The Clean Paper

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

Kriptogrāfisks pierādījums var noslēpt savu noslēpumu, padarot neesošа simulatora pierādīšanu grūtu

Nulles zināšanu pierādījumi ļauj pierādīt apgalvojumu, neatklājot liecinieku. Klasiskā teorija saka, ka pilnā nozīmē to nevar iegūt ar vienu ziņojumu, bez uzticamas sākotnējās iestatīšanas un ar perfektu drošumu pret nepatiesiem apgalvojumiem. Rahul Ilango darbs šo neiespējamību neatceļ. Tas maina mērķi: tā vietā, lai prasītu, lai simulators patiešām eksistētu, pietiek ar to, ka izvēlētā formālā pierādījumu sistēma nespēj efektīvi pierādīt, ka simulatora nav. Pie nozīmīgiem pieņēmumiem par pierādījumu sarežģītību un kriptogrāfiju ar to pietiek, lai pa vienai atgūtu falsificējamas, spēlēs pārbaudāmas nulles zināšanu drošības sekas. Tas nav gatavs interneta protokols. Tā ir pierādījumu teorijas pieeja, kas matemātisku nepierādāmību pārvērš par kriptogrāfisku aizsegu.

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

Triks nav pierādīt, ka noslēpums ir paslēpts

Sāksim ar vienkāršāko nulles zināšanu variantu.

Alise grib pārliecināt Bobu, ka Sudoku mīklai ir atrisinājums. Ja viņa nosūta atrisināto režģi, Bobs ir pārliecināts, taču mīkla vairs nav noslēpums. Alise grib kaut ko dīvainākū: pierādījumu, ka atrisinājums eksistē, neatklājot pašu atrisinājumu.

Tieši to sola **nulles zināšanu pierādījums**. Pierādītājs (Alise) pārliecina pārbaudītāju (Bobu), ka apgalvojums ir patiess, neatklājot neko vairāk par pašu faktu, ka tas ir patiess.

Problēma ir tā, ka šādam solījumam ir cena. Parastam matemātiskam pierādījumam ir divas ērtas īpašības. Tas ir **viens ziņojums**: pieraksti pierādījumu, iedod to otram un vari doties prom. Un tas ir **perfekti drošs pret nepatiesu apgalvojumu pierādīšanu**: nepatiesam apgalvojumam vispār neeksistē derīgs pierādījums. Klasiskie neiespējamības rezultāti rāda, ka nulles zināšanām no šīm īpašībām jāatsakās — turklāt ne tikai no abām reizē; katra atsevišķi ir nepieejama.

Pirmkārt, nulles zināšanu pierādījumam vajadzīga saruna. Ja Alise nosūta tikai vienu ziņojumu un iepriekš nav sagatavota uzticama kopīga iestatīšana, nulles zināšanu garantija sabrūk — neatkarīgi no tā, cik daudz drošuma pret nepatiesiem pierādījumiem jūs būtu gatavi upurēt.

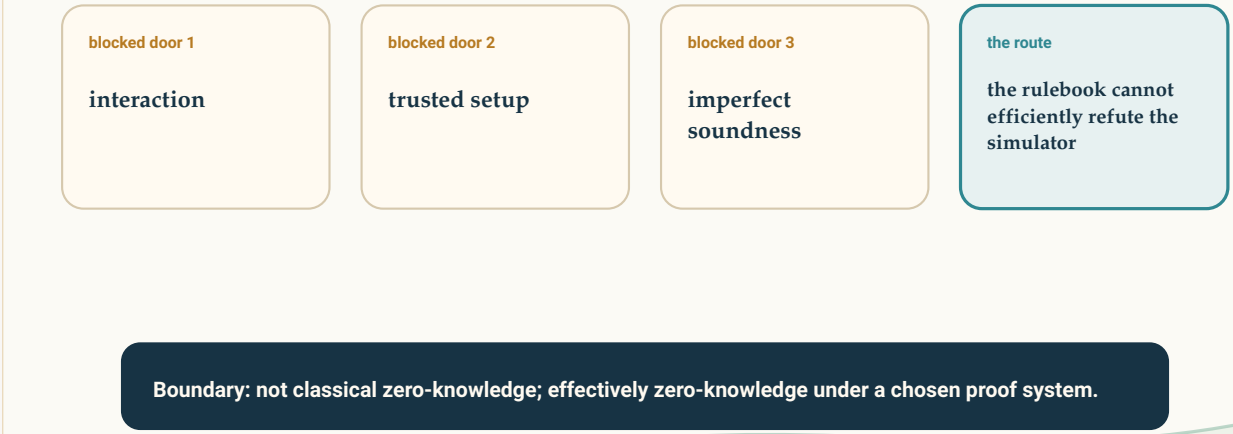
Otrkārt, nulles zināšanu pierādījumam vajadzīga neliela kļūdas tolerance. Prasība pēc perfekta drošuma klusi iznīcina arī mijiedarbības priekšrocību: ja pārbaudītāju nevar piemānīt pie nevienas no tā nejaušajām izvēlēm, tad tas tikpat labi šīs izvēles var fiksēt jau iepriekš. Bet, tiklīdz pārbaudītājs kļūst paredzams, Alise var atbildēt uz visu vienā ziņojumā — tieši tajā gadījumā, kas jau iepriekš izrādījās neiespējams.

