第 15 卷第 6 期 中南大学学报(社会科学版) Vol.15 No6
2009 年 12 月 J. CENT. SOUTH UNIV. (SOCIAL SCIENCE) Dec 2009

哥德尔不完全性定理和“心灵与机器”的关系问题

刘大为，孙明湘

(华南师范大学公共管理学院，广东广州，510006；中南大学哲学系，湖南长沙，410083)

摘要：一些文献在阐述哥德尔不完全性定理的证明过程时，对一些技术细节没有做出明确说明，容易使人误解，因此需要对证明过程中 ω 一致性、系统外证明、元语句可表达性等作出强调。通过系统外证明的启示，分析了由哥德尔定理引起的有关心灵与机器(计算机)关系的争论，得出心灵优于所有目前原理计算机的论点。

关键词：哥德尔不完全性定理；一致性；可证；心灵与机器关系

中图分类号：B81-05 **文献标识码**：A **文章编号**：1672-3104(2009)05−0733−06

1931 年 1 月，库尔特·哥德尔发表了一篇著名的论文《论<数学原理>及有关系统中的形式不可判定命题 I》。其中他证明了以《数学原理》简称 PA 为代表的每个丰富可靠的数学形式系统中，存在着为真且不可证的命题(哥德尔第一定理)；而且 PA 系统的一致性在该系统内是不可证明的(哥德尔第二定理)。

哥德尔的两个结果深刻改变了逻辑学与数学的面貌，尤其影响了第三次数学危机以来关于数学基础的争论与研究，从而开创了现代逻辑与数学发展的新时代。希尔伯特曾计划了以元数学为工具来证明数学系统一致性的方案，而且相信它能成功，但哥德尔 PA 一致性在内部不可证明的结果给希尔伯特的形式主义数学哲学纲领以沉重打击，使他期望以有穷方法来证明算术系统一致性的方案破产。哥德尔不完全性定理真正区分了真与可证，得出了如果一个形式理论足以容纳数论并且是一致的，则它是不完全的，即存在一个数论语句是真的，又是不可证明的。而以前的数学家们，原本期望任何一个真命题都会在某个形式化公理系统内确立起来。这样一来哥德尔定理揭示了形式化公理方法的局限性。不仅如此，哥德尔不完全定理还引发了许多哲学上的反思，如讨论得较多的心灵与机器的关系等等。

许多人都将哥德尔在逻辑学上的地位与亚里士多德相提并论，哥德尔晚年的好朋友著名物理学家爱因斯坦也给予他很高的评价，将他的在数学的成就与自己在物理学的贡献相比。因此，对哥德尔不完全性定理做一个较通俗的介绍还是很有意义的。本文的形式化证明过程主要参考了 A.G.汉密尔顿[1]、朱水林[2]以及郭世铭[3]的文献，包括其中部分的符号系统，但本文着重对原文一些证明过程中容易引起误解、并没有作特别说明的技术细节做了强调和解释。

哥德尔构造了一个自指语句 U 来证明这个定理，为了便于理解，我们先通过直观的自然语言来描述，但直观的自然语言并不严格，仅仅是有助于理解后面形式证明的思路。哥德尔语句 U 的自然语言形式如下：

U：U 是不可证明的

¬U：U 是可以证明的

如果 U 是可以证明的，那么得到 U 是不可证明的，因为 U=U 是不可证明的，矛盾，所以 U 是不可证明的。于是 U 是不可证明的(U)是真的，那么 U 是可以证明的(¬U)是假的，而假命题是不可证明为定理的，所以¬U 也是不可证明的[4](225)。这样就得到了 U 和¬U 都不可证明，而在算术模型中 U 和¬U 必有一真，因此存在着为真且不可证明的命题。

一、哥德尔不完全性定理证明的准备工作

19 世纪 80 年代，皮亚诺所提出了自然算术的一个公理系统，把算术建立在 5 条算术的公理基础之上。而所谓形式算术，就是用一阶逻辑的形式语言陈述皮亚诺公理而得到的形式理论。一般简记为 PA。PA 有以下初始符号：

1.个体变元 x_1，x_2，x_3 …

收稿日期：2009−05−15

作者简介：刘大为(1980-)，男，湖南长沙人，华南师范大学公共管理学院 2007 级科技哲学专业硕士研究生，通信助理工程师，主要研究方向：逻辑哲学与数学哲学.

