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 · 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

# 訣竅不是證明秘密被隱藏了

從零知識最簡單的版本開始。

Alice 想說服 Bob：某個數獨謎題有解。如果她把解發過去，Bob 就信服了，但謎題也就被毀了。她想要的東西更奇怪：一個證明解存在、卻不洩露解本身的證明。

這就是零知識證明的承諾。證明者（Alice）說服驗證者（Bob）某個陳述為真，同時除了該陳述為真這一點之外不洩露任何東西。

問題在於，這個承諾是有代價的。一個普通的數學證明有兩個令人舒適的特徵。它是一**條消息**：你把它寫下來，交出去，然後走開。而且它是**完美可靠的**：一個假的陳述根本沒有任何有效的證明。經典的不可能性結果表明，零知識必須放棄這兩個特徵——而且不僅僅是二者一起放棄；每一個特徵單獨來看都是禁區。

首先，零知識證明需要對話。如果 Alice 發送一條消息，而事先沒有安排任何可信設置，那麼零知識保證就會崩潰——而且無論你願意為交換而放棄多少可靠性，這一點都成立。

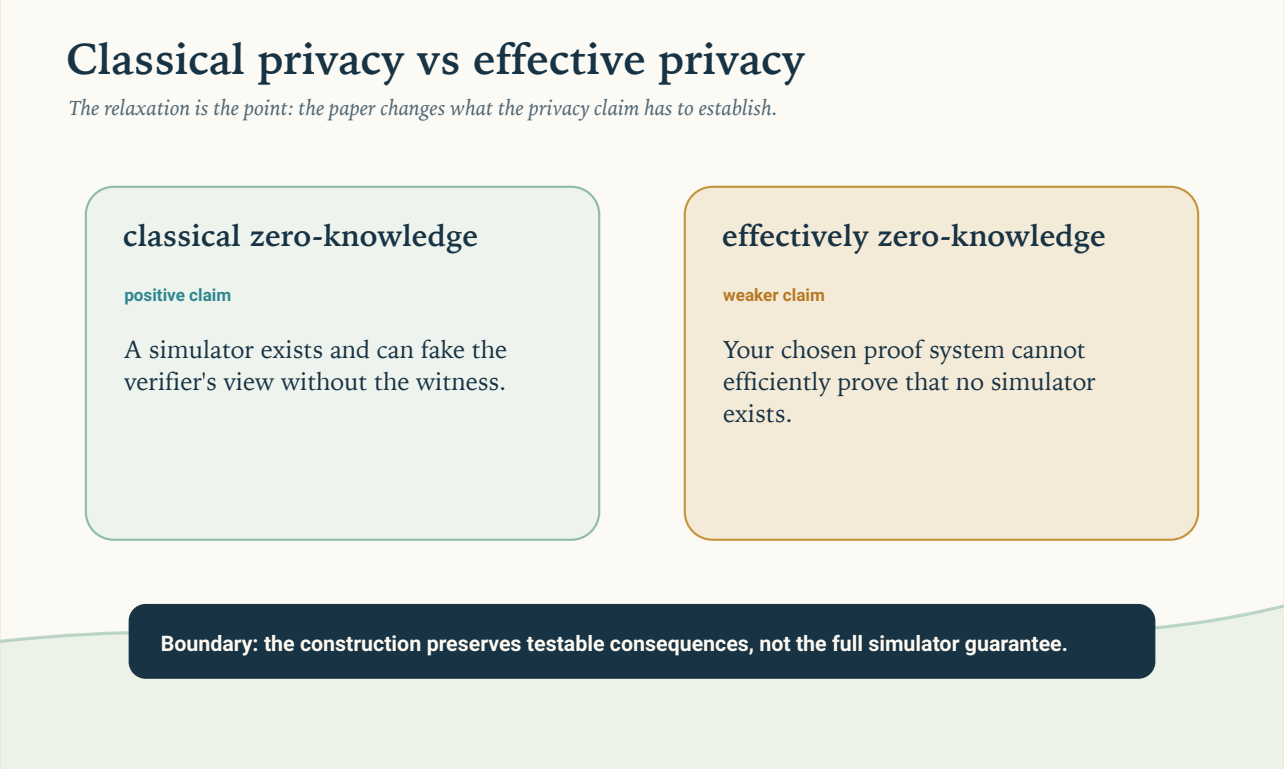
其次，零知識證明需要對錯誤有一點點容忍。事實證明，要求完美可靠性也會悄悄地摧毀交互：一個無論做出哪些隨機選擇都永遠不會被欺騙的驗證者，不如乾脆事先把那些選擇固定下來——而一旦驗證者是可預測的，Alice 就可以在一條消息裡回答所有問題，這恰恰就是前面已經崩潰的那種情形。