Rahul Ilango raksts meklē ceļu garām šai dubultajai sienai. Nevis izliekoties, ka sienas nav, un nevis iegūstot klasiskas nulles zināšanas apstākļos, kuros tas nav iespējams. Gājiens ir smalkāks: vājināt to, ko nozīmē “neatklāj neko”, bet darīt to tā, lai saglabātu tās drošības īpašības, kuras kriptogrāfi patiešām var pārbaudīt.

Rezultātu autors sauc par **efektīvām nulles zināšanām** (*effectively zero-knowledge*).

A proof without leaking the witness

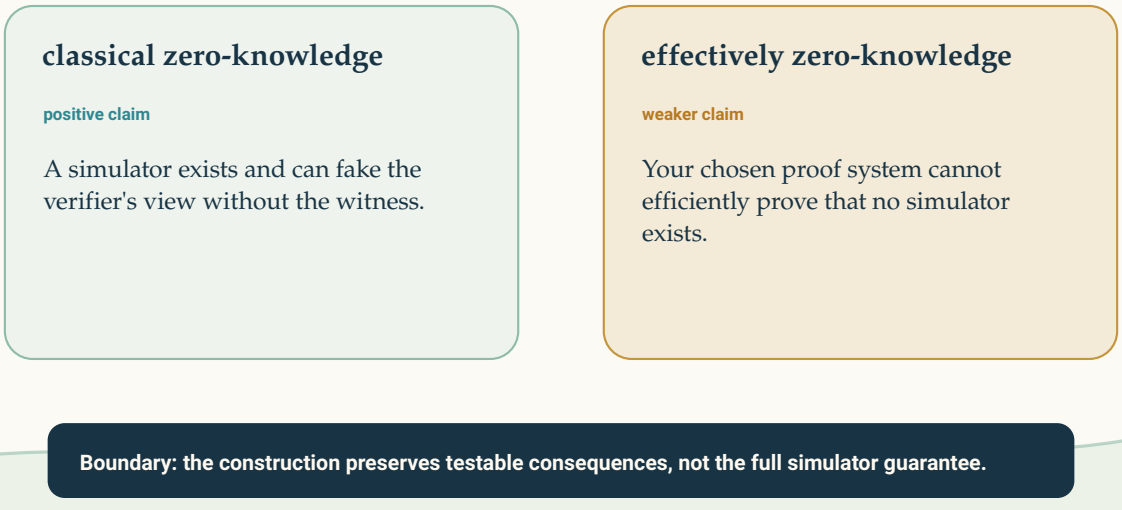
The paper keeps the ordinary-proof shape by weakening what privacy must prove.



Klasiskās nulles zināšanas aizsprosto trīs prasības — viena ziņa bez mijiedarbības, nekāda uzticama sākotnējā iestatīšana un perfekts drošums. Ilango konstrukcija izmanto citu ceļu: formālā noteikumu sistēma nespēj efektīvi pierādīt, ka simulatora nav.

Classical privacy vs effective privacy

The relaxation is the point: the paper changes what the privacy claim has to establish.



Klasiskās nulles zināšanas jautā, vai simulators eksistē; "efektīvās nulles zināšanas" jautā tikai, vai izvēlētā

formālā pierādījumu sistēma spēj efektīvi pierādīt, ka tas nevar eksistēt. Tieši šī vājākā prasība ļauj saglabāt vienu ziņojumu, nekādu sākotnējo iestatīšanu un perfektu drošumu.

Vecais kritērijs: simulators eksistē

Klasiskajā nulles zināšanu formalizācijā izmanto iedomātu palīgu, ko sauc par **simulatoru**.

Ideja ir šāda: iedomājieties Džeinu, kura **nezina** Alises noslēpumu. Ja Džeina pilnīgi pati var ģenerēt pierādījumus, kas izskatās tāpat kā tie, ko Bobs būtu saņēmis no Alises, tad Alises pierādījumi Bobam nav iemācījuši neko jaunu. Džeina jau bez Alises noslēpuma spētu viltot visu šo pieredzi.