2.个体常元 0

3.函数符号 S(后继函数)，+(加函数)，×(积函数)

4.等词=

5.逻辑常项¬(非)，→(蕴含)，∀(全称量词)

6.括号(，)

此外，加上对项和合适公式的定义、列出公理(它是一阶逻辑的扩充，还包括相等公理和前述皮亚诺公理的形式化)、初始变形规则，就构成了完整的形式算术系统 PA。其中来刻画算术的七条特有公理如下：

PA1 $(\forall x_1)\neg(S(x_1)=0)$

PA2 $(\forall x_1)(\forall x_2)(S(x_1)=S(x_2)\rightarrow x_1=x_2)$

PA3 $(\forall x_1)(x_1+0=x_1)$

PA4 $(\forall x_1)(\forall x_2)(x_1+S(x_2)=S(x_1+x_2))$

PA5 $(\forall x_1)(x_1\times 0=0)$

PA6 $(\forall x_1)(\forall x_2)(x_1\times S(x_2)=(x_1\times x_2)+x_1)$

PA7 $A(0)\rightarrow((\forall x_1)(A(x_1)\rightarrow A(S(x_1)))\rightarrow(\forall x_1)A(x_1))$

(也就是数学归纳法)

为了避免混淆，后文都用斜体字表示形式算术系统中的合适公式等符号，而且对于自然数 i，用斜体字 *i* 简化表示在形式算术中的表达式(严格的书写是 S(S(S(0)))…，有 i 个 S)。而后文的哥德尔的不完全性定理就是针对形式算术 PA 的。

哥德尔在证明不完全性定理时，首创了哥德尔编码。简单地说，就是对形式系统中的符号、公式、公式序列都以自然数编码。这个做法不仅在哥德尔定理的证明过程中起了关键作用，而且推动了后来的计算机科学技术和人工智能的发展。

哥德尔编码方法现在看来并不很复杂，对形式算术 PA 编码，就是对 PA 中的每个符号、合适公式、合适公式序列(如证明)，都按确定的规则指派一个自然数。这个自然数就称之为哥德尔数。这种指派要求不同的符号、不同的公式、不同的公式序列，都必须对应不同的自然数，因此，指派的具体方法是多种多样的。

哥德尔使用编码的目的，就是在于把形式算术的元数学语句，转化为关于自然数的语句。例如，“公式序列 $A_1,A_2,\ldots,A_r,A$ 为公式 A 在形式算术中的一个证明，”是关于形式算术的一个元数学语句。它断定的是：公式的一个有限序列和一个特殊公式之间存在着一种证明关系。通过哥德尔编码，假设公式序列 $A_1,A_2,\ldots,A_r,A$ 的哥德尔数为 i，特殊公式 A 的哥德尔数为 j，那么可以将上述元数学语句转化为关于自然数 i、j 之间的关系。如果用元数学符号 Proof 去表示时，它的定义是：定义在自然数上的一个 2 元关系 Proof(i,j)成立，当且仅当，j 是形式算术系统中的某个公式的哥德尔数，i 是该公式在系统中的一个证明序列的哥德尔数。这样，通过哥德尔编码，就可以把关于形式算术的元数学语句转化为关于自然数的断定。

递归函数现在在计算机科学技术中，有着非常重要和广泛的使用，它真正的全面研究，也是从哥德尔开始的。所谓递归函数，就是能由 3 个初始函数(零函数、后继函数、射影函数，它们是递归函数)，通过有限次运用 3 个函数运算规则(复合运算、递归运算、最小数运算)而得到的函数。