Rahul Ilango 的論文講的是一種繞過這道雙重高牆的方法。不是假裝這堵牆不存在，也不是在那個不可能的設定下造出經典的零知識。這一手更微妙：削弱「不洩露任何東西」的含義，但以一種能保留密碼學家實際上可以檢驗的安全性質的方式來削弱它。

這個結果被稱為**有效零知識**。

零知識在三道門前被擋住——交互、可信設置和不完美的可靠性。Ilango 的構造從另一道門溜了過去：規則手冊無法高效地反駁模擬器。

Original diagram —The Clean Paper CC BY 4.0




經典零知識問的是模擬器是否存在；「有效零知識」只問你所選的規則手冊能否高效地證明模擬器不可能存在。正是這個更弱的問題，讓這個構造得以保留一條消息、無設置和完美可靠性。

Original diagram —The Clean Paper CC BY 4.0

# 舊的檢驗：存在一個模擬器

將零知識形式化的經典方法用到了一個虛構的助手，叫作**模擬器**。

想法是這樣的：想像一個叫 Jane 的人，她**不**知道 Alice 的秘密。如果 Jane 完全靠她自己就能生成一些看起來和 Bob 本會從 Alice 那裡收到的證明一模一樣的證明，那麼 Alice 的證明並沒有教給 Bob 任何新東西。Jane 在沒有 Alice 秘密的情況下就已經能偽造出這種體驗了。

所以經典零知識要求有一個真正的模擬器。必須存在一個高效的演算法，它能在不知道秘密——用行話說就是見證——的情況下生成看起來像模像樣的假證明；對數獨來說，見證就是那張解好的網格。

這個定義很強大，但它也恰恰是舊的不可能性發作的地方。直覺是這樣的。一個真正非交互的證明只是一個字串。一旦 Bob 拿到了這個字串，他就可以把它拿給別人看：他獲得了向他人證明該陳述的能力，這聽起來已經不止是「什麼都沒有」了。經典定理把這個直覺銳化成了上面那些不可能性。

