出。我们所缺少的是对于包含真概念的自指性悖论的充分解决方案。哥德尔认为：若我们能解决内涵性的悖论问题，我们将可给出人类心智超越任何图灵机的证明（参见[69]）。

下面我们对不完全性定理是否蕴涵反机械主义论题作个总结。由第一不完全性定理可得$\neg\exists e\,\mathbf{K}(\mathbf{K}=Th(M_e))$；由莱因哈特[52]中定理可得，这并非蕴涵$\neg\exists e(\mathbf{K}=Th(M_e))$；由卡尔森[8]中定理可得，这也并非蕴涵$\neg\mathbf{K}\exists e\,(\mathbf{K}=Th(M_e))$。因此，语句"$\mathbf{K}=Th(M_e)$"前加上的是"$\exists e\,\mathbf{K}$"还是"$\mathbf{K}\exists e$"，这两者间是有本质区别的。在$\mathbf{EA_T}$看来，如下是可能的：$\mathbf{K}\exists e(\mathbf{K}=Th(M_e))$；但如下是不可能的：$\exists e\mathbf{K}(\mathbf{K}=Th(M_e))$。

下面我们讨论科尔纳（Koellner）关于GD的工作（参见[39, 40]）。科尔纳认为，要精确地讨论GD并证明不完全性定理蕴涵GD，首先得在恰当的形式系统中形式化GD。在莱因哈特等人工作的基础上，科尔纳构造了形式系统**DTK**，其也是基于认知算术的一种形式真理论。但与$\mathbf{EA_T}$不同的是，$\mathbf{EA_T}$是一种类型真理论，而**DTK**是一种无类型真理论。[28] 科尔纳在**DTK**中形式化了GD及其两个析取支命题，并证明了如下结论：（1）GD在**DTK**中可证；（2）卢卡斯论证和彭罗斯论证在**DTK**中都不成立；（3）GD的两个析取支命题"人类心智不可机械化"和"存在绝对不可判定命题"在**DTK**中都不可证。最后，科尔纳作出如下结论："人类心智是否可机械化"和"是否存在绝对不可判定命题"这两个问题本身可能就是绝对不可判定命题（参见[40]）。

科尔纳对GD的分析是本人所知的现有文献中对此论题的最全面精确的分析。不管我们是否认同科尔纳对GD的分析，他这一工作的意义在于提出了一个自然的系统并在其中证明了GD，严格分析了卢卡斯论证和彭罗斯论证都不成立的原因，并证明了GD的两个析取支命题在此系统中都不可证。所有这些技术性的成果都有力地支持了哥德尔对于相关问题的看法。

人们常用"深刻"这个词评价不完全性定理。对于何谓"深刻的数学定理"，学界有不同的看法。即使人们对某一定理是否深刻达成共识，人们对此定理"为何深刻"的理由仍可能有不同的看法。对"数学深度"这一概

[28]在**DTK**中，我们可形式化彭罗斯论证，因彭罗斯论证使用了无类型的真概念（参见[39, 40]）。

念，我们没有精确的定义，目前也没有一个判断给定数学定理是否深刻的被广泛接受的一般标准。不完全性定理被广泛认为是逻辑领域“深刻”的数学定理。[14]将不完全性定理作为研究数学定理深度的一个案例，并研究这样两个方法论问题：刻画不完全性定理的深度的标准是什么，如何基于这些标准论证不完全性定理的深度。基于文献中最新研究成果，[14]提出了三个刻画不完全性定理的深度的标准：成果的影响力、结论的丰富性、理论的统一性，并基于这三个标准论证了不完全性定理的深度。

4 不完全性定理的不同证明

哥德尔之后，人们发现了不完全性定理的许多不同的证明。G1和G2在性质上是本质不同的。G1关乎的是一理论是否完全，即此理论是否存在不可判定命题。对于证明G1而言，我们仅需证明存在一不可判定命题，不管此命题是由经典数学方法构造而得还是由纯逻辑方法构造而得。G1的证明就不可判定命题的性质而言可分为两类：一类证明中不可判定命题是由纯逻辑方法构造而得，另一类证明中不可判定命题是经典数学（如分析、数论、组合）中有实在数学含义的命题。G2关乎的是系统的一致性是否在系统内可证。这涉及如何在系统内表达系统的一致性的问题。系统的一致性是个纯粹的逻辑问题，而不是经典数学中关心的主要问题。基于文献中不完全性定理的不同证明，我们提出九个特征对这些不同的证明进行分类。

我们先给出若干定义。我们称G1的证明是构造性的，若存在一个能行的算法将不可判定命题显式地构造出来。G1的非构造性证明仅仅断定不可判定命题的存在性，但没有能行地将此不可判定命题显式地构造出来。我们称G1的证明具有罗瑟特征，若它的证明仅需假设理论是一致的，而不需假设更强的结论（如理论是ω-一致的）。

我们可依据如下特征对文献中不完全性定理的不同证明进行分类：使用证明论方法，使用递归论方法，使用模型论方法，使用算术化方法，使用不动点引理，基于逻辑悖论的证明，构造性的证明，具有罗瑟特征，不可判定命题有实在的数学含义。特别地，哥德尔第一不完全性定理的证明具有如下特征：使用了算术化等证明论方法，没有直接使用不动点引理，证明是构造性的，不具有罗瑟特征，哥德尔语句类似于说谎者悖论中的说谎者语句，哥德尔语句没有实在的数学含义。不完全性定理的这些不同的证明方法揭示了不同领域间的关联：证明论、递归论、逻辑悖论、模型论、经典数学等。关于不完全性定理不同证明的详细讨论，参见[13]。