Tāpēc klasiskās nulles zināšanas prasa reāli eksistējošu simulatoru. Jābūt efektīvam algoritmam, kas var ģenerēt ticamus "viltus" pierādījumus, nezinot noslēpumu — kriptogrāfijas terminoloģijā **liecinieku** (*witness*); Sudoku gadījumā liecinieks ir vienkārši pilnībā aizpildīts pareizais režģis.

Šī definīcija ir spēcīga, bet tieši tajā iekožas vecais neiespējamības rezultāts. Intuīcija ir vienkārša. Patiesi neinteraktīvs pierādījums ir tikai simbolu virkne. Kad Bobs to ir saņēmis, viņš var to parādīt kādam citam: tātad viņš ir ieguvis spēju pierādīt apgalvojumu tālāk, un tas jau izklausās pēc kaut kā vairāk nekā "nekā". Klasiskās teorēmas šo intuīciju precizē līdz iepriekš minētajiem neiespējamības rezultātiem.

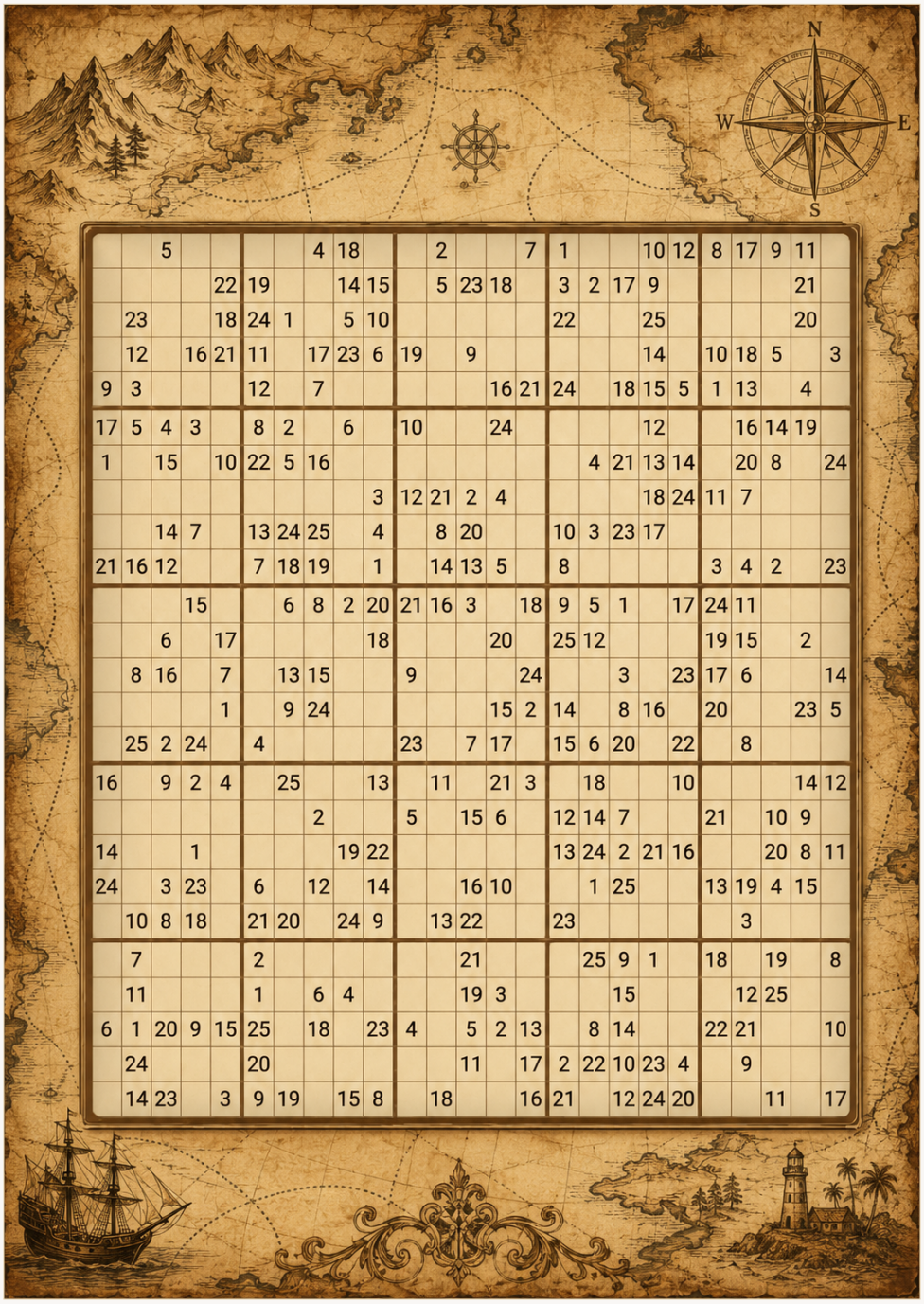
> **Trīs īpašības, kuras šis darbs neatlaiž**
>
> Raksta nosaukums izceļ trīs ierobežojumus:
> **Bez mijiedarbības:** Alise nosūta vienu pierādījuma virkni. Nav turp-atpakaļ protokola.
> **Bez sākotnējas iestatīšanas:** Alise un Bobs nepaļaujas uz uzticamu kopēju atsauces virkni vai citu iepriekš sagatavotu publisku nejaušību. Daudzas sistēmas, ko sauc par "neinteraktīviem nulles zināšanu pierādījumiem", tomēr izmanto iestatīšanas posmu; šeit iestatīšanas nav vispār.
> **Perfekts drošums:** nepatiesam apgalvojumam nav derīga pierādījuma. Nevis "to gandrīz nekad nepieņem", bet gan — derīgs pierādījums neeksistē.
> Tieši šīs trīs īpašības ir parastam rakstītam matemātiskam pierādījumam — un, kā skaidrots iepriekš, klasiskās nulles zināšanas tās visas saglabāt nevar.