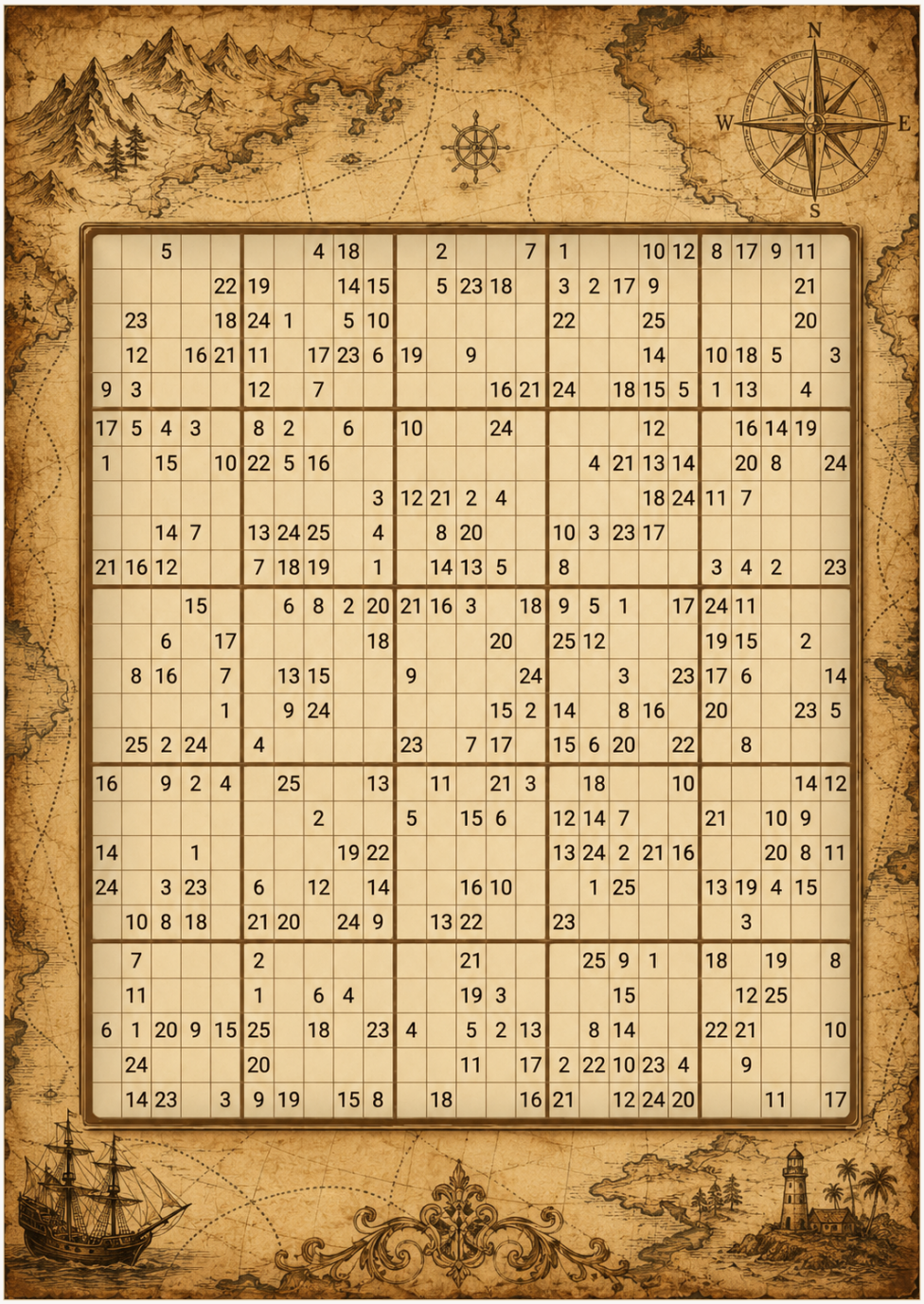
**本文堅持的三個性質**

論文的標題點出了三個約束：
**無交互：** Alice 發送一個證明字串。沒有來回往復的協議。
**無設置：** Alice 和 Bob 不依賴任何可信的公共參考字串或其他事先安排好的公共隨機性。許多被稱為「非交互零知識」的系統仍然依賴設置；本文所說的是零設置。
**完美可靠性：**一個假的陳述沒有任何有效的證明。不是「幾乎從不被接受」；而是根本不存在有效的證明。
這三個性質恰恰是普通的書面數學所具備的——而且，正如上面所解釋的，經典零知識無法保留它們。

# 這個區別的超級數獨版本

下面是一種刻意簡化的、用來感受這個區別的方式。

在這個類比的嚴肅部分，不要用普通的 9×9 數獨。它太小、太有限了：電腦可以直接解出它，或者證明它無解。取而代之，想像一族 **MegaSudoku(n)**（超級數獨）謎題。把通常的規則放大：選擇一個塊大小 $n$，令 $N = n^2$，構建一個 $N \times N$ 的網格，劃分為 $n \times n$ 的塊，使用 $N$ 個符號。普通數獨只不過是那個很小的 $n = 3$、$N = 9$ 的情形：一個 9×9 的網格、3×3 的塊和九個符號。只有當允許 $n$ 增長，並且當網格可以承載額外的小配件（gadget），使它表現得像一個偽裝成數獨謎題的 SAT 公式時，證明複雜度的故事才真正開始。一個 SAT 公式只不過是一列是/否約束：你能否給這些變數賦上真/假值，使得每一個約束都被滿足？




| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 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 25×25 的數獨：它的規則可以在不洩露完成後的網格的情況下被檢驗——這是一個對隱藏解（即見證）進行驗證的證明的視覺替身。

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

### 數獨與 SAT：同一個謎題的兩種裝扮

說數獨可以「表現得像一個 SAT 公式」，這並不是一個比喻。這種轉換在兩個方向上都成立，而且容易的那個方向可以完整地寫出來。

**從數獨到 SAT。** SAT 只會說真/假，所以給每個（行、列、值）三元組配一個布林變數：$x(r,c,v)$ 表示「第 $r$ 行、第 $c$ 列的格子裡裝著值 $v$」。一個 4×4 的數獨（2×2 的塊，值為 1–4）需要 4·4·4

= 64 個變數；經典的 9×9 需要 729 個。於是，每一條數獨規則都變成一批子句。（一個子句是若干變數或它們的否定的「或」（OR）；整個公式是它所有子句的「與」（AND）。）
每個格子至少裝一個值 ——每個格子一條子句：

x(1,1,1) ⊠ x(1,1,2) ⊠ x(1,1,3) ⊠ x(1,1,4)

每個格子至多裝一個值 ——對每一對值都有一條「不能兼得」的子句：

¬x(1,1,1) ⊠ ¬x(1,1,2)  ¬x(1,1,1) ⊠ ¬x(1,1,3)  … 對全部六對依此類推。

每一行都包含每一個值 ——對第 1 行和值 3 來說：至少出現一次，

x(1,1,3) ⊠ x(1,2,3) ⊠ x(1,3,3) ⊠ x(1,4,3)

以及至多出現一次：¬x(1,1,3) ⊠ ¬x(1,2,3)，並對該行中每一對格子依此類推。
列和塊 ——同樣的一批批子句；只是格子的分組變了。對左上角的塊和值 2 來說：

x(1,1,2) ⊠ x(1,2,2) ⊠ x(2,1,2) ⊠ x(2,2,2)

再加上成對的「不能兼得」子句。
印好的提示數字 ——最簡單的部分：每個提示都是一條只含單個變數的子句。左上角印著的一個 3 就變成子句

x(1,1,3)