下面我们对不完全性定理的不同证明作几点说明。第一、以上特征只是强调不完全性定理证明的一个方面，不完全性定理的一种证明可能同时具有以上若干特征。算术化方法也可看作一种证明论方法，但证明论方法包含更广的范围。不动点引理的证明也是构造性的，但特征"使用不动点引理"强调的是证明中是否直接使用了不动点引理，不完全性定理的一个证明可能是构造性的但没有直接使用不动点引理。

第二、以上的任何单个的特征都不是证明不完全性定理的必要条件。对于以上的任何单个特征，文献中我们既可找到具有此特征的证明的例子，也可找到不具有此特征的证明的例子。例如，G1也有非构造性的证明，不直接使用不动点引理的证明，不使用逻辑悖论的证明，具有罗瑟特征的证明，不使用算术化方法的证明等。

第三、很多关于G1和G2的证明都使用了算术化方法。格泽戈尔奇克(Grzegorczyk) 在[33] 中提出一种用来研究不完全性与不可判定性的新理论**TC**，并不直接使用算术化方法证明了**TC**是不完全的（参见[33]）。在**PA**中，我们有可用来相加或相乘的自然数，而在理论**TC**中，我们有可用来拼接的字符串。在哥德尔的证明中，自然数主要被用来给语法对象编码。基于如下原因，我们可使用字符串而不是自然数作为更基本的概念：从数学的层面，我们可不使用自然数定义可计算性的概念；从哲学的层面，使用字符串可得到很好的辩护，我们日常的推理与交流都是使用字符串而不是自然数，很自然我们可考虑通过字符串而不是自然数来定义不可判定性概念。

第四、不完全性定理与可计算性理论中关于不可判定性的若干结果是紧密相联的。可计算性理论中的一个重要结果是停机问题不可判定性定理：不存在一个计算机程序P，使得对于任意输入的程序w，能够判断w是否会在有限时间内结束或者死循环。克莱恩(Kleene) 于1943年给出一个基于停机问题不可判定性定理的不完全性定理的证明。克莱恩指出，若存在一个递归公理化且满足一定一致性条件的完全的算术系统，则停机问题是可判定的，因此这导致矛盾（参见[73]）。斯莫林斯基(Smorynski) 于1977 年给出基于递归不可分集的第一不完全性定理的证明（参见[73]）。

第五、当前关于不完全性定理的研究揭示了不完全性定理与逻辑悖论间的紧密关联。文献中有不少基于逻辑悖论的不完全性定理的证明。哥德尔在[31]中指出：任何关于认知概念和"真"概念的悖论都可用来构造证明不完全性定理的不可判定命题。说谎者悖论是个古老且著名的悖论。我们可将哥德尔第一不完全性定理证明中使用的哥德尔语句看作说谎者语句的某

种变体。哥德尔语句类似于说谎者悖论中的说谎者语句，但与后者有本质的区别。说谎者语句使用的是真概念，而哥德尔语句使用的是可证概念。真概念和可证概念间具有本质的区别。说谎者语句可导致逻辑矛盾，而哥德尔语句并不导致逻辑矛盾。基于逻辑悖论的不完全性定理的证明的关键是恰当地形式化这些逻辑悖论。在文献中，还有很多其他的逻辑悖论被用来证明不完全性定理：如贝瑞（Berry）悖论，格林灵-纳尔逊（Grelling-Nelson）悖论，意外考试悖论，亚布洛（Yablo）悖论等(参见[13])。这些基于逻辑悖论的不同证明证实了哥德尔关于不完全性与逻辑悖论间的关联的观点。

第六、很多关于G1和G2 的证明都使用了自指构造的方法。虽然我们没有一个关于自指性的严格数学定义，但不动点引理通常被视为是使用了自指性的构造。G1也有不使用自指性构造的证明方法。维瑟(Visser)在[67]中由塔斯基真不可定义定理给出G2的一个"非自指性"证明。此研究也揭示了不完全性定理与塔斯基真不可定义定理间的关联。

第七、文献中很多不完全性定理的证明使用的是纯逻辑的方法，构造的不可判定命题没有实在的数学内容。哥德尔之后，一个自然的问题是：我们能否构造具有实在数学内容的在**PA**中不可证的算术真命题。具体不完全性的研究目标正是寻找这样的具有实在数学内容的不可判定命题。在哥德尔之后，人们发现了很多经典数学中的具有实在数学内容的不可判定的算术命题。这些不可判定命题的构造没有使用例如算术化和可证谓词之类的纯逻辑方法。例如，帕里斯-哈林顿(Paris-Harrington)原则是经典数学中具有实在组合内容的不可判定命题。更多关于具体不完全性的研究及经典数学中不可判定命题的例子，参见[7, 10, 13, 27]。

5 不完全性定理的推广形式

本节中，我们简要介绍G1和G2的不同推广形式。第一不完全性定理既可推广到**PA**的扩充理论，也可推广到在解释度意义上比**PA**弱的理论。文献中有很多G1的不同推广形式。这里我们重点介绍几种推广形式。