Atšķirība MegaSudoku valodā

Lūk, apzināti vienkāršota analoģija, kas palīdz sajust atšķirību.

Nopietnajai analoģijas daļai neder parasts 9×9 Sudoku. Tas ir pārāk mazs un galīgs: dators to var vienkārši atrisināt vai pierādīt, ka atrisinājuma nav. Tā vietā iedomāsimies **MegaSudoku(n)** mīklu saimi. Vispārinām ierasto noteikumu: izvēlamies bloka izmēru n, liekam $N = n^2$ un veidojam $N \times N$

režģi, kas sadalīts $n \times n$ blokos un izmanto N simbolus. Parastais Sudoku ir tikai mazais $n = 3$, $N = 9$ gadījums: 9×9 režģis, 3×3 bloki un deviņi simboli. Pierādījumu sarežģītības stāsts sākas tikai tad, kad n drīkst augt un režģī var ievietot papildu konstrukcijas jeb "gadžetus", kas liek tam uzvesties kā SAT formulai Sudoku izskatā. SAT formula ir vienkārši "jā/nē" ierobežojumu saraksts: vai mainīgajiem var piešķirt patiesuma vērtības tā, lai visi ierobežojumi būtu izpildīti?



25×25 Sudoku: tā noteikumus var pārbaudīt, neatklājot aizpildīto režģi — vizuāls analogs pierādījumam, kas pārbauda slēptu atrisinājumu jeb liecinieku.

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

Sudoku un SAT: viena problēma divos tērpos

Apgalvojums, ka Sudoku var "uzvesties kā SAT formula", nav metafora. Pārveidošana darbojas abos virzienos, un vieglāko virzienu var uzrakstīt pilnībā.

No Sudoku uz SAT. SAT runā tikai patiesi/nepatiesi valodā, tāpēc katram trijniekam (rinda, kolonna, vērtība) ieviešam vienu Būla mainīgo: *x(r,c,v)* nozīmē "šūnā rindā *r*, kolonnā *c* atrodas vērtība *v*". 4×4 Sudoku (2×2 bloki, vērtības 1–4) vajag 4·4·4 = 64 mainīgos; klasiskajam 9×9 vajag 729. Katrs Sudoku noteikums tad pārvēršas klauzulu kopā. (Klauzula ir mainīgo vai to noliegumu OR; visa formula ir visu klauzulu AND.)

Katrā šūnā ir vismaz viena vērtība — viena klauzula katrai šūnai:

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

Katrā šūnā ir ne vairāk par vienu vērtību — katram vērtību pārim klauzula "ne abas reizē":

$$\neg x(1,1,1) \vee \neg x(1,1,2) \quad \neg x(1,1,1) \vee \neg x(1,1,3) \quad \text{… un tā tālāk visiem sešiem pāriem.}$$

Katrā rindā ir katra vērtība — 1. rindai un vērtībai 3: vismaz vienu reizi,

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

un ne vairāk kā vienu reizi: $\neg x(1,1,3) \vee \neg x(1,2,3)$, un tā tālāk katram šūnu pārim rindā.

Kolonnas un bloki — tieši tādas pašas klauzulu paketes; mainās tikai šūnu grupa. Augšējam kreisajam blokam un vērtībai 2:

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

plus pāru klauzulas "ne abas reizē".