设 P 为定义在自然数集上的 k 元关系，Q 是自然数的 k 元组集合。关系 P 的特征函数 Cp 定义如下：

$Cp(x_1,\ldots,x_k)=1$，如果 $P(x_1,\ldots,x_k)$为真；

0，如果 $P(x_1,\ldots,x_k)$为假。

集合 Q 的特征函数 Cq 定义如下：

$Cq(x_1,\ldots,x_k)=1$，如果$(x_1,\ldots,x_k)\in Q$；

0，如果$(x_1,\ldots,x_k)$□Q。

我们又得到递归关系和递归集的定义，如果 k 元关系 P 的特征函数 Cp 是递归函数，那么 P 为递归关系。如果 k 元组集合 Q 的特征函 Cq 数是递归函数，那么 Q 为递归集。

为了后文的哥德尔不完全性定理证明的简洁性和直观性，我们接受丘奇论题作为证明依据。丘奇论题：能行可计算函数就是递归函数。

可表达性，是指关于自然数的命题与 PA 公式之间的关系。如果把解释的进程看成是从抽象的形式系统到直观的模型时，那么，表达的进程，是从直观的模型到抽象的形式系统。简单地说，表达是解释的逆过程。具体来看，PA 中的谓词符号、函数符号可以解释为定义在自然数上的关系和函数；现在反过来看，定义在自然数上的关系、函数，能否用 PA 中的公式加以表达，这就是所谓的“是否可表达”问题。下面给出可表达的定义。

设 $P(n_1,\ldots,n_k)$是自然数上的 k 元关系，如果存在恰含 k 个自由变元的 PA 公式 $P(x_1,\ldots x_k)$，使得对任意的自然数 k 元组，

1. 如果 $P(n_1,\ldots,n_k)$为真，那么 PA□$P(n_1,\cdots,n_k)$

2. 如果 $P(n_1,\ldots,n_k)$为假，那么 PA□$\neg P(n_1, \cdots, n_k)$

则称 k 元关系 P 在 PA 中是可表达的，$P(x_1,\ldots x_k)$是表达 $P(n_1,\ldots n_k)$的公式。

设 $f(n_1,\ldots,n_k)$是 k 元自然数函数，如果存在恰含 k+1 个自由变元的 PA 公式 $P(x_1,\ldots,x_{k+1})$使得对任何 k+1 个自然数 $n_1,\ldots,n_{k+1}$，

1.如果 $f(n_1,\ldots,n_k)=n_{k+1}$，那么 PA□$P(n_1,\cdots,n_k, n_{k+1})$

2.而且，PA□$(\exists_1 x_{k+1})P(n_1,\cdots,n_k, x_{k+1})$

则称函数 $f(n_1,\cdots,n_k)$是可表达的，$P(x_1,\cdots,\ x_{k+1})$称为表达 $f(n_1,\cdots,n_k)$的公式。

可以证明，每个递归函数在 PA 中都是可表达的。因为递归函数都是由初始函数通过 3 种函数运算生成的，因此这个证明的过程可以分为两步。首先证明初始函数是可表达的；其次，证明函数运算保持可表达性，即如果运用函数运算前的函数是可表达的，那么运用函数运算后的函数也是可表达的。我们直接给出以下定理：每个递归函数在 PA 中都是可表达的。

最后，我们要证明每个递归关系在 PA 中都是可表达的。由于定义在自然数上的 k 元递归关系 $P(x_1,\cdots,x_k)$的特征函数 $Cp(x_1,\cdots,x_k)$是递归函数，根据前述可表达定理，可由 PA 中的公式 $P(x_1,\cdots,x_k,\ x_{k+1})$加以表达，显然这时，关系 $P(x_1,\cdots,x_k)$在 PA 中可由公式 $P(x_1,\cdots,x_k,1)$表达。因此，得到了推论：

每个递归关系在 PA 中都是可表达的。

二、不完全性定理的证明及其正确理解

前文已介绍完了证明哥德尔不完全性定理的准备工作。这里再将前面所做的工作概括总结一下，其实就是分为两步：第一，通过哥德尔数，把关于形式算术系统 PA 的元数学断定，转化为直观算术中关于自然数的断定；第二，通过可表达性定理，再把直观算术中关于自然数的断定，表达为形式算术中的公式。笔者要提出，一些文献在关于这两步转化的论述中，有含糊和容易令人误解的地方，现把原文摘录如下：“大家可以看到，通过哥德尔码数，确实能把关于一阶算术的断定转换为关于自然数的断定；由递归性作条件，确实能把关于自然数的断定再转换为关于一阶算术的断定。”[2](175)原文使用了两次“关于一阶算术的断定”，需要强调，其实这两个“关于一阶算术的断定”是不同的，第一个是指普通的关于一阶算术的元数学断定；而第二个已经是表达为一阶算术内部的合适公式，从形式上来看，不能再说是一般元语句了。