我们先引入若干定义。令T为$\mathbf{Q}$的一致扩充。我们称理论T是Σ^0_n可定义的，若存在Σ^0_n公式$\phi(x)$使得n是T中某一公理的哥德尔编码当且仅当$\phi(\overline{n})$在算术的标准模型中为真。我们称理论T是算术可定义的，若对某一自然数n，T是Σ^0_n可定义的。我们称$A \subseteq \mathbb{N}$是递归可枚举的，若A是某一递归函数的像集。[29]

[29]即存在一递归函数f使得$A = \{f(x) : x \in dom(f)\}$。

G1告诉我们，$\mathbf{PA}$的任意递归公理化的一致扩充都是不完全的。事实上，G1对$\mathbf{PA}$的递归可枚举的一致扩充也成立。一个自然的问题是：G1是否对$\mathbf{PA}$的非递归可枚举但算术可定义的扩充成立？事实上，在一定条件下，G1也可推广到某些（但并非所有）非递归可枚举但算术可定义的理论。[30]

我们称理论T是Σ^0_n-可靠的，若对任意的Σ^0_n语句ϕ, 如果ϕ 在T中可证，则ϕ在算术的标准模型中为真。我们称理论T是Π^0_n-完全的，若对任意的Π^0_n语句ϕ, 或者$T \vdash \phi$或者$T \vdash \neg\phi$。菊池-仓桥(Kikuchi-Kurahashi)在[38]中证明了若理论T是Σ^0_{n+1}-可定义的且是$\mathbf{Q}$的Σ^0_n-可靠的扩充，则T并非是Π^0_{n+1}-完全的。故G1 可推广到$\mathbf{Q}$的Σ^0_{n+1}-可定义且Σ^0_n-可靠的扩充理论。

解释的概念给我们提供了一种比较不同语言下不同理论的强度的方法。直观地说，给定理论T和S，T在S中的一个解释是从T语言的公式集到S语言的公式集的一个转换函数，使得T中的公理在此转换函数下对应的公式在S中可证。关于解释概念的精确定义，参见[65]。给定理论S和T，我们用$S \trianglelefteq T$表示S 在T中可解释，用$S \lhd T$表示S在T中可解释但T不在S中可解释。我们称理论S 和T是相互可解释的，若$S \trianglelefteq T$且$T \trianglelefteq S$。给定算术理论U和V，我们用$U \leq_T V$表示理论U（作为编码集）是图灵可归约于理论V（作为编码集）；用$U <_T V$表示$U \leq_T V$但$V \not\leq_T U$。

我们也可通过解释的概念以如下方式表述G1：对任意递归公理化的一致理论S，若$\mathbf{Q}$在S中可解释，则S是不完全的。给定一致理论T，我们称G1对T成立，若对任意递归公理化的一致理论S，若T在S中可解释，则S 是不完全的。我们可证明，G1 对理论T成立当且仅当T是实质不可判定的当且仅当T 是实质不完全的（参见[12]）。

我们称理论S 有极小的解释度，若不存在理论T使得$T \lhd S$。我们称理论S 有极小的图灵度，若不存在理论V使得$V <_T S$。本文中，若$S \lhd T$，则我们称理论S在解释度意义上比理论T弱；若$S <_T V$，则我们称理论S在图灵度意义上比理论V弱。通过解释的概念，G1可推广到很多具有不同语言的理论。我们知道G1对$\mathbf{PA}$成立。事实上，G1对很多在解释度意义上比$\mathbf{PA}$弱的理论也成立，[11]中列出了很多G1对其成立的理论的例子。

下面我们引入在不完全性和不可判定性研究中的一个比$\mathbf{Q}$弱的重要算术理论：理论$\mathbf{R}$，其是由塔斯基、莫斯托夫斯基和罗宾逊在[63]中引入的。

定义 2 (理论$\mathbf{R}$) 理论$\mathbf{R}$的语言包含如下非逻辑符号：常元符号$\mathbf{0}$，一元函数符号$\mathbf{S}$，两个二元函数符号$+$ 和$\times$, 及一个二元关系符号$\leq$。理论$\mathbf{R}$ 包含

[30]参见[38, 54]。

如下公理模式**Ax1-Ax5**：[31]

Ax1 $\overline{m} \neq \overline{n}$ 若$m \neq n$;

Ax2 $\overline{m} + \overline{n} = \overline{m+n}$;

Ax3 $\overline{m} \times \overline{n} = \overline{m \times n}$;

Ax4 $\forall x(x \leq \overline{n} \rightarrow x = \overline{0} \vee \cdots \vee x = \overline{n})$;

Ax5 $\forall x(x \leq \overline{n} \vee \overline{n} \leq x)$.

理论**R**的重要性在于：其包含了所有证明G1所需的关于算术的性质。例如，哥德尔在证明G1时用到一个重要结论：所有递归函数都在**PA**中可表示；但事实上，所有递归函数都在**R**中可表示。

我们知道G1对理论**Q**和**R**都成立。理论**R**在解释度意义上比**Q**更弱。人们常认为理论**R**是使得G1成立的在解释度意义上的最弱理论。事实上，G1对很多在解释度意义上比理论**R**更弱的理论也成立。关于使得G1成立的解释度比**R**弱的理论的例子，参见[11]。本人在[11]中证明了不存在图灵度意义上的使得G1成立的极小的递归可枚举理论。G1丰富的推广形式表明不完全性现象在逻辑中是普遍存在的。

