The Clean Paper

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

一個密碼學證明可以通過讓缺失的模擬器難以被證明來隱藏其秘密

零知識證明允許某人證明一個陳述而不揭示見證。經典理論說，在完整意義上，你無法用一條消息、無需可信設置和完美可靠性來做到這一點。*Rahul Ilango* 的論文沒有使這種不可能性消失。它改變了目標：不再要求模擬器確實存在，而是要求沒有選定的證明系統能夠高效地證明模擬器不存在。在重要的證明複雜性和密碼學假設下，這足以逐屬性地恢復零知識的可證偽、博弈式後果。重點不是一個即插即用的互聯網原語。它是將數學不可證明性轉化為密碼學掩護的一種證明論方法。

Written by Cinzia Vaglio · Cross-checked by Vera Rubin · Edited by Michele Renda · Figures by Nova Vela · AI-assisted, human-reviewed

Paper: *Gödel in Cryptography: Effectively Zero-Knowledge Proofs for NP with No Interaction, No Setup, and Perfect Soundness*, Rahul Ilango, FOCS 2025 / IACR ePrint 2025/1296 · DOI: 10.1109/FOCS63196.2025.00058

關鍵不在於證明秘密已被隱藏

先從零知識最簡單的版本開始。

Alice 想讓 Bob 相信一個數獨有解。如果她把解答傳給 Bob，Bob 會相信，但數獨也就被破壞了。她想要的是更奇怪的東西：證明解答存在，卻不透露解答本身。

這就是零知識證明的承諾：證明者（Alice）讓驗證者（Bob）相信某個陳述為真，除了這個陳述確實為真的事實之外，什麼也不透露。

問題在於，這個承諾是有代價的。一般數學證明有兩個令人安心的特點。第一，它是**單一訊息**：你把證明寫下來、交出去，然後離開。第二，它是**完美健全的**：虛假陳述根本沒有任何有效證明。經典的不可能性結果指出，零知識必須放棄這兩項特點——而且不只是兩者不能同時保留；每一項單獨都不行。

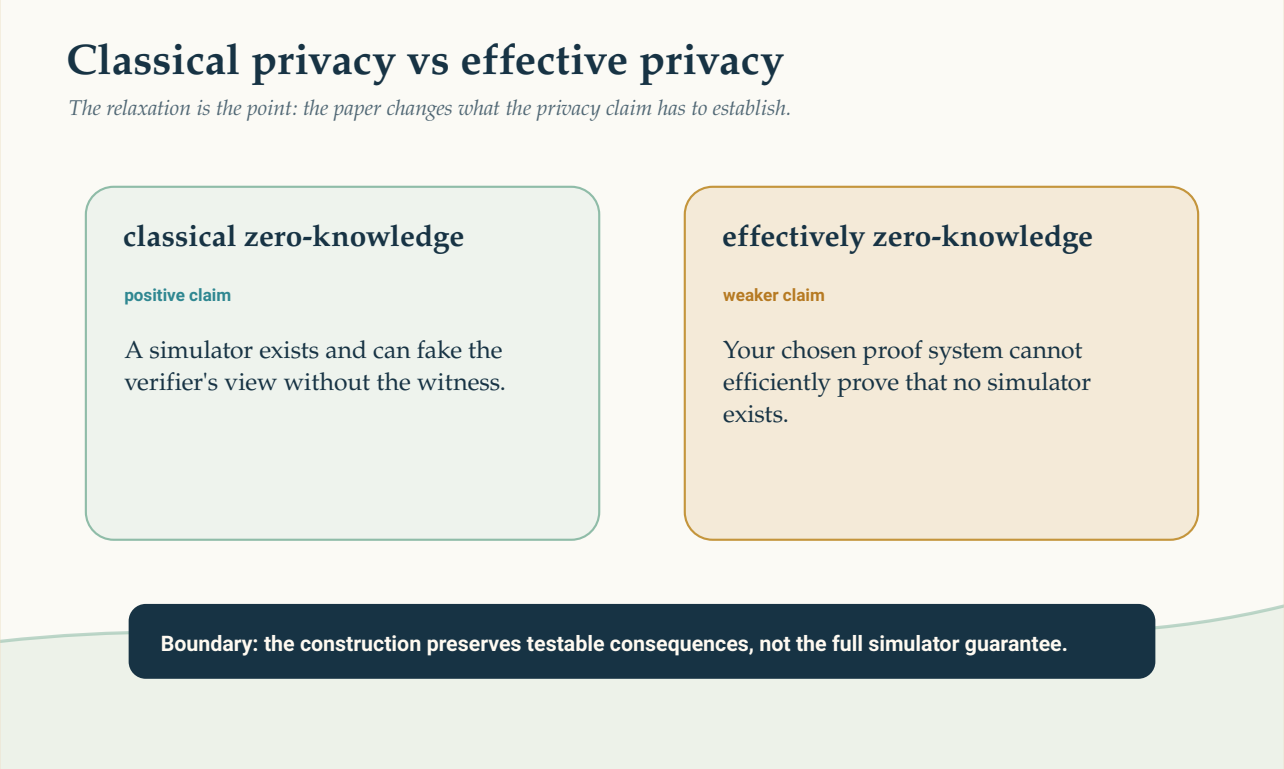
首先，零知識證明需要對話。如果 Alice 在事先沒有安排任何可信設定的情況下只傳送一則訊息，零知識保證就會崩潰——不論你願意以多少健全性作為交換，這點都成立。