哥德尔不完全性定理形式化的证明还需要引入ω一致性的概念，ω一致性是比一致性更强的要求(ω一致性蕴含一致性，反之不一定)。PA 系统是ω一致的，如果在系统中不存在任何公式 $P(x_1)$，其中 x_1 自由出现，使得$\neg\forall x_1P(x_1)$(也就是$\exists x_1\neg P(x_1)$)是可证的，而且同时对每个数 n，总有 P(n)是可证的(即 P(0),P(1),P(2),⋯ 都可证)。需要强调的是，有些论者在这里容易犯错，认为对每个数 n，总有 P(n)是可证的，就等价于$\forall x_1P(x_1)$是可证的，这是不正确的。事实上，我们的形式系统并没有任何内部规则使得，对每个数 n，总有 P(n)是可证的，得到$\forall x_1P(x_1)$是可证的[5](80)。反过来也可以这样看，假如对每个数 n，总有 P(n)是可证的，可以得到$\forall x_1P(x_1)$也是可证的，那么根据定义，ω一致性就完全等同于一致性了，这显然不是哥德尔的引入ω一致性的目的，哥德尔在证明第一定理的第二部分时，正是要用到ω一致性比一致性更强的性质。

接着还要引入一个非常重要的二元关系 W(m,n)。定义在自然数上的二元关系 W(m,n)成立，当且仅当，PA 中的某一合适公式 $P(x_1)$(其中 x_1 自由出现)的哥德尔数为 m，P(m)在 PA 中的一个证明的哥德尔数为 n。这里要强调，为了更好地理解这个在证明中非常重要的二元关系 W(m,n)，可以将它更清楚地细化为 Proof(n,Sub(m,m))。其中 Sub(m，m)就表示将数 m 代入哥德尔数为 m 的公式的自由变项后，得到的新的公式的哥德尔数。Proof(i，j)就表示具有哥德尔数 i 的公式序列是哥德尔数 j 的公式的证明。因为 W(m,n)是能行可判定的，根据前述的丘奇论题，它是递归的，所以也是可表达的。也就是说，在 PA 中存在一个带 2 个自由变元的合适公式 $W(x_1,x_2)$，对任意自然数 m,n，有

如果 W(m,n)为真，则 PA□W(m,n)

如果 W(m,n)为假，则 PA□¬W(m,n)

需要注意的是，W(m,n)在 PA 中可用 W(m,n)表达，而 W(m,n)在直观算术模型中的解释是 W(m,n)。

我们来构造自指的哥德尔语句 U，先构造含有一个自由变元的公式 $P(x_1)$，它的具体形式是：

$\forall x_2\neg W(x_1,x_2)$

现假设 $P(x_1)$的哥德尔数为 i，它在 PA 中的相应数项是 i，用 i 代入自由变元 x_1 后，就得到公式：

P(i)：$\forall x_2\neg W(i,x_2)$

这就是我们要构造的哥德尔自指语句 U，对

U：$\forall x_2\neg W(i,x_2)$

应作出如下解释：“当 i 是某个公式 $P(x_1)$的哥德尔数时，对于每个自然数 n，都不是公式 P(i)在 PA 中一个证明的哥德尔数。”我们注意到公式 P(i)就是 U 本身，因此对 U 的解释就是：“对每个自然数 n，n 都不是 U 在 PA 中任一个证明的哥德尔数。”

哥德尔第一定理：如果 PA 是一致的，则 U 在 PA 中不可证；如果 PA 是ω一致的，则¬U 在 PA 中不可证。(1936 年，罗塞尔构造了另一个语句，只要求一致性即可。)