G2也有多种推广形式：G2不仅可推广到**PA**的非递归可枚举但算术可定义的扩充理论，也可推广到解释度比**PA**弱的理论。这里我们仅举三个典型例子（更多关于G2的推广形式的例子，参见[13]）。

第一、和G1一样，G2也可推广到**PA**的非递归可枚举但算术可定义的扩充理论。[32] 第二、我们也可通过解释的概念推广G2。普德拉克(Pudlák)的一个经典结果是：不存在递归公理化的一致理论T使得$\mathbf{Q} + \mathbf{Con}(T)$ 在T中可解释。由此可得，对任意递归公理化的一致理论T，若**Q**在T中可解释，则G2对T成立（即$\mathbf{Con}(T)$在T中不可证）。

第三、洛布(Löb)定理是G2的一种重要推广形式：对**Q**的任意递归公理化的一致扩充理论T，对任意标准可证谓词$\mathbf{Pr}_T(x)$和公式ϕ，若$T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner) \rightarrow \phi$，则$T \vdash \phi$。作为洛布定理的一个推论，G2对$T$成立。应用洛布定理到**PA**上，若$\mathbf{PA} \nvdash \phi$，则$\mathbf{Pr}_{\mathbf{PA}}(\ulcorner \phi \urcorner) \rightarrow \phi$ 是在算术标准模型中为真但在**PA**中不可证的语句。因此，由洛布定理我们可得到很多在**PA**中不可判定的算术语句。

[31]对任意自然数n, 我们如下递归定义项$\overline{n}$: $\overline{0} = \mathbf{0}$, $\overline{n+1} = \mathbf{S}\overline{n}$。

[32]参见[38, 54]。

洛布定理揭示了不完全性定理与可证性逻辑间的关联。可证性逻辑是研究不完全性及算术的元数学的一种重要有用的工具。历史上，可证性逻辑的起源与不完全性定理紧密相关（如亨金（Henkin）问题，可导性条件，洛布定理等）。在这种意义上，我们可以说不完全性定理是连接一阶算术和模态逻辑的一种桥梁。算术解释的概念为我们提供了一种建立可证性逻辑和算术的元数学间的关联的重要工具。[33] 令人惊讶的是，可证性逻辑**GL**和**GLS**的索洛维（Solovay）算术完备性定理刻画了"在**PA**中可证"和"在算术的标准模型中为真"这两个概念间的区别。令T 是**Q**的Σ^0_1-可靠的递归公理化的扩充。[34] **GL**的索洛维算术完备性定理告诉我们：对任意的模态公式A，A 在**GL**中可证当且仅当对任意的算术解释f, $f(A)$ 在**PA**中可证。而**GLS**的索洛维算术完备性定理告诉我们：对任意的模态公式A，A 在**GLS**中可证当且仅当对任意的算术解释f, $f(A)$ 在算术的标准模型中为真。

我们知道G2的证明取决于可证谓词的性质。可证性逻辑是关于可证谓词的性质的逻辑。不同的可证谓词可能对应不同的可证性逻辑。可证性逻辑给我们提供了一种理解不完全性的新视角和重要工具。

总之，G1和G2都既可推广到**PA**的扩充理论，也可推广到某些比**PA**弱的理论。这一事实告诉我们：不完全性现象在逻辑和数学中是普遍存在的。不完全性定理这些丰富的推广形式揭示了不完全性定理的解释力量和广泛的可应用性。

6 不完全性定理成立的限度

不完全性定理成立的限度研究不完全性定理在何种情形下成立，在何种情形下不成立。对不完全性定理成立的限度的研究，极大的加深了我们对不完全性定理的可应用性范围的理解。我们先从三个方面说明G1成立的限度。

第一、一理论是否完全与理论的语言有关。在语言$L(0,S), L(0,S,<)$和$L(0,S,<,+)$中分别存在递归公理化的完全理论。[35] 对于G1而言，包含

[33]算术解释是模态公式集到一阶算术语句集的函数。关于算术解释的精确定义，参见[6]。

[34]我们称算术理论T是Σ^0_1-可靠的，若任意在T中可证的Σ^0_1语句在算术的标准模型中都为真。

[35]例如，语言$L(0,S,<)$中的如下理论是完全的，其公理由如下语句组成：$x \neq \mathbf{0} \to \exists y(x = \mathbf{S}(y))$; $x < Sy \leftrightarrow x \leq y$; $x \nless 0$; $x < y \vee x = y \vee y < x$; $x < y \to y \nless x$; $x < y \wedge y < z \to x < z$.

或能解释足够的关于算术的信息是至关重要的。例如，欧式几何是关于点、线、面而非关于算术的理论，塔斯基证明了欧式几何理论是完全的。若一理论仅包含关于算术的加法信息而不包含乘法信息，则此理论有可能是完全的。例如，普雷斯堡(Presburger)算术是语言$L(0, S, <, +)$ 中的完全理论（参见[46, 定理3.2.2]）。

第二、我们知道，G1对$\mathbf{Q}$的某些算术可定义的一致扩充成立；但并非$\mathbf{Q}$的所有算术可定义的一致扩充都是不完全的。关于$\mathbf{Q}$的某些算术可定义的完全扩充的例子，参见[54].