所有這些的「與」（AND）恰好在數獨有解時可滿足——而且一個可滿足的賦值就是那個解：讀出哪些 x(r,c,v) 為真，然後把網格填上。對一個 9×9 來說，這總共是 729 個變數和幾千條子句，一個現代 SAT 求解器在幾毫秒內就能處理完。注意那條提示子句 x(1,1,3)：它說的是「這個格子恰好等於 3」，而不是「這些格子兩兩不同」——正是這種不對稱性，會迫使下面那個協議說明裡對提示格子採用額外的技巧。
**從 SAT 到數獨**。論文需要的是相反的、更難的那個方向：給定一個任意的 SAT 公式，構造一個恰好在該公式有解時才有解的超級數獨。數獨原生的規則只會說「這些格子兩兩不同」，所以任意的邏輯約束必須被構造出來——而這正是那些 gadget 所做的事情。一個 gadget 是一小簇預先製造好的格子，公式的每一條子句對應一簇，其中指定的一些格子扮演變數的角色（它們所裝的符號編碼真或假），而這簇格子的內部約束被精心設計，使得它唯一合法的填法對應於滿足該子句的那些賦值。這是 NP 完全性證明中的標準手藝；對於廣義數獨，它由 Yato 和 Seta 在 2003 年完成。
把這兩個方向合起來，就是說 N×N 數獨和 SAT 是同一個問題穿著不同的戲服。正是這一點，才准許這篇文章——以及這篇論文——用網格和符號來講一個關於整個 NP 的故事。

見證仍然很容易想像。Alice 知道超級數獨的一種完整的合法填法。Bob 想要被說服這樣一種填法存在，但 Alice 不想把它洩露出來。如果她把整個填法發過去，Bob 就信服了，但秘密也就沒了。

在經典零知識的版本裡，Alice 和 Bob 進行交互。一個老派的思維模型用到了蓋住的方塊。Alice 把解好的網格藏起來，在每一輪之前偷偷地給符號重新命名，然後讓 Bob 檢查一個隨機選中的局部約束：一行、一列、一個宮，或者一個 gadget。如果被揭開的格子顯示出的符號兩兩不同，Bob 就增加了幾分信心。然後一切又被蓋起來，符號也被重新命名一遍。（有一個小麻煩：謎題給定的提示需要一個

額外的技巧，因為給符號重命名也會把它們藏起來。下面的說明解釋了經典協議如何解決這個問題；對於接下來的內容，這個玩具式的圖景已經足夠了。）

**經典協議究竟如何處理提示格子**

重命名這個技巧有一個盲點。行、列、宮的規則都說「這些格子兩兩不同」，而兩兩不同在符號的任何重命名下都能保持。但一個提示說的是「這個格子裡恰好裝著 5」，而在重命名之後，Bob 只看到 $\sigma(5)$——某個被掩蓋的符號——卻不知道重命名 $\sigma$。他什麼都檢查不了。如果不加修補，Alice 就可以證明某個合法網格存在，卻完全無視印好的提示，而這對這個謎題什麼都證明不了。經典文獻裡有兩種標準的修補辦法。

**調色板**。給隱藏的網格加上額外的一行 N 個格子——一塊調色板，Alice 按照固定的、公開的順序把符號 1…N 填進去，然後和其他一切一起重命名，於是它裝著 $\sigma(1)…\sigma(N)$。現在 Bob 的隨機挑戰多了一個選項。除了挑一行、一列、一個宮或一個 gadget 來揭開之外，他還可以挑**這塊調色板加上一個提示格子**。Alice 把兩者都揭開；調色板揭示了那一輪的重命名，Bob 檢查那個提示格子顯示的恰好是印好的提示經過重命名後的版本。這仍然是零知識的，因為 Bob 只學到 $\sigma$——它每一輪都是新抽的，單獨來看毫無價值——以及一個他早已從謎題中知道其值的格子的值。關於秘密格子的任何資訊都沒有洩露，而一個模擬器可以通過抽一個隨機的 $\sigma$ 來偽造這個視圖。它是可靠的，因為一個作弊的 Alice 每一輪都以固定的機率被抓住，而且輪次會一直重複到疑慮可以忽略為止。

**把提示編譯掉**。一個更具結構性的變體，不是增加而是移除這個特殊的挑戰。與其驗證提示的值，不如用差異約束把它強制出來：把提示格子和除了承載它自己那個值之外的每一個調色板格子都連起來——「與 $\sigma(1)$ 不同、與 $\sigma(2)$ 不同、……、與除 $\sigma(5)$ 以外的一切都不同」。這個格子唯一能合法裝的符號就是提示的那個值。現在每一條約束又都是「這兩個不同」那一類了——在重命名下不變，可以完全像檢查一行那樣來檢查。這與經典圖著色協議中對預先著色的頂點所用的手法是同一個，也正是上面 *gadget* 一詞的精神所在：在「把超級數獨當作 SAT」的圖景裡，提示和其他每一條約束一樣，都被編譯成了不等式 gadget。