Drukātie sākuma skaitļi — vienkāršākā daļa: katrs dotais skaitlis ir klauzula ar vienu mainīgo. Drukāts 3 augšējā kreisajā stūrī kļūst par klauzulu

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

Visu šo nosacījumu AND ir izpildāms tieši tad, ja Sudoku ir atrisinājums — un izpildošs mainīgo piešķirums *ir* atrisinājums: nolasa, kuri *x(r,c,v)* ir patiesi, un aizpilda režģi. 9×9 Sudoku tas nozīmē 729 mainīgos un dažus tūkstošus klauzulu, ko mūsdienīgs SAT risinātājs apstrādā milisekundēs. Pievērsiet uzmanību dotā skaitļa klauzulai x(1,1,3): tā saka "šī šūna ir tieši 3", nevis "šīs šūnas visas ir atšķirīgas". Tieši šī asimetrija vēlāk piespiedīs izmantot papildu triku ar dotajām šūnām.

No SAT uz Sudoku. Rakstam vajadzīgs pretējais, grūtākais virziens: no *patvaļīgas* SAT formulas uzbūvēt MegaSudoku, kam ir atrisinājums tieši tad, ja formula ir izpildāma. Sudoku pamatnoteikumi prot pateikt tikai "šīs šūnas ir visas atšķirīgas", tāpēc patvaļīgi loģiski ierobežojumi ir *jāuzbūvē* — tam arī domāti gadžeti. Gadžets ir neliels iepriekš izstrādāts šūnu klasteris, pa vienam formulas klauzulai, kurā noteiktas šūnas spēlē mainīgo lomu (simbols tajās kodē patiesi vai nepatiesi), bet iekšējie ierobežojumi ir izveidoti tā, lai legālie aizpildījumi atbilstu tieši tiem piešķirumiem, kas klauzulu apmierina. Tas ir standarta paņēmiens NP-pilnības pierādījumos; vispārinātam Sudoku šādu konstrukciju 2003. gadā izstrādāja Yato un Seta.

Abi virzieni kopā nozīmē, ka N×N Sudoku un SAT pēc sarežģītības ir viena un tā pati problēma dažādos tērpos. Tas ļauj gan šim rakstam, gan pašam zinātniskajam darbam stāstīt par visu NP ar režģu un simbolu palīdzību.

Liecinieku joprojām ir viegli iztēloties. Alise zina pilnīgi pareizu MegaSudoku aizpildījumu. Bobs grib pārliecināties, ka šāds aizpildījums eksistē, bet Alise negrib to atklāt. Ja viņa nosūta visu režģi, Bobs ir pārliecināts, taču noslēpums ir zudis.

Klasiskajā nulles zināšanu variantā Alise un Bobs mijiedarbojas. Viena veca un intuitīva analoģija izmanto aizklātas kartītes. Alise paslēpj atrisināto režģi, pirms katras kārtas slepeni pārdēvē simbolus un ļauj Bobam apskatīt vienu nejauši izvēlētu lokālu ierobežojumu: rindu, kolonnu, bloku vai gadžetu. Ja atklātajās šūnās redzami savstarpēji atšķirīgi simboli, Boba pārliecība pieaug. Tad viss atkal tiek aizklāts, un nākamajā kārtā simbolus pārdēvē no jauna. (Ir viena nianse: sākotnēji dotajiem mīklas skaitļiem vajag papildu triku, jo simbolu pārdēvēšana paslēpj arī tos. Nākamā piezīme paskaidro, kā klasiskie protokoli to risina; mūsu vienkāršotajam stāstam ar to pietiek.)

Kā klasiskie protokoli patiesībā pārbauda dotās šūnas

Pārdēvēšanas trikam ir aklā zona. Rindu, kolonnu un bloku noteikumi visi saka "šīs šūnas ir savstarpēji atšķirīgas", un īpašība *visi atšķirīgi* saglabājas pie jebkuras simbolu pārdēvēšanas. Taču dotais skaitlis saka "šajā šūnā ir tieši 5", un pēc pārdēvēšanas Bobs redz tikai $\sigma(5)$ — kādu maskētu simbolu — nezinot pārdēvēšanu σ. Viņš to nevar pārbaudīt. Ja šo problēmu neatrisinātu, Alise varētu pierādīt, ka eksistē *kāds* derīgs režģis, pilnīgi ignorējot uzdrukātos sākuma skaitļus, un tas neko nepierādītu par *šo* konkrēto mīklu. Klasiskajā literatūrā ir divi tipiski risinājumi.