第三、我们已知很多在解释度意义上比$\mathbf{PA}$弱的使得G1成立的理论。一个自然的问题是：是否存在使得G1成立的极小系统？这个问题的答案取决于我们对"极小"的理解。若理论T是极小的是指理论T的公理仅有一条，则存在使得G1成立的极小系统；若理论T是极小的是指理论T仅有一个公理模式，那么也存在使得G1成立的极小系统（参见[14]）。当我们说某理论是极小的，我们需说明此理论是相对于哪种度结构而言是极小的。图灵度和解释度是两种常见的度结构。我们已知对任意非递归的递归可枚举图灵度$\mathbf{d}$，存在递归可枚举的理论T使得G1对T成立且T的图灵度是$\mathbf{d}$（参见[12]）。因此不存在使得G1成立的在图灵度意义上的极小递归可枚举理论。

下面我们讨论G2成立的限度。无论数学上或哲学上，G2与G1是本质上不同的。对于证明G1而言，要证明一理论是不完全的，仅需证明存在一不可判定命题：我们既可构造类似于哥德尔语句之类的纯逻辑构造的不可判定命题，也可构造具有实在数学内容的经典数学中的不可判定命题。而对于G2而言，要证明T的一致性在T中不可证，我们就需恰当地形式化"T是一致的"这个概念。因此，对于G2，我们需关心如何在T中表达T的一致性。

我们称G2对理论T成立，若"T的一致性"在T中不可证。然而，这个定义是模糊的：G2对理论T是否成立取决于我们如何在T中形式化"T的一致性"。我们称这一现象为G2的内涵性问题。在我们可不使用算术化和可证谓词等纯逻辑方法来构造有实在数学内容的经典数学中的不可判定命题的意义上，我们可以说G1是外延性的。然而，G2是内涵性的，与G1有本质的区别。现有文献中有不少关于G2的内涵性问题的讨论（参见[34, 35, 65]）。下面我们基于文献中最新研究成果讨论G2对理论T是否成立取决于哪些因素。

我们称公式$\phi(x)$是理论T的公理集的表示公式，若对任意的自然数n，$\mathbf{PA} \vdash \phi(\overline{n})$ 当且仅当n 是T的某一公理的编码。下文中，若未作特殊说明，我们作出如下假设：(1) 理论T是$\mathbf{Q}$的递归公理化的一致扩充；(2) 我们默认

使用哥德尔编码法；(3) 我们使用的可证谓词是标准可证谓词；(4) 我们使用典范一致性命题$\mathbf{Con}(T) \triangleq \neg\mathbf{Pr}_T(\ulcorner\mathbf{0} \neq \mathbf{0}\urcorner)$表达理论$T$的一致性；[36] (5) 理论$T$的公理集的表示公式是$\Sigma_1^0$公式。我们知道，若理论$T$满足如上假设条件(1)-(5)时，则G2对$T$成立。

理论T公理集的表示公式的复杂性是指此表示公式在算术层级中的复杂性（例如是Σ_1^0公式还是Π_1^0公式）。基于文献中关于G2的研究成果，下面我们说明G2对理论T是否成立取决于如下因素：(A) 理论T的选择；(B) 编码方法的选择；(C) 可证谓词的选择；(D) 表达一致性的方式；(E) 理论T公理集的表示公式的复杂性。

我们对这些因素作两点说明：第一，这些因素间不是完全相互独立的，但它们强调的侧重点有所不同。例如，给定一种表达一致性的模式，选择不同的可证谓词也导致不同的表达一致性的方式。但(D)强调的是表达一致性的整体方式，而不是具体可证谓词的选择。而(C)强调的是具体可证谓词的选择。第二，当我们研究G2是否成立是如何取决于以上某种因素时，我们假定其他的因素以我们如上假设的方式保持不变，而仅仅讨论此因素的变动是如何影响G2是否成立的。以因素(C)为例，当我们说G2对理论T是否成立取决于可证谓词的选择时，我们是说：假设T是$\mathbf{Q}$的递归公理化的一致扩充，使用的是哥德尔编码法，使用典范一致性命题表达理论的一致性，且T的公理集的表示公式是Σ_1^0公式，则使用不同于标准可证谓词的可证谓词可能会导致G2不成立。

下面我们简要地讨论G2是否成立是如何取决于以上五个因素。关于G2的内涵性问题的详细讨论，参见[13]。第一，G2对理论T是否成立取决于理论T的选择。证明G2相对于证明G1而言，需要更多关于算术的信息。若理论T并非包含关于算术的足够信息，G2可能不成立：即T的一致性在T中可证。例如，威拉德（Willard）研究G2对理论T不成立的边界情形，并构造了一些算术理论，它们不能证明后续函数是全函数，但可证明自身的一致性（参见[70, 71]）。

第二，G2对理论T是否成立取决于编码方法的选择。巴尔塔萨(Balthasar)定义了一类可接受的编码方法，并证明了在假定其他因素不变的情况下，对于这类可接受的编码方法，G2是成立的；然而，巴尔塔萨指出，若一种编码方法是不可接受的，则G2可能不成立（参见[3]）。

第三，G2对理论T是否成立取决于可证谓词的选择。维瑟论证道，一

[36] 典范一致性命题使用哥德尔编码和标准可证谓词。

致性命题不是一个绝对的概念，而是与其使用的可证谓词有关（参见[66]）。我们知道，在假定其他因素不变的情况下，G2对标准可证谓词成立。然而，对于非标准可证谓词，G2可能不成立。一种重要的非标准可证谓词是罗瑟可证谓词，其是罗瑟在证明G1时引入的。我们可证明，用罗瑟可证谓词表达的一致性命题，在理论T中是可证的：即若$\mathbf{Pr}_T(x)$ 是罗瑟可证谓词，则其对应的一致性命题$\neg\mathbf{Pr}_T(\ulcorner\mathbf{0}\neq\mathbf{0}\urcorner)$在$T$中可证。