其次，零知識證明需要容許一點錯誤。事實證明，要求完美健全性也會悄悄摧毀互動性：一個無論隨機選了什麼都永遠不會被騙過的驗證者，還不如預先固定那些選擇——而一旦驗證者變得可預測，Alice 就能用單一訊息回答所有問題，這正是前一種情況已經失敗的地方。

Rahul Ilango 的論文提出了一條繞過這道雙重高牆的路。不是假裝高牆不存在，也不是在不可能的設定中製造經典零知識。這個做法更細緻：削弱「什麼都不透露」的含義，但以一種保留密碼學家實際能檢驗的安全性質的方式削弱它。

這個結果稱為**有效零知識**。

零知識被三道門擋住——互動、可信設定，以及不完美健全性。Ilango 的構造從另一道門溜過去：規則書無法有效率地證明模擬器不可能存在。



經典零知識問的是模擬器是否存在；「有效零知識」只問你選定的規則書能否有效率地證明模擬器不存在。正是這個較弱的問題，讓該構造得以保留單一訊息、無設定和完美健全性。

舊測試：模擬器存在

將零知識形式化的經典方法，會引入一個虛構的助手，稱為**模擬器**。

想法是這樣：想像有一位不知道 Alice 秘密的 Jane。如果 Jane 完全靠自己，就能產生看起來與 Bob 從 Alice 那裡收到的證明相同的內容，那麼 Alice 的證明就沒有教給 Bob 任何新東西。即使沒有 Alice 的秘密，Jane 也能預先偽造出同樣的經歷。

因此，經典零知識要求一個真實存在的模擬器。必須有一個有效率的演算法，能在不知道秘密——術語稱為 *witness*，也就是「見證」；對數獨而言，見證就是解出的棋盤——的情況下，產生看似真實的證明。

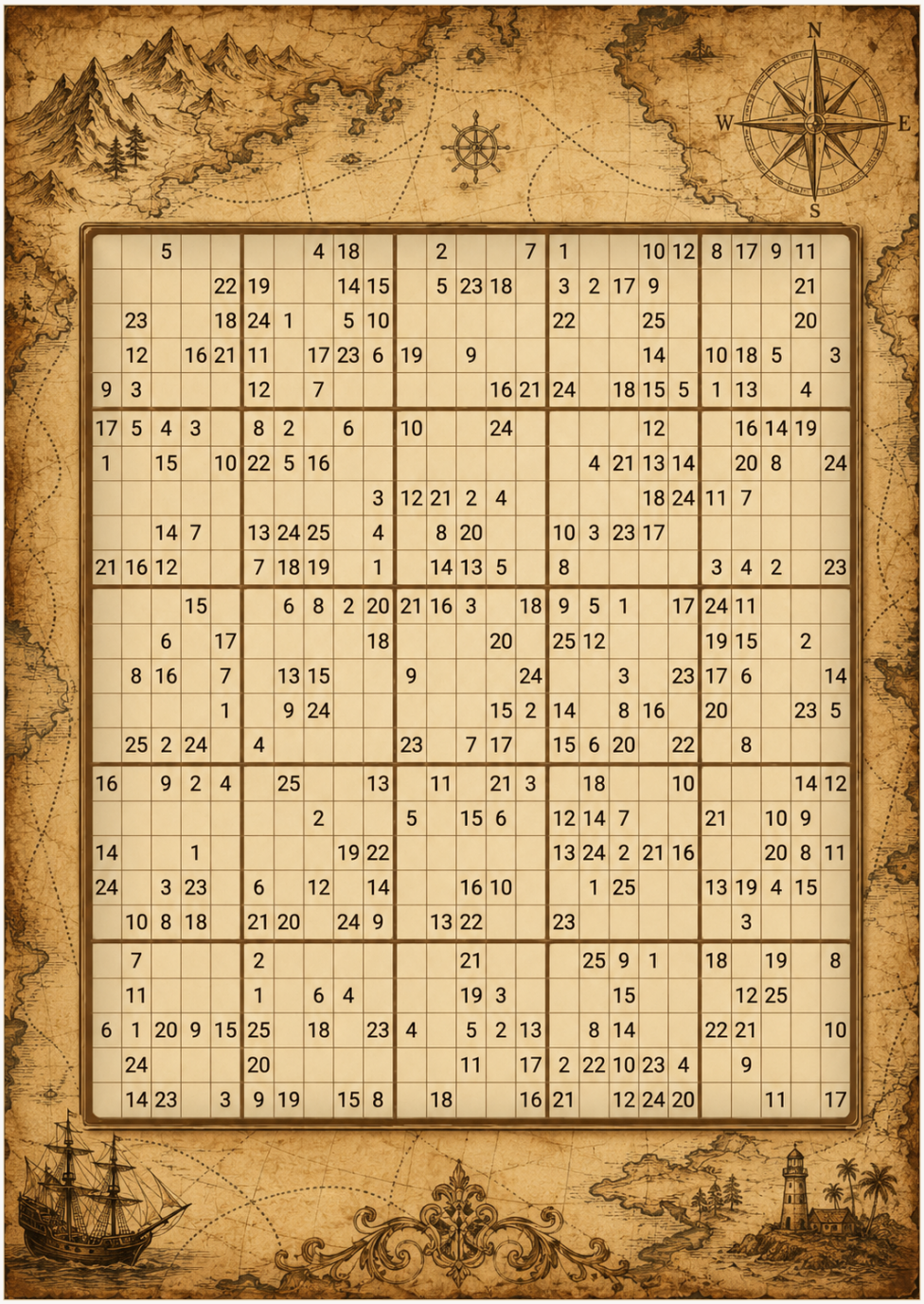
這個定義很強，但也正是舊有不可能性結果發揮作用的地方。直觀來說，真正非互動的證明只是一串字串。Bob 一旦取得那串字串，就可以把它展示給別人：他獲得了向其他人證明該陳述的能力，而這聽起來已經不只是「什麼都沒有」了。經典定理把這個直覺進一步精確化為前述的不可能性。