Palete. Slēptajam režģim pievieno vēl vienu rindu ar N šūnām — paleti, ko Alise publiski zināmā secībā aizpilda ar simboliem 1…N un pēc tam pārdēvē kopā ar visu pārējo, tāpēc tajā atrodas $\sigma(1)…\sigma(N)$. Boba nejaušajam izaicinājumam tagad ir vēl viena iespēja. Papildus rindas, kolonnas, bloka vai gadžeta atvēršanai viņš var izvēlēties **paleti plus vienu doto šūnu**. Alise atklāj abas; palete parāda šīs kārtas pārdēvēšanu, un Bobs pārbauda, vai dotajā šūnā ir tieši uzdrukātā skaitļa pārdēvētā versija. Nulles zināšanu īpašība saglabājas, jo Bobs uzzina tikai σ — tas katrā kārtā tiek izvēlēts no jauna un pats par sevi neko neatklāj — un vērtību šūnā, kuru viņš jau zināja no mīklas. Nekas par slepenajām šūnām nenoplūst, un simulators šo skatu var atdarināt, izvēloties nejaušu σ. Drošums rodas tāpēc, ka krāpjošu Alisi katrā kārtā ar noteiktu varbūtību pieķer, un kārtas atkārto, līdz kļūdas iespēja kļūst niecīga.

Doto skaitļu "iekompilēšana" ierobežojumos. Strukturālāks variants īpašo pārbaudi vispār noņem. Tā vietā, lai tieši *pārbaudītu* dotās šūnas vērtību, to piespiež ar nevienādības ierobežojumiem: doto šūnu sasaista ar katru paletes šūnu, izņemot to, kurā atrodas tās pašas vērtības simbols — "atšķiras no $\sigma(1)$, atšķiras no $\sigma(2)$, …, atšķiras no visa, izņemot $\sigma(5)$". Vienīgais simbols, ko šūna drīkst legāli saturēt, ir dotais. Visi ierobežojumi atkal ir tipa "šīs divas vērtības atšķiras" — tie nemainās pie simbolu pārdēvēšanas un ir pārbaudāmi tāpat kā rindas noteikums. Tas ir tas pats paņēmiens, ko klasiskajā grafa krāsošanas protokolā izmanto iepriekš iekrāsotiem virsotņu punktiem, un tas atbilst iepriekš minētajai gadžetu idejai: MegaSudoku kā SAT attēlojumā dotie skaitļi tiek iekompilēti nevienādību gadžetos gluži tāpat kā citi ierobežojumi.

Fiziskais protokols. Reāls Sudoku kartīšu protokols (Gradwohl, Naor, Pinkas un Rothblum, 2007) vispār neizmanto simbolu pārdēvēšanu un ar dotajiem skaitļiem tiek galā jau pirms slēpšanas. Katrai šūnai Alise noliek trīs identiskas kartītes ar šūnas vērtību — slepenajām šūnām ar attēlu uz leju, bet **dotajām šūnām ar attēlu uz augšu**, lai Bobs pats redzētu, ka

> sākuma skaitļi ievēroti, pirms kartītes apgriež. Tad pa vienai kartītei no katras šūnas nonāk rindas paketē, pa vienai — kolonnas paketē un pa vienai — bloka paketē; katru paketi sajauc un atklāj, un Bobs pārbauda, ka tajā ir visi N simboli. Sajaukšana iznīcina informāciju par pozīcijām — tā ir nulles zināšanu daļa — bet dotie skaitļi jau bija nostiprināti kartīšu izlikšanas brīdī.
> Abos variantos mācība ir tā pati, pie kuras šis raksts atgriežas vairākkārt: nulles zināšanu protokols ir rūpīga uzskaite par to, *kuri* fakti izdzīvo paslēpšanu. Pārdēvēšana saglabā "visi atšķirīgi", bet izdzēš "vienāds ar 5", tāpēc "vienāds ar 5" jāatjauno citā veidā.

Tas nav Ilango rakstā izmantotais protokols. Tā ir tikai intuitīva aina klasiskajām nulles zināšanām:

- Alise un Bobs sazinās turp un atpakaļ.
- Bobs izvēlas nejaušas pārbaudes.
- Alise atklāj tikai lokālu konsekvenci, nevis visu atrisinājumu.
- Privātuma pierādījums balstās uz to, ka Boba redzēto varētu ģenerēt arī bez Alises slepenā atrisinājuma.

Tātad klasiskās nulles zināšanas balstās pozitīvā faktā:

> Simulators patiešām eksistē.