第四，G2对理论T是否成立取决于表达一致性的方式。何谓一个理论的一致性命题？不同学者有不同的看法。例如，阿特莫夫（Artemov）论证道："在希尔伯特纲领中，初始的表达一致性的方式是'不存在一个矛盾式的演绎序列'，这一表达方式是关于公式的有限序列，而与算术化和编码无关"（参见[1]）。在研究文献中，我们常用理论T语言中的算术公式表达T的一致性。然而，即使是用算术语句表达T的一致性，仍存在不同于典范一致性命题的表达一致性的方式使得G2不成立：即在这种表达方式下，其对应的一致性命题在理论T中可证（参见[62]）。

最后，G2对理论T是否成立取决于理论T公理集的算术复杂性。给定理论T公理集的表示公式，我们可定义其对应的可证谓词和一致性命题。我们已知，若理论T公理集的表示公式是Σ^0_1公式，则其对应的一致性命题在T中不可证。若理论T的公理集具有不同复杂性的表示公式，则在此种表示公式下，G2可能不成立，即其对应的一致性命题可能在T中可证。例如，费弗曼构造了理论T公理集的一个复杂性为Π^0_1的表示公式，使得其对应的一致性命题在T中可证，即G2不成立（参见[19]）。

当我们讨论G2的哲学含义时，我们需知道G2成立的范围和限度。以上讨论告诉我们，G2对理论T是否成立至少取决于以上五个因素，其可能还取决于其他因素，如形式化的方法，使用的逻辑（是经典逻辑还是非经典逻辑）等。因此，为了使对G2的哲学含义的讨论变得有意义，我们需要一种绝对版本的G2。一个有趣的问题是：我们能否找到更一般的条件使得若理论T满足这些条件时，则G2对T成立。

7 结论

哥德尔之后对不完全性定理的研究极大地更新和深化了我们对不完全性现象的认识。人们研究发现了使用不同领域中方法的不完全性定理的不同证明，不完全性定理不同的推广形式及其成立的范围和限度。对不完全性定理的深入研究揭示了其在不同领域中的影响力、结论的丰富性和建立

起的与不同领域间的关联：如逻辑与经典数学中不同领域间的关联，不完全性与不可判定性理论间的关联，不完全性与逻辑悖论间的关联，及不完全性与可证性逻辑间的关联等。哥德尔发表不完全性定理时可能自己都没预想到：不完全性定理会对数学基础、哲学、经典数学和理论计算机科学产生如此重要深刻的影响，不完全性定理会有如此多的不同证明和推广形式，第二不完全性定理是否成立会取决于这么多的因素。

关于不完全性定理在不同领域中的影响，费弗曼总结道："不完全性定理在数理逻辑及计算理论中的影响是最重要的；此外，不完全性定理的哲学相关性很重要，但这种相关性的程度是尚不清楚的。最后，不完全性定理在逻辑之外对于数学的相关性不是很明显的，但不完全性定理在数学中的意义和影响是值得持续努力研究的方向（参见[22], p.434）"。

如费弗曼所总结的，不完全性定理在数学基础及理论计算机科学中的影响是明显且巨大的，它对这些领域的影响将会是持久深远的。不完全性定理的哲学意义及其在哲学中的应用一直是哲学研究中的一个热门话题，但学界对不完全性定理的哲学意义及其在哲学中的应用也充满了争议。因此，客观地说，学界对不完全性定理在哲学上的意义及其在哲学中的应用尚未达成共识。不完全性定理在数学中的影响当前主要体现在关于具体不完全性的研究。人们已经发现了很多经典数学中的独立性命题。在我们看来，不完全性定理在数学中的重要性及其对数学的影响程度在很大程度上取决于哈维-弗里德曼关于具体不完全性的研究纲领的实现程度。如果未来我们能揭示不完全性现象在经典数学的几乎每一个领域中都是普遍存在的，那么不完全性定理将对经典数学产生更大的影响。

References

[1] Sergei Artemov. The Provability of Consistency. See arXiv:1902.07404v5, 2019.

[2] Matthias Baaz, Papadimitriou H. Christos, Putnam W. Hilary, Scott S. Dana, and Harper L. Charles (Edited). *Kurt Gödel and the Foundations of Mathematics: Horizons of Truth*. Cambridge University Press, 2014.

[3] Grabmayr Balthasar. On the Invariance of Gödel's Second Theorem with regard to Numberings. *The Review of Symbolic Logic*, Volume 14, Issue 1, March 2021, pp. 51-84, 2020.

[4] Lev D. Beklemishev. Gödel incompleteness theorems and the limits of their applicability I. *Russian Math Surveys*, 2010.

[5] Paul Benacerraf. God, the devil and Gödel. *The Monist* 51, pp. 9-32, 1967.

[6] George Boolos. *The Logic of Provability.* Cambridge University Press, 1993.

[7] Andrey Bovykin. Brief introduction to unprovability. *Logic Colloquium: Lecture Notes in Logic 32*, 2006.

[8] Timothy J. Carlson. Knowledge, machines, and the consistency of Reinhardt's strong mechanistic thesis. *Annals of Pure and Applied Logic*, 105 (1-3):51-82, 2000.