> **本論文堅持的三項性質**
>
> 論文標題點出了三項限制：
> **無互動：**Alice 傳送一串證明字串。沒有來回進行的協定。
> **無設定：**Alice 和 Bob 不依賴可信的共同參考字串，或其他預先安排的公開隨機性。許多稱為「非互動零知識」的系統仍依賴設定；本論文所說的是完全不需要設定。
> **完美健全性：**虛假陳述沒有有效證明。不是「幾乎不會被接受」；而是不存在任何有效證明。
> 這三項性質正是一般書面數學所擁有的——而且如上所述，經典零知識無法同時保留它們。

超大型數獨版本的差異

以下用一個刻意簡化的例子，幫助我們感受其中差異。

嚴肅討論這個類比時，不要使用普通的 9×9 數獨。它太小，也太有限：電腦可以直接解出來，或證明它無解。請改想像一族 **MegaSudoku(n)** 數獨。把通常的規則放大：選定區塊大小 n，令 $N = n^2$，再建立一個 $N \times N$ 的棋盤，分成 $n \times n$ 的區塊，使用 N 種符號。普通數獨只是很小的 $n = 3$、$N = 9$ 情況：9×9 棋盤、3×3 區塊和九種符號。只有當 n 可以增長，而且棋盤能攜帶額外的小工具，使它表現得像一個披著數獨外衣的 SAT 公式時，證明複雜度的故事才開始。SAT 公式只是一份真假限制清單：能不能替變數指定真／假值，使每項限制都得到滿足？



		5					4	18			2			7	1			10	12	8	17	9	11	
				22	19			14	15		5	23	18		3	2	17	9					21	
	23			18	24	1		5	10						22			25					20	
	12		16	21	11		17	23	6	19		9						14		10	18	5		3
9	3				12		7						16	21	24		18	15	5	1	13		4	
17	5	4	3		8	2		6		10			24					12			16	14	19	
1		15		10	22	5	16									4	21	13	14		20	8		24
									3	12	21	2	4					18	24	11	7			
		14	7		13	24	25		4		8	20			10	3	23	17						
21	16	12			7	18	19		1		14	13	5		8					3	4	2		23
			15			6	8	2	20	21	16	3		18	9	5	1		17	24	11			
		6		17					18				20		25	12				19	15		2	
	8	16		7		13	15			9				24			3		23	17	6			14
				1		9	24						15	2	14		8	16		20			23	5
	25	2	24		4					23		7	17		15	6	20		22		8			
16		9	2	4		25			13		11		21	3		18			10				14	12
							2			5		15	6		12	14	7			21		10	9	
14			1					19	22						13	24	2	21	16			20	8	11
24		3	23		6		12		14			16	10			1	25			13	19	4	15	
	10	8	18		21	20		24	9		13	22			23						3			
	7				2							21				25	9	1		18		19		8
	11				1		6	4				19	3				15				12	25		
6	1	20	9	15	25		18		23	4		5	2	13		8	14			22	21			10
	24				20							11		17	2	22	10	23	4		9			
	14	23		3	9	19		15	8		18			16	21		12	24	20			11		17

