人或電腦會有靈魂嗎？

II 論心靈程式主義及可行性問題

● 王　浩

一　心靈程式主義和哥德爾定理

電腦會思考嗎？這種問法顯然太含糊了。「思考」包括的東西太多，所以，要使問題有確定意義，或者要（至少在起始階段）認真研究這問題中某些確定部分，看來必須先選出某些有代表性的思考方式來討論。這已經有不少嘗試，例如圖林探究電腦能否與人類或其他電腦作智性對話的「模擬遊戲」，便是一個有名的例子。另一個例子是問：電子計算機能計算的數學，是否跟人類一樣多？

如果能找到一些數學工作，例如證明定理、發現公設、設定猜想等等，是人類做得到而電腦卻做不到的，便可以下結論說，在數學上，心靈的確有某些方面比電腦優越；而且，可以進一步推斷，心靈和電腦的能力一般是不對等的。這時我們還可以明確地回答原先的問題說：電腦不會思考，即電腦不能在所有方面都像人那樣思考。

甲　波斯特和彭羅斯的嘗試

根據哥德爾定理，任何有相當豐富內涵的形式系統、理論證明機器、或者程序，都有其所無法處理的命題，以及較其能證明更多理論的其他系統或編碼程序。在1921年，波斯特(Post)設想過大致相近的理論，並據而推斷人類心靈比電子計算機優越①：

> 這樣，數學家就遠不止僅僅比機器更靈巧，能更快地做到機器至終可以做到的事。我們看到，機器永不可能提出完備的邏輯，因為機器一旦製成，

我們總能證明一個它不會證明的定理。

然而，經過反思，波斯特不久就修正了這草率的推論：

「人不是機器」這結論不能成立。我們能説的只是，人無法製造出一部能作出人類所有思考的機器。要説明這點，我們可以想像製造一部能夠證明相類於其自身心理運作的定律的「人—機器」複合體。

波斯特的第一個看法在1961年被盧卡斯(John Lucas)重新提出，但卻被其他人據波斯特第二個看法加以批評。最近，彭羅斯(Penrose)進一步發揚盧卡斯式的論證，由是在許多人之間引起爭論②。由於彭羅斯和他的批評者似乎無法溝通，看來似乎有必要研究一下彭羅斯對哥德爾定理出奇強力的運用——特別是其較確定部分。

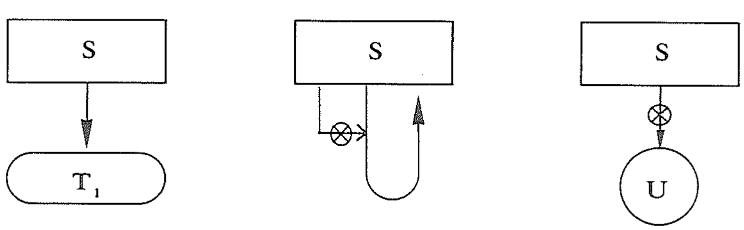
用彭羅斯自己的話説，除了哥德爾定理之外，他又應用了「由『內省』所得的非決定性和軼事式論證來支持我的説法」，但他又説：「我從不認為證明……可以靠這類軼事式論證」：

哥德爾定理的作用是關鍵性的，……只有利用它……才能以類乎數學的嚴謹性證明，我們的理解一定不是程式的。

然而，那麼多人認為他的論證不可信，他的證明是否真具有「類乎數學的嚴謹性」，就大大值得懷疑了。

哥德爾定理

是維也納哲學家和邏輯學家哥德爾(Kurt Gödel, 1906–78)於1931年發現的。它的大意是數學在任何形式系統(formal system)或計算機程式中都有「不可窮盡性」，因為定下了系統公設和求證方法之後，總是可以建構在直覺上為真但在這系統中不可能推導出來的命題。特別是：「該系統具有自洽性(self-consistency)」這一命題(假定為真)便是可表達而不可證明的。粗略地用圖象表示，假如下圖(甲)表明形式系統 S(以長方塊代表)能證明(以箭頭代表)數學定理 T_1(以長橢圓代表)，則(乙)表明 S 不能自證自洽性(自洽性以自指箭頭代表，「不能證明」以 —⊗→ 代表)，(丙)表明 S 不能窮盡宇宙中所有命題 U(以圓代表)。