**物理協議**。數獨的現實世界紙牌協議（Gradwohl、Naor、Pinkas 和 Rothblum，2007）完全不用重命名，並且在藏起來之前就把提示搞定了。對每一個格子，Alice 放下三張寫著該格子值的相同的牌——秘密格子的牌面朝下，但**提示格子的牌面朝上**，這樣在牌被翻過來之前，Bob 就親眼看到提示得到了遵守。然後每個格子拿出一張牌放進它所在行的包，一張放進它所在列的包，一張放進它所在宮的包；每個包都被洗牌然後揭開，Bob 檢查它包含全部 N 個符號。洗牌摧毀了位置資訊（這就是零知識的所在），但提示在發牌的時候就已經被釘死了。

無論哪種辦法，教訓都是這篇文章反覆回到的那一個：一個零知識協議就是一份小心翼翼的賬目，記著哪些事實在隱藏之後還能存活。重命名保住了「兩兩不同」，抹去了「等於 5」——所以「等於 5」必須靠其他手段偷偷帶回來。

這並不是論文裡的協議。它是經典零知識的思維模型：

- Alice 和 Bob 來回往復。
- Bob 選擇隨機的檢查。
- Alice 只揭示局部的一致性，而不是整個解。
- 隱私性的證明是通過表明 Bob 所看到的視圖本可以在沒有 Alice 秘密解的情況下被生成來實現的。

所以經典零知識是圍繞一個肯定的事實建立起來的：

> 一個模擬器真的存在。

現在把那些令人舒適的部分去掉。Alice 發送一個證明字串，然後走開。沒有可信設置，沒有事先準備好的共享隨機字串，而且 Bob 絕不能接受一個假的謎題。這就是經典零知識無法存活的那個設定。

在這個訣竅之前，還需要一個角色。固定一個**規則手冊**：邏輯學家意義上的一個形式證明系統——一組固定的公理，加上用來檢查書面數學證明的機械規則。ZFC，即數學的標準公理，就是那個典範的例子。從這裡開始的一切，都是相對於一個事先選定的規則手冊來陳述的，而且這個選擇是靈活的：這個構造對你固定的任何規則手冊都成立，包括 ZFC。

(一個關於用詞的說明，借自論文本身：這裡的「證明系統」始終指的是這個規則手冊——那個檢查數學證明的形式系統——而絕不是 Alice 發送的消息。Alice 和 Bob 的那套機制被稱為「證明者和驗證者」。)

哥德爾風格的版本保留了超級數獨的故事，但改變了證明。

挑選第二個顯示尺寸相同的約束系統，把它叫作 **D**。就故事而言，S 和 D 是兩個格式相同的 MegaSudoku(n) 謎題。在幕後，D 可能一開始是一個尺寸不同的困難邏輯公式；如果需要，可以用一些無害的虛設約束把它填充起來，讓它適配同一個網格。D 是由一個實際上**不可滿足**的邏輯公式構造出來的：不存在任何一種可能的賦值能讓它所有的約束都為真，就像一個壞掉的謎題沒有合法的完成網格一樣。一個玩具式的例子會是一個既要求「X 為真」又要求「X 為假」的公式。所以 D 沒有合法的填法。

但 D 不能是一個容易被揭穿的壞掉的謎題。上面那個玩具例子就不合格：任何規則手冊都能用一行反駁「X 且非 X」。D 必須以一種所選規則手冊無法用簡短論證來證實的方式為假。如果規則手冊能用一個簡短的證明反駁 D，下面的故事就會崩塌：那條本可能在沒有 Alice 秘密的情況下產生證明的替代路線，就可以被形式地排除掉，隱私保證也隨之一起排除掉。所以 D 是從一個固定規則手冊無法高效反駁的族中挑選出來的：在那個規則手冊內部，不存在關於 D 無解的簡短證明。

於是 Alice 的一條消息證明所針對的是一個二選一的陳述：

> 要麼真正的超級數獨 S 有解，要麼誘餌 D 有解。

這就是那個邏輯上的聯繫。D **並不是**以某種能讓 S 為真的魔法方式生成的。這個證明並不是在論證「D 無解，因此 S 有解」。它證明的是析取式 **S 或 D**。完美可靠性說的是一個假的析取式不可能有有效的證明。既然 D 在現實中為假——它沒有解——那麼這個析取式能為真的唯一途徑就是 S 為真。所以如果證明被接受，S 就必然有解。誘餌沒法讓一個假的 S 變成真的。