一個 25x25 數獨：它的規則可以在不透露完成棋盤的情況下檢查——可視為驗證隱藏解答（見證）的證明之視覺替身。

AI-generated editorial thumbnail —The Clean Paper CC BY 4.0

數獨與 SAT：同一個謎題的兩種外衣

數獨可以「表現得像 SAT 公式」這個說法不是比喻。兩個方向都能進行翻譯，而且較容易的方向可以完整寫出來。

從數獨到 SAT。SAT 只談真假，因此對每個（列、欄、值）三元組給一個布林變數：$x(r,c,v)$ 表示「第 r 列、第 c 欄的儲存格包含值 v」。一個 4×4 數獨（2×2 區塊，值為 1–4）需要 4·4·4 = 64

個變數；經典的 9×9 數獨則需要 729 個。接著，每條數獨規則都會變成一批子句。（子句是變數或其否定的 OR；整個公式則是所有子句的 AND。）
每個儲存格至少包含一個值——每個儲存格一個子句：

$$x(1,1,1) \lor x(1,1,2) \lor x(1,1,3) \lor x(1,1,4)$$

每個儲存格至多包含一個值——每一對值各有一個「不可同時成立」子句：

$$\neg x(1,1,1) \lor \neg x(1,1,2) \quad \neg x(1,1,1) \lor \neg x(1,1,3)$$ …其餘六對也依此類推。

每一列包含每個值——以第 1 列的值 3 為例，至少出現一次：

$$x(1,1,3) \lor x(1,2,3) \lor x(1,3,3) \lor x(1,4,3)$$

而至多出現一次：$\neg x(1,1,3) \lor \neg x(1,2,3)$，該列內每一對儲存格都依此處理。
欄與區塊——批次形式完全相同；只有儲存格群組改變。以左上方區塊的值 2 為例：

$$x(1,1,2) \lor x(1,2,2) \lor x(2,1,2) \lor x(2,2,2)$$

再加上成對的「不可同時成立」子句。
印出的提示——最簡單的部分：每個提示都是只有一個變數的子句。左上角印出的 3 會變成子句

$$x(1,1,3)$$

所有這些內容的 AND，在且僅在數獨有解時才可滿足；而一個滿足指派就是解答：讀出哪些 x(r,c,v) 為真，將它們填回棋盤。對 9×9 數獨而言，這會得到 729 個變數和幾千個子句，現代 SAT 求解器在幾毫秒內就能處理完。注意提示子句 x(1,1,3)：它說的是「這個儲存格恰好等於 3」，而不是「這些儲存格彼此不同」——正是這種不對稱性，會迫使我們在稍後協定的提示儲存格中加入額外技巧。
從 SAT 到數獨。論文需要的是相反、也較困難的方向：給定一個任意的 SAT 公式，建立一個恰好在該公式成立時有解的超大型數獨。數獨原生規則只能表達「這些儲存格彼此不同」，因此必須把任意邏輯限制建造出來——而這正是小工具的作用。小工具是一小組預先製作的儲存格，每個公式子句對應一組；其中指定的儲存格扮演變數角色（它們所持有的符號編碼真或假），而群組內部的限制經過設計，使其唯一合法的填法恰好對應於滿足該子句的指派。這是 NP 完全性證明中的標準技藝；對一般化數獨而言，Yato 和 Seta 於 2003 年完成了這項工作。
兩個方向合在一起，表示 N×N 數獨與 SAT 是穿著不同外衣的同一個問題。這正是本文——以及該論文——得以用棋盤與符號講述整個 NP 故事的理由。

這個見證仍然很容易想像。Alice 知道超大型數獨的一份完整、有效填法。Bob 想被說服這樣的填法存在，但 Alice 不想透露它。如果她傳送整份填法，Bob 會相信，但秘密也就消失了。

在經典零知識版本中，Alice 和 Bob 會互動。一種老派的心智模型使用覆蓋的牌片。Alice 藏起已解出的棋盤，在每一輪之前秘密重新命名符號，讓 Bob 檢查一個隨機選出的局部限制：一列、一欄、一個區塊，或一個小工具。如果揭開的儲存格顯示彼此不同的符號，Bob 就更有信心。接著所有東西再次覆蓋，符號也重新命名。（有一個細節：謎題給出的提示需要額外技巧，因為重新命名也會把提示藏起來。下方的註解會說明經典協定如何處理；接下來的討論只需要這個玩具模型。）

經典協定實際上如何處理提示儲存格

重新命名技巧有一個盲點。列、欄和區塊的規則都說「這些儲存格彼此不同」，而彼此不同在任何符號重新命名下都不變。但提示說的是「這個儲存格包含恰好 5」，重新命名後 Bob 只看得到 $\sigma(5)$——某個被遮蔽的符號——卻不知道重新命名 σ。他什麼都無法檢查。若不加處理，Alice 可以證明「存在某個有效棋盤」，卻完全忽略印出的提示，這對這個謎題什麼也沒證明。經典文獻有兩種標準修補方式。