（甲）　（乙）　（丙）

關於哥德爾可參閱王浩〈探索永恆：哥德爾和愛因斯坦〉，《二十一世紀》第2期，頁71 (1990)。

而且，彭羅斯又利用出處不明的資料，宣稱哥德爾同意他的說法：

> 正如我將會說明，我的論證實在並沒有問題——儘管，坦白地說，它大部分都是「舊」的，肯定在30年代哥德爾本人早已想到過，並且此後從未被真正否定過。

這並不正確，雖然很自然地，哥德爾想用自己的定理作為「心靈優於電腦」證明的一個要素。

乙 定理證明機的存在與證明問題

事實上，在1972年，哥德爾曾經直接和詳盡地解釋清楚下列大家熟知的共識：即單憑他的理論不能得出所需結論③：

> 另一方面，根據現有證明，仍然可能有(甚或真可以發現)實際上相當於數學直覺(即心靈的數學能力)的定理證明機器，但這卻不可能證實；我們甚至不可能證明，這樣一部機器在證明有限數論定理時一定產生正確結果。

換言之，哥德爾定理並沒有排除製造出實際上相當於數學心靈的機器M的可能性。但若真有這麼一部機器，哥德爾用他的定理推斷出兩個結果：

(A)「不可能證明 M 真的有此能力」；

(B)「也不可能證明它只產生正確定理」。

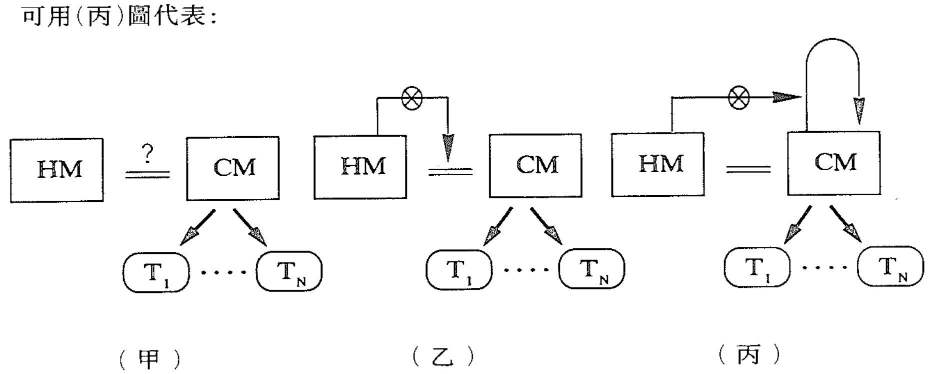
在我看來，(B)很直接，但(A)則較為複雜。

假設我們可以證明M只產生正確定理。這樣的話，我們應該可以證明M的自洽性：但根據關於M的假設(即它的能力相當於數學心靈)，M就應該能夠證明它自己的自洽性。但是，這恰恰違反哥德爾定理。因此，(B)必然成立。

再假設我們可以證明M真具有心靈的數學能力。同時，有理由假定所謂「能力」是指做得「正確」，不會做錯；而所謂「證明」，則可以假定是「數學地證

定理證明機

若以HM代表人類的數學心靈，CM代表定理證明機器，等號代表相當關係，箭號代表(某種關係的)證明，則定理證明機的存在與否可用以下(甲)圖代表：文中結果(A)可用(乙)圖代表：結果(B)(相當於不能證明機器的自洽性)可用(丙)圖代表：



明」。在這兩個假定下，本段開頭的假設就等同於：我們可以數學地證明，M只產生正確定理——但這與(B)矛盾。因為我們已用哥德爾定理證明(B)，所以(A)也成立。

我相信，這是哥德爾心目中(A)證明方法的正確重構。他所說的是：他的理論本身並不否定可能找到M，他同時提供了M具有所需性質的可能理由。(A)所否定的是一種(M的確具有所需性質的)強證明。因此(A)相對於(B)是(他的定理的)一個較弱的推論結果，因為(A)所否定的結論比(B)強。