但對於零知識風格的那一部分，問一問如果 D 確實有解會發生什麼。那個誘餌的解會充當一個替代的見證。它會讓某人在不知道 Alice 真正的超級數獨解的情況下也能產生證明——換句話說，就是一個模擬器。在現實中 D 沒有解，所以這條模擬器路線是封閉的。關鍵在於，規則手冊無法高效地證明它是封閉的。

所以 D 有兩項工作。對於**可靠性**，D 為假，因此一個關於「S 或 D」的有效證明強制 S 成立。對於**有效零知識**，D 難以反駁，因此規則手冊無法迅速排除那條本會讓模擬成為可能的誘餌路線。

所以安全性的檢驗不再是：

> 我們能否證明一個模擬器真的存在？

它變成了：

> 你的規則手冊能否高效地證明這個模擬器不可能存在？

如果答案是否定的，就會得出某種出人意料地強的結論：每一個滿足如下兩條的安全保證——(a) 可以通過運行一個測試來觀察到，並且 (b) 在那個規則手冊內部，可證明地從模擬器的存在推出——實際上都成立。對它們中任何一個的成功攻擊，本身就相當於那個缺失的簡短反駁，而那個缺失的簡短反駁並不存在。這就是「有效零知識」裡「有效」的那一部分。

所以課堂上的對比是：

**經典零知識：**這些證明是安全的，因為存在一個模擬器。

**哥德爾風格的有效零知識：**對於可觀察的安全性測試，這些證明被當作是安全的，因為規則手冊無法高效地證明這個模擬器不可能存在。

第二個斷言更弱。它也正是論文能夠保留那三個打垮了經典版本的特徵的原因：一條消息、無設置和完美可靠性。

# 新的檢驗：你無法證明模擬器不存在

Ilango 的鬆弛改變了問題。

經典零知識問的是：

> 是否存在一個模擬器？

有效零知識問的是某種更弱的東西：

> 你所選的規則手冊能否高效地證明不存在任何模擬器？

這聽起來像是一個技術性的躲閃，但它正是核心思想。這個構造處於一種奇怪的狀態：模擬器實際上並不存在——論文對這一點說得很明確——但你所固定的那個規則手冊無法高效地證明它不存在。如果你所關心的每一個壞後果都需要這樣一個反駁，那麼對於那些後果來說，這個系統仍然表現得像零知識。

這裡就是哥德爾登場的地方。不是作為裝飾，也不是「哥德爾讓密碼學變安全」。這個聯繫是證明論意義上的。一個規則手冊被稱為**最優的**，如果它在一個精確的意義上是可能的最好的那一個：每當任何一個規則手冊能用一個簡短的證明反駁某種相關類型的公式時，那個最優的規則手冊也能，用一個至多多項式地更長的證明。Krajíček 和 Pudlák 在 1989 年猜想**不存在最優證明系統**：無論你固定哪個規則手冊，總有另外某個規則手冊能遠為簡潔地證明某一族真陳述。這是證明複雜度的核心未解猜想

之一，也是哥德爾不完備性定理在有限的、複雜度理論意義上的表親：有些真陳述在你所固定的規則手冊裡沒有簡短的證明——不是因為它們原則上不可證明，而是因為每一個固定的規則手冊都會留下一些簡短的真理，它們沒有簡短的證明。

論文假設這個猜想（以一個稍強的「無窮頻繁」形式，這是把猜想用於密碼學時的標準做法）。根據 Krajíček 和 Pudlák 的一個定理，回報是具體的：對每一個規則手冊，都存在一個公式序列，它們確實不可滿足，規則手冊無法用簡短的證明反駁它們——而且，至關重要的是，一個高效的演算法能夠生成它們。最後那個性質，即一致性（uniformity），正是把整個想法從一個存在性斷言變成 Alice 真正能運行的一個演算法的東西：她的誘餌 D 是從流水線上下來的，而不是憑空變出來的。

密碼學上的這一手，就是把那種證明能力的短缺利用起來。

# 這個構造在做什麼

下面是論文的構造，被剝到只剩它的骨架形狀。

固定一個規則手冊——比如說 ZFC。在證明複雜度假設之下，存在一個可高效生成的公式序列，它們實際上不可滿足，但規則手冊沒有關於它們不可滿足的簡短證明。

現在構建一個如下形式的一條消息證明：

> 要麼真正的陳述可滿足，要麼這個特殊的困難公式可滿足。

這個特殊的困難公式是不可滿足的。所以如果底層的證明機制是完美可靠的，接受這條消息仍然意味著真正的陳述為真。這就給出了完美可靠性。

但對於類似零知識的安全性，想像這個特殊的困難公式確實可滿足。那麼它的見證就可以被用來在不知道真正見證的情況下模擬證明。這個公式在現實中不可滿足——但規則手冊無法高效地證明這一點。所以它無法高效地證明這個模擬器不可能存在。