符號表。在隱藏棋盤旁增加一列 N 個儲存格——這是一個符號表，由 Alice 按固定且公開的順序填入符號 1…N，然後和其他一切一起重新命名，因此內容會是 $\sigma(1)…\sigma(N)$。Bob 的隨機挑戰現在多了一個選項。除了選擇一列、一欄、一個區塊或一個小工具來揭開之外，他還可以選擇**符號表加上一個提示儲存格**。Alice 同時揭開兩者；符號表揭示本輪的重新命名，Bob 再檢查提示儲存格是否恰好顯示印出提示的重新命名版本。這仍然是零知識，因為 Bob 只學到 σ——它每輪都重新隨機抽取，單獨沒有價值——以及一個他早已從謎題知道的儲存格值。秘密儲存格沒有洩漏任何資訊，模擬器可以透過抽取隨機 σ 來偽造這個視圖。它是健全的，因為作弊的 Alice 每輪都有固定機率被抓到，而重複進行足夠多輪後，作弊未被發現的機率會變得可忽略。

編譯掉提示。較結構性的變體不增加特殊挑戰，而是移除它。不要驗證提示值，而是用差異限制強制指定它：將提示儲存格連到除承載其自身值的符號表儲存格之外的每一個符號表儲存格——「不同於 $\sigma(1)$、不同於 $\sigma(2)$、……不同於 $\sigma(5)$ 以外的一切」。該儲存格唯一能合法持有的符號就是提示指定的符號。現在每項限制又都是「這兩者不同」的形式；它對重新命名保持不變，也能像列一樣檢查。這正是經典圖著色協定處理預先著色頂點時使用的同一招，也就是上文 *gadgets* 一詞的精神：在「超大型數獨即 SAT」的圖像中，提示被編譯成與其他限制一樣的不等式小工具。

實體協定。現實世界的數獨卡牌協定（Gradwohl、Naor、Pinkas 和 Rothblum，2007）完全不使用重新命名，而是在隱藏開始之前先處理提示。對每個儲存格，Alice 放下三張相同、印有該儲存格值的牌——秘密儲存格面朝下，**提示儲存格則面朝上**，讓 Bob 親眼看到提示在翻牌前已獲遵守。接著每個儲存格的一張牌分別放入其列、欄和區塊的牌堆；每堆洗牌並揭開，Bob 檢查其中包含全部 N 種符號。洗牌破壞了位置資訊（這就是零知識），但提示在發牌時就已經固定下來。

無論採用哪一種方式，教訓都和本文不斷回到的重點相同：零知識協定是在仔細記帳，記錄隱藏後究竟哪些事實仍然存在。重新命名保留「彼此不同」，卻抹去「等於 5」——所以必須用其他方式把「等於 5」偷渡回來。

那不是論文中的協定，而是經典零知識的心智模型：

- Alice 和 Bob 來回互動。
- Bob 選擇隨機檢查。
- Alice 只揭示局部一致性，而不是完整解答。
- 隱私之所以成立，是因為可以證明 Bob 所看到的內容，原本就能在沒有 Alice 秘密解答的情況下產生。

所以，經典零知識建立在一個正面事實上：

> 模擬器確實存在。

現在移除那些令人安心的部分。Alice 傳送一串證明字串後離開。沒有可信設定，沒有事先準備好的共享隨機字串，而且 Bob 絕不能接受虛假謎題。這正是經典零知識無法存活的設定。

在技巧出現前，還需要另一個角色。固定一份**規則書**：以邏輯學家的意思來說，一個形式化證明系統——由固定公理加上檢查書面數學證明的機械規則組成。ZFC，即標準的數學公理，是典型例子。從這裡開始，所有內容都相對於事先選定的規則書陳述，而選擇具有彈性：無論你固定哪一本規則書，這個構造都能運作，包括 ZFC。

（關於用詞的一點說明，取自論文本身：「證明系統」在此一律指這本規則書——檢查數學證明的形式系統——絕不指 Alice 傳送的訊息。Alice 和 Bob 使用的機制稱為「證明者與驗證者」。）

哥德爾風格的版本保留超大型數獨的故事，但改變證明方式。

選擇第二個表面上大小相同的限制系統，稱為 **D**。在故事中，S 和 D 是採用相同格式的兩個 MegaSudoku(n) 謎題。在幕後，D 可能起初是另一種尺寸的困難邏輯公式；若有需要，可以填入無害的虛擬限制，使它符合相同棋盤。D 是由一個實際上**不可滿足**的邏輯公式建成：不存在任何值指派能讓它的所有限制都為真，就像一個損壞的謎題沒有合法的完整棋盤一樣。玩具例子可以是一個同時要求「X 為真」和「X 為假」的公式。因此 D 沒有有效填法。