丙　心靈優於機器嗎？

假定整個宇宙的運作像一台電子計算機(就叫它做U)。那麼人類心靈所能做的所有數學，是U的運作的一部分。根據哥德爾定理和上述結果(A)的證明，即使我們知道了U的程序P，也不能證明P包括我們的所有數學能力。因此，我們不能證明P真的是宇宙的程序。更有甚者，在這樣的宇宙內，雖然我們看不到「隱藏」的計算結構，但並沒有任何精神或物質過程是真正「不可計算」(以圖林的意義論)的。

從哥德爾定理直接推斷心靈比電腦優越，自然很有吸引力。若要證明心靈和電腦對等，那麼，為了讓電腦具有心靈的數學能力，便需要一部能夠滿足哥德爾定理條件的強力定理證明電腦，但根據同一定理，這部電腦卻不能證明自身的自洽性。然而，對這部電腦來說，自洽性的命題自然是應該成立，也因此應該可以證明的。從這一點看，自不免覺得心靈比電腦優越了。不過，根據同一定理，心靈同樣不可能證明自身的自洽性。用哥德爾自己的話說：

> 人類心靈不能夠把它所有的數學直覺公式化或機械化，亦即，它如果能把某些直覺公式化，這事實本身就會產生新的直覺性知識，例如這公式化過程的自洽性。這可稱為數學的「不可完全性」(incompletability)。

1951年，哥德爾主講第二十五屆吉斯(Gibbs)講座，題為「數學基礎的某些基本定理及其含意」，並發表為文。演講直接涉及心靈與電腦的部份相當簡略，在長稿中只佔第9至12頁。它非闡釋性部分其實主要是為數學中的柏拉圖主義辯解。

在這演講中，哥德爾從他的定理所得到(關於心靈與電腦)的主要結論是一個非排斥性的選言判斷(nonexclusive disjunction)：要嗎(在做數論時)心靈比所有機器優越，要嗎就會有心靈沒法給出結論的數論問題。他在1972年告訴我這個結論，並且認為，基於「理性的樂觀主義」，可以排除上述選言判斷的第二個可能④。

二　複雜性和可行性

由於電腦的急速發展和廣泛應用，複雜性已成為今日的熱門題目。到目前

為止，我只是比較電腦和心靈原則上能做甚麼。從這個角度看，我們相信電腦會做的一切，心靈都能做。然而，有些事情是心靈不方便，甚至實際上不可能做的，那就要用到電腦。對心靈和對電腦，「實際可行」和「理論可能」之間都有差別。

甲 哥德爾的計算時間假想

可行性是相對於心靈和電腦發展階段而言的。由於近數十年來電腦顯然進步迅速，這種相對性顯然傾向於電腦這邊。如果有方法M和多項式p，使每一個(表達式)長度為m的問題，其答案可以用方法M以少於p(m)個步驟求得，則這問題集稱為「可在多項式時間計算」；如果多項式是一次或二次，則稱為可在線性或平方時間計算。問題集可在多項式時間計算的話，則稱為屬 P。若有(非決定性的)方法M和多項式p，使問題集內每一長度為m的問題，都可以在「題解樹」(包括所有以M找到的可能題解)內找到短於p(m)的題解徑，則這個集屬NP。問題集S稱為「NP完全」(NP complete)假如：

(1) S屬NP，同時

(2) NP內每一問題集都能「多項式地化約到 S」，即能以S問題的題解為基礎，在多項式時間內求解。

在一封1956年3月20日寫給奈曼(John von Neumann)的信裏，哥德爾提出(用今日的語言說)一個「NP完全」的問題，並推測這問題可以在線性或平方時間內求解⑤：