证明：用反证法，设 U 在 PA 中可证，就在 PA 中存在 U 的一个证明序列，令 j 是这个证明序列的哥

德尔数，再令 i 是公式 $P(x_1)$：$\forall x_2 \neg W(x_1,x_2)$的哥德尔数，此时参照定义，W(i,j)为真，由于 W 是递归关系，在 PA 中可表达，因此有

$PA \vdash W(i,j)$

另一方面，由假设 U 是定理，有

$PA \vdash \forall x_2 \neg W(i,x_2)$

应用一阶逻辑的公理有：

$PA \vdash \forall x_2 \neg W(i,x_2) \rightarrow \neg W(i,j)$

两式分离，得到

$PA \vdash \neg W(i,j)$

与 PA 的一致性矛盾，所以 U 不可证。

这里也需要强调的是，有些错误观点认为，上述的整个证明过程不就是证明了 U 不可证明？而 U 的解释正是 U 不可证明，那岂不是就是证明了 U 本身？其实，这里要注意区分系统内的证明和系统外的证明，系统内的证明是依据系统内部的公理和推理规则推导，而这个证明用到了一个系统外的假设，即 PA 系统是一致的，因此上述证明是在系统外证明了 U 为真，与 U 不能在 PA 系统内部证明并不矛盾[6](29)。

然后证明¬U 不可证，由于 U 在 PA 中不可证，也就是对任意自然数 j 来说，都不是 U 在 PA 中任一个证明的哥德尔数，也就是$\forall x_2 \neg W(i,x_2)$在 PA 中不可证，那么 W(i,j)对任意自然数 j 都为假。根据可表达定理，有

$PA \vdash \neg W(i,j)$，对每个自然数 j

再参照 ω 一致性的定义，得到

$\neg \forall x_2 \neg W(i,x_2)$

不可证，也就是¬U 不可证。

这里要强调，有人错误认为由 $PA \vdash \neg W(i,j)$，对每个自然数 j，能推出 $PA \vdash \forall x_2 \neg W(i,x_2)$，于是得到 $PA \vdash U$，与前面 U 不是定理矛盾，所以整个证明有误。其实如上文已经所说过，由 $PA \vdash \neg W(i,j)$，对每个自然数 j，是不能推出 $PA \vdash \forall x2 \neg W(i,x2)$，因为系统内没有一条推导规则能得出这个结论。考察我们的 PA 系统，可以看到只有符合公理 PA7(即数学归纳法)才能得到这一结果，而这里的情况是无法符合 PA7 公理的要求的。事实上，只有我们从外部考察这个系统时，才能看出这个结论，依据 PA 系统内现有的规则是证不出来的[7](291)。

哥德尔第二定理：如果 PA 是一致的，那么 PA 的一致性不能在 PA 内部证明。

证明：根据第一定理证明的第一部分，得到如果 PA 系统是一致的，那么 U 是不可证明的。上述过程在 PA 中形式化，就得到：

$PA \vdash cons(PA) \rightarrow U$

其中 cons(PA)是“PA 是一致的”一个纯形式的表达。假设能够在 PA 系统内部证明它的一致性，就有 $PA \vdash cons(PA)$，应用分离规则，就有 $PA \vdash U$。这与第一定理矛盾，所以要放弃假设，即 PA 的一致性不能在 PA 的内部证明。这里要强调，哥德尔当时并没有给出第二定理的完整证明，只是提出了证明的思路，因为要将第一定理在系统内形式化($PA \vdash cons(PA) \rightarrow U$)，需要满足更强的条件，完整的证明是由希尔伯特和贝奈斯于 1939 年完成的。还应指出，对于能够在 PA 系统内部证明它的一致性($PA \vdash cons(PA)$)，是证明前提做出的一个预设，如不这样，那么第二定理的结论 PA 的一致性不能在 PA 的内部证明也就没有很多意义了。