這就是那個樞紐。這個系統不是通過產生一個經典的模擬器來隱藏秘密的。對於一大類可觀察的安全性測試來說，它把秘密藏在了規則手冊無力證實模擬器不存在的背後。

# 論文主張了什麼

主定理是分層的。核心結果是這樣的：

在一個標準的密碼學假設——**非交互見證不可區分證明**的存在，這是一類被充分研究過的對象，可從若干已確立的假設包中推出——之下，以及在**不存在（無窮頻繁）最優證明系統**這個證明複雜度猜想之下，論文為每一個規則手冊的選擇都構造了一個針對 NP/SAT 的一條消息證明者和驗證者，它具有**完美可靠性**、無設置，並且相對於那個規則手冊是有效零知識的。（NP/SAT 是各種謎題式問題的標準「最難公分母」；超級數獨是它所穿的一件戲服。）

對於那個關於保留可證偽安全性質的更寬泛的主張，論文再添加一個標準假設，即去隨機化的信念 **P**

**= BPP**（大致是說：隨機性沒有給演算法帶來任何本質上的額外能力）。

把定理的語言翻譯過來：

- 證明是一條消息。
- 沒有可信設置。
- 假的陳述無法被證明。
- 這個證明者不是經典零知識的——它沒有模擬器。
- 但經典零知識的每一個可證偽的、基於博弈的安全性後果，都可以在這個設定中實現。

「可證偽」這一點很重要。它意味著一次安全失敗可以通過在一場博弈中運行一個對手來測試。許多密碼學的安全定義都有這種形式：對手能否區分兩個密文、求逆一個函數、恢復一個見證，或者贏得某個指定的實驗？這個定理為每一個可證偽的性質各給出一個證明者，一次一個。一個同時享有每一個可證偽性質的單一證明者很可能是不可能的——舊的可重用性攻擊（「Bob 可以把證明拿給別人看」）本身就是一個可證偽的性質，而它在這裡確實失敗了。論文的提議是，一個單一的證明者有理由能夠涵蓋所有自然的可證偽性質——那些在密碼學實踐中真正出現的性質——但那一部分是一個條件性的定理，依賴於一個非正式的「自然」概念，外加一個明確的猜想。這個保證針對的是可觀察的失敗，而不是保密性的每一種哲學意義或基於模擬的意義。

有一個具體的推論值得點名：這個構造給出了第一個帶有一致證明者的非交互見證隱藏證明——「一個謎題的證明並不能幫你找到它的解」，沒有交互也沒有設置——這是一個聽起來不起眼、卻抵抗了數十年構造嘗試的對象。

# 這不是在說什麼

這是讓整篇文章保持誠實的那一節。

它**並不**是說舊的不可能性定理是錯的。這個構造是通過改變定義來繞開它們的。

它**並沒有**給出無交互、無設置和完美可靠性的普通的、經典的零知識。論文明確地說，所構造的證明者沒有模擬器。

它**並不**意味著這個證明不能被重用。一條消息的證明仍然可以拿給別人看；論文並沒有保留可否認性那一類的性質。（帶可信設置的非交互零知識也有同樣的局限。）

它**並不**意味著這是一個可以部署的實用協議。這是複雜度理論和密碼學基礎。這個結果依賴於來自證明複雜度和密碼學的重大假設，而這個構造關乎的是原則上什麼是可能的。

它**並沒有**把「哥德爾」變成一個魔法般的安全原語。哥德爾的聯繫是通過證明系統、最優證明系統以及不完備性的有限類比來實現的。可用的直覺不是「不完備性保護你的密碼」。而是：如果一個規則手冊無法高效地證明一個模擬器不可能存在，那麼那些需要該證明的攻擊就可以在安全定義的層面被擋住。

# 為什麼它無論如何都很有趣

密碼學常常把困難性轉化為安全性。因數分解是困難的，所以 RSA 式的假設變得有用。格問題是困難的，所以格密碼學變得有用。這裡的困難性更奇怪：不是「難以計算出一個秘密」，而是「難以證明某個證明對象不可能存在」。

這就是為什麼這篇論文讓人感覺不同尋常。它幾乎把公理和規則手冊當作密碼學資源來對待。通常的不可能性說的是，可靠性和模擬之間存在一種張力。Ilango 這一手，是把這種張力放到一道證明論的帷幕後面：模擬器不存在，但形式系統無法高效地揭露這種不存在。

對一個讀者來說，令人驚訝的部分並不是這將取代今天的零知識系統。它大概不會，至少不會直接取代。令人驚訝的部分在於，一個來自數理邏輯的局限性可以被建設性地使用：不只是作為一堵牆，而是作為一種掩護。