> 顯然我們很容易構造一部圖林機(Turing machine)，它能就每個一階謂語邏輯(first-order predicate logic)公式 F 與自然數 n，判斷F是否有長度為n(長度=符號數目)的證明。設 h(F,n)是機器作判斷所需的步驟數目，同時設 $g(n)=\mathrm{Max}_F h(F,n)$。問題是：就理想機器而言，g(n)隨 n上升的速度如何。我們可以證明 $g(n) \geq Kn$。如果有一部機器可以使 $g(n) \sim Kn$(甚至只是$\sim Kn^2$)，那將會有了不得的重大後果。具體說，它顯然等於：(用機器)雖不能解決可決定性問題，但就「是」「否」問題而言，則機器可以完全取代數學家的智力工作(建構公設除外)。只要 n 選得夠大，那麼機器做不出結果，就不必再去考慮該問題。然而，我現在看來，g(n)只緩慢增加是完全有可能的。……在其他有限問題中(例如在重複應用倒易律計算平方餘數時)，也經常碰到這種力量的遞減。了解這類問題，例如，在決定某數是否質數時如何；或者在一般組合問題中，這與無遺嘗試比較，可以節省多少步驟，那將會是很有意思的。

(在上文 g 的定義中，運算符號 Max 指給定 n 時，用適當的 F 以使 h(F,n)的值為最大。這並不要求對每一公式 F(長度為 n 或更少，因為顯然較長公式不可能有長度 n 的證明)計算 h 值。但只用 h(F,n)而避免用g(n)可能更簡單。輸入數 n 假定用單元系符號，所以長度就是 n——譬如說以 n 個「1」符構成的字串的形式出現。)

乙 假想背後的意念

哥德爾的基本想法是這樣的：一個命題如果並沒有合理地簡短的證明，那麼證明的嘗試只是徒勞。假如有快捷方法去預先判斷有沒有簡短證明（所謂「簡短」的意義，很容易按情況決定），則電腦可以在兩個重要方面幫數學家的忙，即（在沒有簡短證明時）預先排除徒勞的嘗試，和（在有簡短證明時）提供困難層次的上限。可惜，我們並不能證明 g(n) 只會緩慢增長或者它是 n 的多項式。其實，多數人認為 g(n) 並不受限於任何多項式。

如果哥德爾的推想正確，那麼由（數目有限的）給定公設來證明定理（譬如說公式 F ）便成為另一種問題。就給定的F來說，如果知道有一個證明它的方法短於 n，則不管有沒有從給定公設得出這證明，我們已經知道 F 可證，因此為真。不止如此，只要證明「F 的證明存在」，已可以視為是 F 的（另一種）證明，同時，F 的證明既然存在，把這證明直接寫出來便只是例行公事而已。

但哥德爾的推想其實有些令人迷惑。為要確定 F 有長度為 n（或小於 n）的證明，似乎得去考察所有長度為 n（或小於n）的證明。認為這問題比實際尋求長度為 n（或小於 n）的 F 的證明更容易，是很奇怪的——我覺得他推想的有希爾伯特推想(Hilbert's conjecture)的味道，即可能替強系統(strong system)找到有限而自洽的證明，而那是哥德爾在1931年否證的。

即使哥德爾的推想正確，「建構公設」的工作，如他所指出，也不能指派給電腦。如所周知，哥德爾對新公設極感興趣。也許他的推想可視為指向這樣一個方向：把心靈從例行運算釋放出來，使它集中於較不明確（因此較高級）的工作，例如找尋新公設。無論如何，我們自不必認為他的推想目標在於證明心靈（就數學能力而言）不比電腦優越。

丙 可行性不完全的假想

1976年6月5日，哥德爾告訴我一個心靈力量大於電腦的假想。當時我看不出這個推想有甚麼說服力。以下是他說法不太精確的重構⑥：

> 去證明判斷相對地短的命題的最短決定程序需要很長時間，將是很有意義的。具體地說，有可能證明：對每一個「可決定系統」(decidable system)和它每一個決定程序，總有某種長度小於200但其最短證明大於10^{20}的公式。這樣的結果會表明機器不能代替人的心靈，因為後者能想出新方法來做短證明。