但 D 不能是那種一眼就能揭穿的壞謎題。上面的玩具例子就不合格：任何規則書都能用一行反駁「X 且非 X」。D 必須是假的，卻不能讓所選規則書用簡短論證證明它為假。若規則書能用短證明反駁 D，後面的故事就站不住腳：那條本可在不知道 Alice 秘密的情況下產生證明的替代路徑，會被形式化地排除，隱私保證也會隨之消失。因此，D 要從一族固定規則書無法有效率反駁的公式中選出：在這套規則書內，沒有短證明能證明 D 無解。

Alice 的單一訊息證明，於是針對以下二選一陳述：

> 真正的超大型數獨 S 有解，或誘餌 D 有解。

這就是邏輯連結。D **不是**以某種神奇方式產生，讓 S 因而為真。證明並不是在論證「D 無解，所以 S 有解」。它證明的是析取命題 **S 或 D**。完美健全性表示，虛假的析取命題不可能有有效證明。既然 D 在現實中是假的——它沒有解——該析取命題為真的唯一方式，就是 S 為真。因此，如果證明被接受，S 就必定有解。誘餌不能讓虛假的 S 變成真的。

但對零知識風格的部分，要問的是如果 D 真的有解，會發生什麼事。那個誘餌解答會充當替代見證。它能讓某人不用知道 Alice 真正的超大型數獨解答，也產生證明——換句話說，就是一個模擬器。現實中 D 沒有解，所以這條模擬路徑被封閉了。重點是，規則書無法有效率地證明它已被封閉。

因此 D 有兩項任務。對**健全性**而言，D 是假的，所以「S 或 D」的有效證明會迫使 S 為真。對**有效零知識**而言，D 很難被反駁，所以規則書無法迅速排除那條原本能讓模擬成為可能的誘餌路徑。

所以安全性測試不再是：

> 我們能證明模擬器確實存在嗎？

而是：

> 你的規則書能有效率地證明模擬器不可能存在嗎？

如果答案是否定的，就會得到一個出人意料地強的結論：任何安全保證，只要（a）能藉由執行測試觀察到，且（b）能在該規則書內證明會由模擬器的存在推出，就確實成立。因為，若能成功攻破其中任何一項，本身就等於給出了那份不存在的短反駁。這正是「有效零知識」中「有效」的含義。

因此，課堂式的對照是：

經典零知識：證明是安全的，因為模擬器存在。

哥德爾風格的有效零知識：對可觀察的安全測試而言，證明被視為安全，因為規則書無法有效率地證明模擬器不可能存在。

第二項主張較弱。也正因如此，論文才能保留讓經典版本失敗的三項特徵：單一訊息、無設定和完美健全性。

新測試：你無法證明模擬器不存在

Ilango 的放寬版本改變了問題。

經典零知識問的是：

> 模擬器存在嗎？

有效零知識問的是一個較弱的問題：

> 你選定的規則書能有效率地證明不存在任何模擬器嗎？

這聽起來像技術性的迴避，但它正是核心概念。這個構造處在一種奇怪狀態：模擬器實際上不存在——論文明確如此說明——但你固定的規則書無法有效率地證明它不存在。如果你關心的每個壞結果都需要這樣的反駁，那麼對那些結果而言，系統仍然表現得像零知識。

哥德爾就是在這裡登場。不是裝飾，也不是「哥德爾讓密碼學安全」。兩者的連結是證明論上的。一份規則書若在精確意義下是可能的最佳規則書，就稱為**最佳**：只要有任何規則書能以短證明反駁某種相關公式，這份最佳規則書也能做到，而且證明長度至多增加多項式倍數。Krajíček 和 Pudlák 於 1989 年猜想，**不存在最佳證明系統**：無論你固定哪一本規則書，都會有另一份規則書能用短得多的證明，證成某一族真陳述。這是證明複雜度的核心未解猜想之一，也是哥德爾不完備定理在有限、複雜度理論上的近親：某些真陳述在你固定的規則書中沒有短證明——不是因為它們原則上不可證明，而是因為每一份固定規則書都會留下某些沒有短證明的真理。

論文假設這項猜想成立（採用一種稍強的「無窮多次」形式，這是將猜想用於密碼學時的標準做法）。根據 Krajíček 和 Pudlák 的定理，其具體收益是：對每一本規則書，都存在一串確實不可滿足的公式，規則書無法用短證明反駁它們——更關鍵的是，有一個有效率的演算法可以產生這串公式。最後這項性質稱為一致生成性，它把整個想法從存在性主張變成 Alice 真能執行的演算法：她的誘餌來自生產線，而不是憑空出現。