Tagad noņemam ērtās īpašības. Alise nosūta vienu pierādījuma virkni un aiziet. Nav uzticamas sākotnējās iestatīšanas, nav iepriekš sagatavotas kopīgas nejaušas virknes, un Bobs nekad nedrīkst pieņemt nepatiesu mīklu. Tieši šādos apstākļos klasiskās nulles zināšanas nevar izdzīvot.

Pirms galvenā trika vajag vēl vienu tēlu. Fiksēsim **noteikumu grāmatu**: formālu pierādījumu sistēmu loģiķu nozīmē — noteiktu aksiomu kopu un mehāniskus noteikumus rakstītu matemātisku pierādījumu pārbaudei. Kanoniskais piemērs ir ZFC, standarta matemātikas aksiomas. Viss turpmākais tiek formulēts attiecībā pret iepriekš izvēlētu noteikumu grāmatu, un izvēle ir elastīga: konstrukcija darbojas jebkurai fiksētai sistēmai, arī ZFC.

(Vārdu lietojuma piezīme, pārņemta no paša raksta: "pierādījumu sistēma" šeit vienmēr nozīmē šo noteikumu grāmatu — formālo sistēmu, kas pārbauda matemātiskus pierādījumus — un nekad ne Alises sūtītos ziņojumus. Alises un Boba mehānismus saucam par "pierādītāju" un "pārbaudītāju".)

Gēdeļa stila variants saglabā MegaSudoku stāstu, bet maina pašu pierādījumu.

Izvēlamies otru tāda paša parādītā izmēra ierobežojumu sistēmu un nosaucam to par **D**. Stāstā S un D ir divas vienā formātā uzrakstītas MegaSudoku(n) mīklas. Zem virsmas D var būt sākusies kā cita izmēra sarežģīta loģiska formula; ja vajag, to papildina ar nekaitīgiem tukšiem ierobežojumiem, lai tā ietilptu tādā pašā režģī. D ir uzbūvēta no loģiskas formulas, kas patiesībā ir **neizpildāma**: nav iespējams piešķirt vērtības tā, lai visi tās ierobežojumi būtu patiesi, gluži kā bojātai mīklai nav legāla aizpildījuma. Rotaļlietas piemērs būtu formula, kas reizē prasa "X ir patiess" un "X ir nepatiess". Tātad D nav derīga atrisinājuma.

Taču D nedrīkst būt tāda bojāta mīkla, kuras bojājumu ir *viegli pierādīt*. Iepriekšējais rotaļlietas

piemērs neder: jebkura saprātīga formāla sistēma “X un ne-X” atspēkotu vienā rindā. D jābūt nepatiesai tādā veidā, ko izvēlētā noteikumu sistēma nespēj apliecināt ar īsu argumentu. Ja tā spētu D atspēkot ar īsu pierādījumu, tālāk aprakstītais stāsts sabruktu: formāli varētu izslēgt alternatīvo ceļu, pa kuru pierādījumus teorētiski varētu veidot bez Alises noslēpuma, un līdz ar to pazustu privātuma garantija. Tāpēc D izvēlas no formulas saimes, kuru fiksētā sistēma nespēj efektīvi atspēkot: šajā sistēmā nav īsa pierādījuma, ka D nav atrisinājuma.

Alises vienas ziņas pierādījums tad attiecas uz “vai nu/vai” apgalvojumu:

> vai nu īstajam MegaSudoku S ir atrisinājums, vai arī māneklim D ir atrisinājums.

Te ir loģiskā saite. D **nav** kaut kā maģiski ģenerēts tā, lai padarītu S patiesu. Pierādījums nesaka “D nav atrisinājuma, tātad S ir atrisinājums”. Tas pierāda disjunkciju **S vai D**. Perfekts drošums nozīmē, ka nepatiesai disjunkcijai nevar būt derīgs pierādījums. Tā kā D realitātē ir nepatiesa — tai nav atrisinājuma — vienīgais veids, kā disjunkcija var būt patiesa, ir tas, ka S ir patiesa. Tātad, ja pierādījums tiek pieņemts, S jābūt atrisināmai. Māneklis nevar padarīt nepatiesu S par patiesu.

Taču nulles zināšanām līdzīgajā daļā pajautāsim, kas notiktu, *ja D tomēr būtu atrisinājums*. Šis mānekļa atrisinājums kalpotu kā alternatīvs liecinieks. Tas ļautu veidot pierādījumus, nezinot Alises īsto MegaSudoku atrisinājumu — tātad darbotos kā simulators. Realitātē D nav atrisinājuma, tāpēc šis simulācijas ceļš ir slēgts. Galvenā doma ir tāda, ka noteikumu grāmata nespēj efektīvi pierādīt, ka tas ir slēgts.