三、心灵与机器关系问题的探讨

哥德尔不完全性定理引起了许多哲学上的反思和争论，其中讨论得较多的就有心灵与机器(计算机)的关系问题。由于目前心理学、脑科学、神经科学与人工智能等科学发展的局限性，对于心灵与机器关系的分析与讨论不可避免地需要加上一些形而上学的假设，笔者提出的观点仅仅作为一种哲学上的初探，提供一点新的思路与视角。首先需要说明几个前提，在以下的讨论中，目前原理的所有“机器(计算机)和形式系统这两个概念是可以互换的，因为对任何形式系统来说，总有方法设计一台计算机系统地生成这个形式系统的定理，反过来，对任何设计好的计算机使用某种形式语言生成的语句，都相对应一个形式系统可以在它的定理中得到这些语句”。[8](115)此外，心灵是否优于计算机的问题可以转化为几个子问题：“(1)是否大脑和心的功能等同；(2)是否大脑的运作基本上等同于计算机的运作；(3)是否心的活动都是可计算的。第一个问题即心脑同一论或称物理主义是否正确的问题，第二个问题是大脑算法主义或大脑可计算主义是否成立的问题，第三个问题是心的算法主义或称心的可计算主义正确与否的问题。”[9](105)而能行可计算函数等价于图灵可计算函数，即在图灵机上可计算的函数。

笔者并没有完全否认未来超出形式系统的机器有可能达到心灵的能力，只是否认目前图灵机似的、纯形式系统原理的计算机能够达到心灵的水准。主要论证过程如下：同许多科学家的观点一样，笔者首先接受物理主义的观点(指心灵、意识可以还原成大脑的神经活动)。由于心灵具有的直觉和经验，可以不断引入新的可接受的方法(包括非有穷方法)，在形式系统外

“看出”并证明在系统内无法证明的语句，而机器无法做到这一点，因此心灵优于目前原理的形式系统的机器。此外，根据生物进化论的观点来看，大脑运作也不是对应于一台计算机，因为从简单生物进化到人类，神经系统的底层已经发生了不可想象的根本改变，这一点尤其体现在大脑和神经系统与外部环境的交互，并可随环境发生底层的进化和改变。而目前原理的计算机底层硬件都是固定不变的。最后提出，未来要达到心灵水准的的机器，一个必要条件是如同大脑一样，具有形式系统外的能力，底层硬件指令必须与环境交互，随环境发生改变和进化以适应环境。因此笔者对上述三个问题的回答是：接受物理主义和心脑同一论，心灵不等价于一台计算机，大脑的运作也不等同于计算机。

前人关于心灵与机器的关系主要有两种对立的观点，简要地说，一种是认为哥德尔不完全性定理支持了心灵胜过机器；而另一种相反的观点则认为哥德尔定理并未支持了这一点。

包括内格尔等人的观点认为:“今天的演算机械内建固定的一组指令，这些指令对应于形式化公理系统的固定的一组演绎规则。这种机器通过一步步操作的方式为问题提供答案，其中每一步都由内建的指令来控制。但是，正如哥德尔在他的不完全性定理中所表明的，初等数论中有无穷多的问题超出了固定的公理方法的能力范围，不管按其原理造的机器有多复杂多精妙，运算多么快，它都不具备回答所有这些问题的能力。…大脑所具有的操作规则结构也要比现在所理解的人工机器强大得多。当前还看不出用机器人来替代人类大脑的直接前景。”[10](86)简要的说，这种观点认为，因为计算机要遵循一组固定的指令，即对应着固定的一组公理和推理规则，所以原则上无法在推理方面做到和人类推理一样灵活机动。

侯世达对此做了反驳,“在人工智能研究中正在开发的很多程序与那些生成数论真理的程序——一些具有不变的推导规则和固定的公理集合的程序——没有多少共同之处，而且它们当然是有意做成“心智模型”的。在它们的顶层——“非形式”层次——可能有一些意象的操作、类比的形式、念头的遗忘、概念的混淆以及差别的抹杀等等。不过，这并不与下述事实相矛盾：它们就像大脑依赖其神经原的正确行动一样，依赖于支持它们的硬件的正确行动。所以，人工智能程序依然是‘形式系统的具体实例’”[7](759)。“人的思想，包括它的灵活性和易错性，在原则上讲是可以用“一组固定的指令”模型化的，前提是人们要从那种认为按照算术运算而建造的计算机，除了奴隶般地产生真命题---全部真命题而且只有真命题外，什么也做不了的先入之见中解放出来。…只要深刻地吸取这个教训，我们就绝对能使“一组固定的指令”具有上面提到的(心灵)一些性质”[10](序言 9)。简要的说，侯世达认为，计算机的底层算术硬件是“一组固定的指令”，即形式化的，但利用计算机软件可以在高层实现许多模拟化的实体,包括有易错性和创造性的智能程序等。