密碼學上的做法，就是利用這種證明能力的不足。

這個構造在做什麼

以下是論文構造的骨架版。

固定一份規則書——例如 ZFC。在證明複雜度假設下，存在一串能有效率產生的公式；它們實際上不可滿足，但規則書沒有證明它們不可滿足的短證明。

現在建立一個如下形式的單一訊息證明：

> 真正的陳述可滿足，或這個特殊的困難公式可滿足。

特殊的困難公式並不可滿足。因此，如果底層證明機制是完美健全的，接受這則訊息仍然代表真正的陳述為真。這就提供了完美健全性。

但對零知識式安全性而言，想像特殊的困難公式如果可滿足。那麼它的見證就能在不知道真正見證的情況下，用來模擬證明。公式在現實中不可滿足——但規則書無法有效率地證明這一點。因此，它也無法有效率地證明模擬器不可能存在。

這就是支點。系統不是靠產生經典模擬器來藏起秘密。對一大類可觀察的安全測試而言，它是把秘密藏在規則書無法證明模擬器不存在這件事背後。

論文聲稱了什麼

主要定理分成幾層。核心結果如下：

在一項標準密碼學假設——存在**非互動見證不可區分證明**，這是廣受研究且可由數個已建立的假設套件推出的物件——以及「**不存在（無窮多次的）最佳證明系統**」這項證明複雜度猜想下，對每一份規則書，論文都能為 NP/SAT 構造一個單一訊息的證明者與驗證者；它具有**完美健全性**且無需設定，並相對於該規則書達到有效零知識。（各類謎題問題通常都可歸結為 NP/SAT 這個標準的困難核心；超大型數獨只是它的一種外觀。）

對於保留可證偽安全性質的更廣泛主張，論文再加入一項標準假設，即去隨機化信念 **P = BPP**（大致來說：隨機性不會為演算法帶來本質上的額外能力）。

把定理語言翻譯出來：

- 證明只有一則訊息。
- 沒有可信設定。
- 虛假陳述無法被證明。
- 證明者不是經典零知識——它沒有模擬器。
- 但經典零知識所有可證偽、基於遊戲的安全性後果，都能在這個設定中達成。

「可證偽」很重要。它表示安全性失敗可以透過在一場遊戲中執行攻擊者來測試。許多密碼學安全性定義都採用這種形式：攻擊者能否區分兩個加密結果、反轉某個函數、取回見證，或在指定實驗中獲勝？定理針對每一項可證偽性質，各自給出一個證明者。單一證明者同時享有每一項可證偽性質，

很可能是不可能的——舊有的可重用性攻擊（「Bob 可以把證明展示給其他人」）本身就是一項可證偽性質，而且在這裡確實失敗。論文提出的是，一個單一證明者可以合理地涵蓋所有自然的可證偽性質——也就是密碼學實務中實際出現的那些——但這部分是建立在「自然」這個非正式概念，以及一項明確猜想上的條件性定理。這項保證針對的是可觀察的失敗，而不是每一種哲學上或基於模擬的秘密意義。

有一個具體推論值得點名：這個構造產生了第一批具一致生成證明者的非互動見證隱藏證明——「謎題的證明不會幫助你找出解答」，而且無互動、無設定——這個聽來謙遜的物件，幾十年來一直難以構造。

這不代表什麼

這一節是讓文章保持誠實的地方。

它**不**表示舊有不可能性定理錯了。這個構造是藉由改變定義來避開它們。

它**不**提供無互動、無設定且具完美健全性的普通經典零知識。論文明確指出，構造出的證明者沒有模擬器。

它**不**表示證明無法被重複使用。單一訊息證明仍然可以展示給其他人看；論文沒有保留否認性風格的性質。（具有可信設定的非互動零知識也有相同限制。）

它**不**表示這是一個可以立即部署的實用協定。這是複雜度理論與密碼學基礎研究。結果依賴證明複雜度和密碼學中的重大假設，而這個構造談的是原則上可能做到什麼。

它**不**把「哥德爾」變成神奇的安全原語。哥德爾的連結是透過證明系統、最佳證明系統和不完備性的有限類比建立的。可用的直覺不是「不完備性保護你的密碼」。而是：如果規則書無法有效率地證明模擬器不可能存在，那麼需要這項證明的攻擊，就能在安全性定義層次被擋下來。

為什麼仍然有趣

密碼學經常把困難性轉化成安全性。因數分解很難，所以 RSA 式假設變得有用。格問題很難，所以格密碼學變得有用。這裡的困難性更奇特：不是「難以計算出秘密」，而是「難以證明某種證明物件不可能存在」。

這就是這篇論文令人不尋常的地方。它幾乎把公理和規則書當成密碼學資源來處理。通常的不可能性結果說，健全性和模擬之間存在張力。Ilango 的動作，是把這種張力放到證明論的帷幕後：模擬器不存在，但形式系統無法有效率地揭露模擬器不存在這件事。

對讀者而言，令人意外的部分不是它將取代今日的零知識系統。至少不會直接取代。真正意外的是，數理邏輯的一項限制可以被建設性地使用：不只是作為一道牆，也可以作為某種掩護。