[9] J. David Chalmers. Minds, machines, and mathematics: A review of Shadows of the mind by Roger Penrose. *Journal Psyche*, 2, 1995.

[10] Yong Cheng. *Incompleteness for Higher-Order Arithmetic: An Example Based on Harrington's Principle.* Springer series: Springerbrief in Mathematics, Springer, 2019.

[11] Yong Cheng. Gödel's incompleteness theorem and the Anti-Mechanist Argument: revisited. *Studia Semiotyczne (a special issue titled 'People, Machines and Gödel')*, VOL 34 NO 1, pp. 159-182, 2020.

[12] Yong Cheng. Finding the limit of incompleteness I. *Bulletin of Symbolic Logic*, Volume 26, Issue 3-4, December 2020, pp. 268-286.

[13] Yong Cheng. Current research on Gödel's incompleteness theorems. *Bulletin of Symbolic Logic*, Volume 27, Issue 2, June 2021, pp. 113-167.

[14] Yong Cheng. On the depth of Gödel's incompleteness theorems. *Philosophia Mathematica*, Volume 30, Issue 2, June 2022, Pages 173-199.

[15] Yong Cheng and Ralf Schindler. Harrington's Principle in higher order arithmetic. *The Journal of Symbolic Logic*, Volume 80, Issue 02, pp 477-489, June 2015.

[16] J. Dawson. The Reception of Gödel's Incompleteness Theorems. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1984, vol. II, pp. 253-271, 1985.

[17] Michael Detlefsen. On Interpreting Gödel's Second Theorem. *Journal of Philosophical Logic* 8, 297-313, 1979.

[18] Herbert B. Enderton. *A mathematical introduction to logic* (2nd ed.). Boston, MA: Academic Press, 2001.

[19] Solomon Feferman. Arithmetization of mathematics in a general setting. *Fundamenta Mathematicae*, vol. 49, pp. 35-92, 1960.

[20] Solomon Feferman. Transfinite recursive progressions of axiomatic theories. *Journal of Symbolic Logic*, vol. 27 (1962), pp. 259-316.

[21] Solomon Feferman. Hilbert's Program Relativized: Proof-Theoretical and Foundational Reductions. *The Journal of Symbolic Logic*, Vol. 53, No. 2, pp. 364-384, 1988.

[22] Solomon Feferman. The Impact of the Incompleteness Theorems on Mathematics. *Notices of the AMS*, Volume 53, Number 4, p.434-439, 2006.

[23] Solomon Feferman. Gödel, Nagel, minds, and machines. *The Journal of Philosophy*, CVI(4):201-219, 2009.

[24] Solomon Feferman, John W. Dawson, Warren Goldfarb, Charles Parsons and Robert M. Solovay (Edited). *Kurt Gödel's Collected Works (vol. 1, pp. 145-195)*, Oxford University Press, New York and Oxford, 1995.

[25] Torkel Franzen. *Gödel's Theorem: an incomplete guide to its use and abuse.* A.K.Peters, 2005.

[26] Harvey M. Friedman. My Forty Years On His Shoulders. *Proceddings of the Göedel Centenary meeting in Vienna* 2006, Horizons of Truth, 69 pages. Supercedes earlier versions, 2009.

[27] Harvey M. Friedman. *Boolean Relation Theory and Incompleteness.* To appear in Lecture Notes in Logic, Association for Symbolic Logic.

[28] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatsh. Math. Phys.* 38:1, 173-198, 1931.

[29] Kurt Gödel. On completeness and consistency. In: *Gödel 1986-2003 Vol. I* (235-237), 1932.

[30] Kurt Gödel. On undecidable propositions of formal mathematical systems?. In: *Gödel 1986-2003 Vol. I* (346-371), 1934.

[31] Kurt Gödel. *Kurt Gödel's Collected Works, Vol. 1*, pp. 145-195.

[32] Kurt Gödel. Some basic theorems on the foundations of mathematics and their implications. In *Collected Works, Volume III: Unpublished Essays and Lectures*, pp. 304-323, Oxford University Press, 1951.

[33] Andrzej Grzegorczyk. Undecidability without arithmetization. *Studia Logica*, 79(2):163-230, 2005.

[34] Volker Halbach and Albert Visser. Self-reference in arithmetic I. *Review of Symbolic Logic* 7(4), 671-691, 2014.

[35] Volker Halbach and Albert Visser. Self-Reference in Arithmetic II. *Review of Symbolic Logic* 7(4), 692-712, 2014.

[36] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*, Vol. 1 (1934), Vol II (1939), Berlin: Springer.

[37] Leon Horsten and Philip Welch. *Gödel's Disjunction: The scope and limits of mathematical knowledge.* Oxford University Press, 2016.

[38] Makoto Kikuchi and Taishi Kurahashi. Generalizations of Gödel's incompleteness theorems for Σ_n-definable theories of arithmetic. *Rew. Symb. Logic*, 10(4): 603-616, 2017.

[39] Peter Koellner. On the Question of Whether the Mind Can Be Mechanized Part I: From Gödel to Penrose. *Journal of Philosophy*, Volume 115, Issue 7, Pages 337-360, 2018.

[40] Peter Koellner. On the Question of Whether the Mind can be Mechanized Part II: Penrose's New Argument. *Journal of Philosophy*, Volume 115, Issue 9, Pages 453-484, 2018.

[41] Henryk Kotlarski. The incompleteness theorems after 70 years. *Annals of Pure and Applied Logic*, 126(1-3):125-138, 2004.