這假想可以視為一種可行性不完全定理(feasibility incompleteness theorm)。要證明這樣一個想法，問題首先是，把它直接和清楚地以適當詞語表達出來。例如，肯定需要列出「可決定系統」須滿足甚麼條件，上列推想才為真。此外，上述舉例說法用了200和10^{20}這兩個數字，而這可能會引出有趣的反證，因此應該用其他直接或間接的限值。無論如何，我對這個推想還不曾想得很仔細，因此無從加以評價，或看出甚麼明顯的反對理由。

註釋

① 1941年波斯特寫了*The Undecidable* (Martin Davis, ed., Raven Press, 1965), 內收"Absolutely Unsolvable Problems and Relatively Undecidable Propositions—Account for an Anticipation", 其中即包括了他由1921年起的部分筆記。以下兩段引文出自該書頁417和423。

② 盧卡斯第一篇公開發表的論文是"Minds, Machines, and Gödel", *Philosophy* **36**, 112 (1990)。作者在其本人的*From Philosophy to Mathematics* (1974), pp. 315–26對這個問題有進一步詳細討論。本文大部分有關哥德爾的斷言均引自此書pp. 324–26。*Behavior and Brain Science* **13** (1990)是Roger Penrose: *The Emperor's New Mind* (Penguin, 1991)一書的專題討論，其中包括彭羅斯自己替該書寫的撮要、37個專家的評論和彭羅斯的答辯。彭羅斯以哥德爾定理來直接證明心靈比電腦優越，討論便集中在這一點上。本文以下兩段彭羅斯的引文引自該書頁699和693。

③ 作者的*From Philosophy to Mathematics*在324頁提及哥德爾這個說法，並經哥德爾本人首肯。

④ 同③，頁324、325。

⑤ 雷賓(Michael Rabin)最近告訴我，哥德爾曾在1973年10月向他提及關於長度 n 的證明這個問題，與及他在40年代給紐曼的信(但我可能記錯了發信日期)。庫克(S.A. Cook)曾在下列文章應用(一個固定系統的)公設性集論討論類似問題：*Journal of Symbolic Logic* **44**, 36–37 (1979); "Can Computers Routinely Discover Mathematical Proof?", *Proceeding of the American Philosophical Society* **128**, 41 (1984)。他提醒作者注意這些文章，並澄清若干技術問題，謹在此表示謝意。

⑥ 作者的*Reflections on Kurt Gödel* (MIT Press, 1987), p. 197錯誤引用了這一段(把「對每一個可決定系統」改成「有某些可決定理論，以致」)，並問：為甚麼它不可能吸收新觀念，發展新的決定程序？首先，正如庫克對我指出的，弱推測的正確性看來已得到確認——例如，通過費希爾(M.J. Fischer)和雷賓他們對普雷斯伯格(Presburger)算術(只是自然數加法)在下列文章的論證來確認：*SIAM–AMS Proceedings* **7**, 27–41 (1974) (R.M. Karp, ed.)。此外，昂加爾(A.M. Ungar)對我的問題的評論也是正確的：「當然新觀念可以被吸收並發展出新的決定程序，但新程序在決定其他短命題所費的時間仍然會是長得驚人。」

羅奇 譯

本文的翻譯及說明方塊未經王浩教授過目，如有錯誤由本刊編輯室負責。

王 浩 王浩是當代卓越的數理邏輯學家及哲學家，在50年代即開始探究利用電子計算機證明邏輯命題這個嶄新的領域，並作出許多開拓性的貢獻，因此除榮膺英、美兩國國家科學院院士外，復在1983年獲得第一屆「米斯東(Milestone)自動化定理證明獎」的殊榮。王教授1921年在濟南出生，在西南聯大和哈佛大學攻讀數學和哲學，1967年迄今在紐約洛克菲勒大學擔任邏輯學教授。王教授著作等身，除百餘篇專業論文之外，還著有《邏輯、電算機和集》、《從數學到哲學》、《分析哲學之外》等六本專書。