# 證據有多強？

這是一篇定理論文，所以「證據」的含義不同於一篇生物學或天文學論文。問題不在於一個實驗是否被複現。問題在於那些定義、假設和證明鏈條是否支持這個主張。

證明是形式化的，而且論文對它的假設說得很明確。這些假設不是隨隨便便的。非交互見證不可區分證明是密碼學中的標準對象，可從若干已確立的假設包中推出。不存在最優證明系統這個猜想是證明複雜度中的一個核心猜想。P = BPP 是一個標準的去隨機化信念，只用於那個更寬泛的可證偽性質定理。

論文還論證說，這些假設是恰當的代價，而不是一個隨意搭起來的腳手架：它證明了一個逆命題，表明這些假設本質上是必要的——如果像這樣的構造根本存在，那麼非交互見證不可區分證明就必然存在，並且（在承認標準單向函數的前提下）不可能存在最優證明系統。而且這些假設是「雙贏」的：反駁它們中的任何一個，本身就會是證明複雜度、密碼學或複雜度理論中的一個里程碑式的發現。

但由於這個結果是條件性的，它的可信度也是條件性的。如果那些假設不成立，定理的解讀就會改變。而且即便那些假設成立，這個保證也不是完整的經典零知識；它是論文那個鬆弛過的、證明論意義上的版本。

所以恰當的可信度是：對於論文確立了一個自洽的條件性可能性結果這一點，是高的；對於它的假設描述了我們實際所處的密碼學世界這一點，是中等的；而對於任何即時的實用後果，則是低的。

# 為什麼這很重要

這篇論文打開了一條本應封閉的路線。

經典理論說：完整的零知識不可能是無設置的一條消息，也不可能是完美可靠的。Ilango 的論文說：如果我們要的是零知識那些可以在安全博弈中被測試的後果，並且如果我們允許安全定義取決於一個規則手冊能或不能高效反駁什麼，那麼許多有用的行為就可以被恢復——用一條消息、無設置和完美

可靠性。

這並不是一個小小的定義上的微調。它是一種思考密碼學保證的不同方式。不要只問什麼存在，而要問你的規則手冊能排除什麼。不要把不可證明性當作一種哲學上的煩擾，而要把它當作一種結構來使用。

實用的世界可能不會在明天就改變。但概念上的地圖變了。現在有了一種形式上的意義，在這種意義下，「沒有人能高效地證明秘密洩露了」可以強到足以恢復許多我們本想從「秘密沒有洩露」中得到的、基於博弈的保護。

這就是為什麼哥德爾理應出現在標題裡。

# 乾淨的總結

零知識證明讓一個證明者能夠說服一個驗證者某個陳述為真，而不洩露見證。經典的不可能性結果說，零知識不可能被塞進無設置的一條消息裡，也不可能具有完美可靠性。Rahul Ilango 的論文並沒有反駁那些不可能性。它定義了一個更弱的概念，即有效零知識：它不是要求一個模擬器真的存在，而是要求一個所選的證明系統——一個像 ZFC 那樣的形式規則手冊——無法高效地證明不存在任何模擬器。在來自密碼學（非交互見證不可區分證明）和證明複雜度（不存在最優證明系統）的重大假設之下，論文為 NP/SAT 構造了無設置、完美可靠的一條消息證明者，它們逐個性質地實現了零知識那些可證偽的、基於博弈的後果。一個單一的證明者涵蓋所有「自然的」這類性質，是一個更進一步的、部分帶有猜想性的擴展——而要涵蓋字面意義上每一個可證偽的性質，則很可能是不可能的，因為證明仍然是可重用的。這個結果是理論性的、條件性的，不是一個已部署的原語，但它展示了一種把證明論意義上的不可證明性用作密碼學資源的新方法。

# 不扯淡核查

**論文展示了什麼：**在所陳述的假設之下，人們可以為 NP/SAT 構建一條消息、無設置、完美可靠的證明者，它們相對於任何所選的證明系統都是有效零知識的，並且實現了經典零知識每一個可證偽的、基於博弈的後果。

**什麼是有道理但並未被無條件證明的：**所需要的證明複雜度和密碼學假設成立。它們是嚴肅的、被充分研究過的假設——而且論文表明它們不僅充分，還本質上是必要的——但仍然是假設。

**它沒有展示什麼：**無交互、無設置和完美可靠性的經典零知識；一個可以部署的實用系統；證明的可否認性或不可重用性；或者哥德爾不完備性定理本身就能保障密碼學的安全。

**主要局限：**這個保證是零知識的一個鬆弛；最寬泛的版本依賴於多個假設；單一通用證明者的那些主張仍然部分帶有猜想性；而且這個結果主要是基礎性的。

**一個普通讀者應該有多少信心？**如果接受那些定義，那麼對於這是一個重要的條件性理論結果這一點，信心是高的。對於這些假設抓住了現實這一點，是中等的。對於即時的實用部署，是低的。安全的要點是：這篇論文並沒有打破零知識的不可能性；它找到了一種新的、證明論意義上的方法，繞過

它們當中對許多安全博弈來說重要的那些部分。

## 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