笔者结合了两种观点，认为目前原理的计算机原则上也完全不能达到人类推理的能力，但所依据的论据在内格尔等人的基础上更进一步。笔者赞同侯世达的这一部分论据，即现在的计算机底层硬件是完全形式化的，但利用软件可以在高层实现表现为非形式化的智能程序，事实上许多程序已经做到了这一点。但笔者认为这一事实并不能论证通过这种程序，可以达到心灵的水准，最多只相当于一个好的摆设而已。我们知道人之所以是智慧生命，其中很重要的一点就是在于与环境的交互，包括交换物质、能量和信息等等，而且环境还可以分为自然环境和社会环境。也就是说人与自然、人与社会都是一个更大的生态圈，进行全面的交互与融合，也只有从这个视角看，才体现人是智慧生命。而在这个过程中，包括人从婴儿长到成人，不断地进行学习和适应环境，大脑神经系统底层的神经元也在不断地进行成长和改变。更放宽地看，根据生物进化论，由单细胞生物进化到人类，其底层神经系统更是发生了不可预料的更本改变。这一切都是与环境相互作用和自然选择的结果。因此，底层的指令关键不仅仅在于数量的多少,更在于是否与环境交互，随环境发生变化和升级。而我们目前原理的计算机底层硬件指令都是固定不变的(如对应于形式系统固定的公理与推理规则)，由哥德尔定理得到它对哥德尔语句是无法证明的。而大脑与环境交互的灵活性使其运作不对应于一台计算机,具有超出形式系统外的能力，再根据物理主义的观点，就可以得到心灵同样是优于计算机的。前面内格尔等人认为因为计算机要遵循一组固定的指令，所以原则上无法在推理方面做到和人类推理一样灵活机动，结论正确，但没有更深入，还缺乏较强的说服力，侯世达已经采用高层软件来驳斥了这个论点。但他们都忽视了底层与环境全面交互和改变这一重要因素。

此外，卢卡斯(Lucas)做了另一个论证，同样认为心灵优于任何机器。他说道,“无论我们构造出多么复杂的机器，只要它是机器，就都对应于一个形式系统。接着就能找到一个在该系统内不可证的公式而使之受

到哥德尔过程的打击。机器不能保持着真理性地把这个公式产生出来，尽管人类心智会看出它是真的。…幸亏哥德尔定理，永远是心智说最后一句话。”[7](625)

包括 T.Franzén 等人反对卢卡斯的观点认为，他有一个主要前提是，对于形式系统来说，哥德尔不完全性定理确定了一个公式，系统不可证它，而心灵却可以看出它为真。但是哥德尔定理只告诉我们，如果系统是一致的，那么 U 为真(U 在系统中不可证)，而并没有直接证明 U 为真，因而心灵也不知道 U 为真。而在哥德尔第二定理的证明过程中，我们看到在 PA 内部可证明 PA□cons(PA)→U，也就是说 PA 也能够得到同样的结论。因此，在这一点上，心灵并未超过机器[8](117)。