[42] Per Lindström. *Aspects of Incompleteness.* Lecture Notes in Logic v. 10, 1997.

[43] Per Lindström. Remarks on Penrose's new argument. *Journal of Philosophical Logic*, 35:231-237, 2006.

[44] J. R. Lucas. Minds, machines, and Gödel: A retrospect. In *Machines and thought: The legacy of Alan Turing, Volume 1* (P. J. R. Millican and A. Clark, editors), Oxford University Press, Oxford, 1996.

[45] P. Mancosu. Between Vienna and Berlin: The Immediate Reception of Gödel's Incompleteness Theorems. *History and Philosophy of Logic*, 20: 33-45, 1999.

[46] Roman Murawski. *Recursive Functions and Metamathematics: Problems of Completeness and Decidability, Gödel's Theorems.* Springer Netherlands, 1999.

[47] Roger Penrose. *The Emperor's New Mind: Concerning Computeres, Minds, and the Laws of Physics.* Oxford University Press, 1989.

[48] Roger Penrose. *Shadows of the Mind: A Search for the Missing Science of Consciousness.* Oxford University Press, 1994.

[49] P. Pudlák. A note on applicability of the incompleteness theorem to human mind. *Annals of Pure and Applied Logic* 96 (1999), pp. 335-342.

[50] Panu Raattkainen. On the philosophical relevance of Gödel's incompleteness theorems. *Revue Internationale de Philosophie* 59 (4):513-534, 2005.

[51] William N. Reinhardt. Absolute versions of incompleteness theorems. *Noûs*, 19(3):317-346, September 1985.

[52] William N. Reinhardt. The consistency of a variant of Church's thesis with an axiomatic theory of an epistemic notion. In Special Volume for the *Proceedings of the 5th Latin American Symposium on Mathematical Logic*, 1981, volume 19 of Revista Colombiana de Matematicas, pages 177-200, 1985.

[53] William N. Reinhardt. Epistemic theories and the interpretation of Gödel's incompleteness theorems. *Journal of Philosophical Logic* 15 (1986), pp. 427-474.

[54] Saeed Salehi and Payam Seraji. Gödel-Rosser's Incompleteness Theorem, generalized and optimized for definable theories. *Journal of Logic and Computation*, Volume 27, Issue 5, Pages 1391-1397, 2017.

[55] Stewart Shapiro. Incompleteness, Mechanism, and Optimism. *The Bulletin of Symbolic Logic*, Vol. 4, No. 3, pp. 273-302, 1998.

[56] Stewart Shapiro. Mechanism, truth, and Penrose's new argument. *Journal of Philosophical Logic*, XXXII(1):19-42, 2003.

[57] Peter Smith. *An Introduction to Gödel's Theorems.* Cambridge University Press, 2007.

[58] C. Smoryński. The Incompleteness Theorems. In: J. Barwise (Ed.), *Handbook of Mathematical Logic*, North-Holland, Amsterdam, pp. 821-865, 1977.

[59] Raymond M. Smullyan. *Gödel's Incompleteness Theorems.* Oxford Logic Guides, Oxford University Press, 1992.

[60] Krajewski Stanislaw. On the Anti-Mechanist Arguments Based on Gödel Theorem. In a special issue of *Semiotic Studies*, 2020.

[61] Simpson G. Stephen. Partial Realizations of Hilbert's Program. *The Journal of Symbolic Logic*, Vol. 53, No. 2, pp. 349-363, 1988.

[62] Kurahashi Taishi. Rosser provability and the second incompleteness theorem. *Symposium on Advances in Mathematical Logic 2018 proceedings*, 2020.

[63] Alfred Tarski, Andrzej Mostowski and Raphael M. Robinson. *Undecidabe theories*. Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam, 1953.

[64] A. Turing. Systems of logic based on ordinals. *Proceedings of the London Mathematical Society*, vol. 45 (1939), pp. 161-228.

[65] Albert Visser. Can we make the second incompleteness theorem coordinate free? *Journal of Logic and Computation* 21(4), 543-560, 2011.

[66] Albert Visser. The Second Incompleteness Theorem: Reflections and Ruminations. In *Gödel's Disjunction: The scope and limits of mathematical knowledge*, edited by Leon Horsten and Philip Welch, Oxford University Press, 2016.

[67] Albert Visser. From Tarski to Gödel: or, how to derive the second incompleteness theorem from the undefinability of truth without self-reference. *Journal of Logic and Computation*, Volume 29, Issue 5, Pages 595-604, 2019.

[68] Hao Wang. *Reflections on Kurt Gödel*. The MIT Press, 1990.

[69] Hao Wang. *A logical journey: from Gödel to philosophy*. The MIT Press, 1997.

[70] D. E. Willard. Self-verifying axiom systems, the incompleteness theorem and related reflection principles. *Journal of Symbolic Logic*, 66(2):536-596, 2001.

[71] D. E. Willard. A generalization of the second incompleteness theorem and some exceptions to it. *Ann. Pure Appl. Logic*, 141(3):472-496, 2006.

[72] Richard Zach. Hilbert's Program Then and Now. *Philosophy of Logic: Handbook of the Philosophy of Science*, Pages 411-447, 2007.

[73] "Gödel's Incompleteness Theorems" in Wikipedia.

[74] "Gödel's Incompleteness Theorems" in Stanford Encyclopedia of Philosophy.