證據有多強？

這是一篇定理論文，因此「證據」的含義不同於生物學或天文學論文。問題不是實驗是否重現，而是定義、假設和證明鏈是否支持該主張。

證明是形式化的，論文也明確列出假設。這些假設並非隨意。非互動見證不可區分證明是密碼學中的標準物件，能由數個已建立的假設套件推出。不存在最佳證明系統是證明複雜度中的核心猜想。P = BPP 則是標準的去隨機化信念，只用於更廣泛的可證偽性質定理。

論文也主張，這些假設是恰當的代價，而非任意搭建的腳手架：它證明了一個逆命題，顯示這些假設在本質上是必要的——如果這類構造確實存在，那麼非互動見證不可區分證明必須存在，而且（假設標準的單向函數存在）最佳證明系統就不可能存在。這些假設還是「雙贏」的：反駁其中任何一項，本身都會成為證明複雜度、密碼學或複雜度理論中的里程碑發現。

但由於結果是條件性的，它的可信度也同樣是條件性的。如果那些假設失敗，定理的解讀就會改變。即使假設成立，這項保證也不是完整的經典零知識，而是論文所提出的、經過放寬的證明論版本。

因此，合理的信心是：論文確實建立了一個連貫的條件性可能性結果，這點的信心很高；它的假設描述了我們實際身處的密碼學世界，信心中等；任何立即的實務後果，信心偏低。

為什麼重要

這篇論文開啟了一條原本應該封閉的路。

經典理論說：完整零知識在沒有設定時不能只有一則訊息，也不能具備完美健全性。Ilango 的論文則說：如果我們要求的是那些能在安全性遊戲中測試的零知識後果，並讓安全性定義取決於規則書能否有效率地反駁某件事，那麼許多有用的行為都能被恢復——同時保留單一訊息、無設定和完美健全性。

這不是小小的定義調整，而是思考密碼學保證的一種不同方式。不要只問什麼存在，也要問你的規則書能排除什麼。不要把不可證明性視為哲學上的麻煩，而要把它當成結構使用。

實際世界或許不會在明天改變。但概念地圖改變了。現在有一種形式化意義上的說法：「沒有人能有效率地證明秘密已經洩漏」可能強大到足以恢復我們原本希望從「秘密沒有洩漏」取得的許多基於遊戲的保護。

這就是哥德爾會出現在標題中的原因。

重點摘要

零知識證明讓證明者能使驗證者相信某個陳述為真，卻不透露見證。經典不可能性結果指出，零知識無法在沒有設定的情況下被壓縮成單一訊息，也不能具備完美健全性。Rahul Ilango 的論文沒有反駁這些不可能性，而是定義了較弱的概念，即有效零知識：不要求模擬器真的存在，而要求選定的證

明系統——像 ZFC 一樣的形式規則書——無法有效率地證明不存在任何模擬器。在密碼學的重大假設（非互動見證不可區分證明）和證明複雜度的假設（不存在最佳證明系統）下，論文為 NP/SAT 構造了無設定且具完美健全性的單一訊息證明者，逐項達成零知識中可證偽、基於遊戲的安全性後果。一個涵蓋所有「自然」此類性質的單一證明者，是更進一步、部分依賴猜想的延伸；而涵蓋字面上每一項可證偽性質則很可能不可能，因為證明仍可被重複使用。這項結果是理論性且條件性的，不是已部署的原語，但它展示了把證明論上的不可證明性當作密碼學資源的新方式。

務實檢視

論文展示了什麼：在既定假設下，可以為 NP/SAT 建立單一訊息、無設定、具完美健全性的證明者；相對於任何選定的證明系統，它們是有效零知識，並能達成經典零知識的每一項可證偽、基於遊戲的後果。

合理但尚未無條件證明的內容：所需的證明複雜度和密碼學假設確實成立。它們是嚴肅且經過充分研究的假設——論文也顯示它們不只是充分，而且本質上必要——但終究仍是假設。

它沒有展示的內容：無互動、無設定且具完美健全性的經典零知識；可以立即部署的實用系統；證明的否認性或不可重用性；或哥德爾不完備定理本身能保障密碼學。

主要限制：這項保證是零知識的放寬版本；最廣泛的版本依賴多項假設；單一通用證明者的主張仍部分建立在猜想上；而且結果主要屬於基礎研究。

一般讀者應有多大信心？如果接受這些定義，對這是一項重要的條件性理論結果，信心很高。對假設能否反映現實，信心中等。對立即實務部署，信心偏低。安全的結論是：這篇論文沒有打破零知識的不可能性；它找到了一種新的證明論方式，繞過其中對許多安全性遊戲而言真正重要的部分。

來源

Ilango, IACR ePrint 2025/1296
https://eprint.iacr.org/2025/1296

FOCS 2025, DOI 10.1109/FOCS63196.2025.00058
https://doi.org/10.1109/FOCS63196.2025.00058