笔者也赞同反对卢卡斯观点的论据，即哥德尔不完全性定理并没有直接证明 U 为真，而是说如果 PA 是一致的，那么 U 为真，而且根据哥德尔第二定理，PA 本身也同样能得到这一结论 PA□cons(PA)→U。但笔者不同意心灵并未超过目前的机器这一论点。从哥德尔定理的证明受到启发，必须分清系统外证明与系统内证明，因为第二定理已经证明了 PA 不可能从内部证明自身的一致性，但如果心灵能从系统外知道 PA 是一致的，那么就能得出 U 是为真且在 PA 中不可证，可见，系统外这一概念是至关重要的。1936 年，希尔伯特学派的根茨，在放宽了元数学中对证明方法的限制后，用超限归纳法证明了 PA 的一致性，从而可以得到了哥德尔语句为真。虽然我们目前不能得出心灵能够从外部知道任何一个形式系统的一致性和哥德尔语句为真，但人们可以根据自身的需要，加上直觉和经验，不断拓展证明的要求(包括非有穷的方法)，最终达到自己的目的。即使退后一步讲，至少在 PA 的形式系统内，心灵是优于机器的，因为学界基本认可了证明 PA 的一致性。而对于数学基础的直觉主义学派来说，更是相信直觉能够看出数学系统的一致性。因此，心灵是有形式系统外部分的，可以包括直觉、经验等等目前人工智能远远未涉及的领域。依靠这些直觉、经验等，心灵完全可以在此基础上找到一种学界公认接受的方式从外部得知某个形式系统是一致的，而目前的机器根本做不到这一点，因为要将直觉、经验完全形式化，已经超出了当前计算机的能力。包括我们已经沿用了上千年的欧几里得公理系统，也不能在系统内部证明它的一致性，但直觉和经验有理由使我们相信它是一致的。事实上，数学家们已沿用了它很多年了，而它对在现实世界的实际成功应用更是不用多说了。简要地说，心灵用直觉、经验等系统外的方法得知某个形式系统的一致性是客观存在的，而这种能力是以目前原理设计的计算机完全不具备的。

总之，笔者认为以目前原理设计的计算机从理论上也完全达不到心灵的能力。如前所述，很多论者都忽略了计算机与环境的交互，这里绝不是指简单地提高目前计算机的输入输出能力，而是指与整个环境完全融合。我们知道目前计算机的大脑 CPU 的指令都是一组固定的硬件指令，未来如果想要达到心灵水准的计算机，其中一个必要条件是如同人类大脑一样，计算机的底层硬件能随环境改变和进化，包括计算机的 CPU 必须与环境交互，也就是说 CPU 的指令等也可以随环境变化而变化，自动升级并且自主发生进化等等。这样的计算机就不再是目前原理的计算机了，也不对应于一个严格的形式系统了，环境的许多无穷的随机因素都可能干扰计算机内部的形式系统，而计算机本身也可以升级以适应和改变环境。只有整个计算机的所有部分都与环境全开放式地交互和互相作用，才能出现许多形式系统外的不可预知因素，才可能从理论上拥有真正的系统外能力，到了那时候，实现具备心灵能力的计算机才不是梦想。

参考文献：

[1] A.G.汉密尔顿, 骆如枫等译.数学家的逻辑[M]. 北京: 商务印书馆, 1989

[2] 朱水林. 哥德尔不完全性定理[M]. 沈阳: 辽宁教育出版社, 1987.

[3] 张清宇, 郭世铭, 李小五. 哲学逻辑研究[M]. 北京: 社会科学文献出版社, 1997.

[4] S·C·克林. 元数学导论[M]. 北京: 科学出版社, 1984.

[5] 张家龙. 公理学、元数学与哲学[M]. 上海: 上海人民出版社, 1983.

[6] 陈慕泽. 正确理解哥德尔不完全性定理[J]. 湖南科技大学学报, 2008, 11(2): 27−30.

[7] 侯世达. 哥德尔、艾舍尔、巴赫——集异璧之大成[M]. 北京: 商务印书馆, 1996.

[8] Franzén·Gödel’s Theorem:An Incomplete Guide to Its Use and Abuse[M]. Wellesley: A.K.Peters, 2005.

[9] 刘晓力. 理性的生命——哥德尔思想研究[M]. 长沙: 湖南教育出版社, 2000.

[10] 欧内斯特·内格尔, 詹姆士·R·纽曼. 哥德尔证明[M]. 北京: 中国人民大学出版社, 2008.

[11] 邢滔滔. 哥德尔定理正反观[J]. 科学文化评论, 2008, 5(2): 86−108.

(下转 817 页)