Tātad D pilda divus uzdevumus. **Drošumam** D ir nepatiesa, tāpēc derīgs “S vai D” pierādījums piespiež S būt patiesai. **Efektīvajām nulles zināšanām** D ir grūti atspēkojama, tāpēc noteikumu sistēma nevar ātri izslēgt mānekļa ceļu, kas būtu padarījis simulāciju iespējamu.

Drošības kritērijs vairs nav:

> Vai varam pierādīt, ka simulators patiešām eksistē?

Tas kļūst par:

> Vai jūsu noteikumu sistēma spēj efektīvi pierādīt, ka simulators nav iespējams?

Ja atbilde ir nē, seko pārsteidzoši spēcīga sekas: katra drošības garantija, kuru (a) var novērot, izpildot konkrētu testu, un kura (b) šajā noteikumu sistēmā pierādāmi seko no simulatora eksistences, patiešām ir spēkā. Veiksmīgs uzbrukums jebkurai no šīm īpašībām pats dotu trūkstošo īso atspēkojumu — bet šāda īsa atspēkojuma nav. Tieši tas ir vārda “efektīvi” saturs jēdzienā “efektīvas nulles zināšanas”.

Tātad mācību klases kontrasts ir šāds:

Klasiskās nulles zināšanas: pierādījumi ir droši, jo simulators eksistē.

Gēdeļa stila efektīvās nulles zināšanas: novērojamajos drošības testos pierādījumus var uzskatīt par drošiem, jo noteikumu sistēma nespēj efektīvi pierādīt, ka simulators nav iespējams.

Otrais apgalvojums ir vājāks. Tieši tāpēc raksts var saglabāt trīs īpašības, kas salauza klasisko variantu: vienu ziņojumu, nekādu sākotnējo iestatīšanu un perfektu drošumu.

Jaunais kritērijs: nevar pierādīt, ka simulatora nav

Ilango relaksācija maina jautājumu.

Klasiskās nulles zināšanas jautā:

> Vai simulators eksistē?

Efektīvās nulles zināšanas jautā kaut ko vājāku:

> Vai izvēlētā noteikumu sistēma spēj efektīvi pierādīt, ka simulatora nav?

Tas var izklausīties pēc tehniskas izvairīšanās, taču tā ir centrālā ideja. Konstrukcija atrodas dīvainā stāvoklī: simulators patiesībā neeksistē — raksts to pasaka skaidri —, bet fiksētā formālā sistēma nespēj efektīvi pierādīt, ka tas neeksistē. Ja katra nevēlamā sekas, kas jums rūp, prasītu tieši šādu atspēkojumu, sistēma attiecībā uz šīm sekām joprojām uzvedas kā nulles zināšanu sistēma.

Te parādās Gēdelis. Nevis kā dekorācija un nevis nozīmē "Gēdelis padara kriptogrāfiju drošu". Saikne ir pierādījumu teorijā. Noteikumu sistēmu sauc par **optimālu**, ja tā precīzā nozīmē ir labākā iespējamā: ja kāda cita sistēma spēj attiecīgā tipa formulu atspēkot ar īsu pierādījumu, optimālā sistēma to spēj izdarīt ar pierādījumu, kas ir ne vairāk kā polinomiāli garāks. Krajíček un Pudlák 1989. gadā izvirzīja hipotēzi, ka **optimāla pierādījumu sistēma neeksistē**: lai kuru formālo sistēmu jūs fiksētu, kāda cita sistēma kādu patiesu apgalvojumu saimi pierādīs daudz īsāk. Tā ir viena no centrālajām atvērtajām hipotēzēm pierādījumu sarežģītībā un galīgās sarežģītības teorijas radiniece Gēdeļa nepilnības teorēmai: ir patiesi apgalvojumi, kuriem jūsu fiksētajā noteikumu sistēmā nav īsu pierādījumu — nevis tāpēc, ka tos principā nevar pierādīt, bet tāpēc, ka katra fiksēta sistēma atstāj kādu īsi formulējamu patiesību bez īsa pierādījuma.

Raksts pieņem šo hipotēzi nedaudz stiprākā "bezgalīgi bieži" (*infinitely often*) formā, kā tas kriptogrāfiskos pielietojumos ir ierasts. Pēc Krajíček un Pudlák teorēmas ieguvums ir konkrēts: katrai noteikumu sistēmai eksistē formulu virkne, kuras patiesībā ir neizpildāmas, bet kuras šī sistēma nespēj atspēkot ar īsiem pierādījumiem — un, kas ir ļoti svarīgi, efektīvs algoritms šo virkni var *ģenerēt*. Tieši šī vienmērīgā ģenerējamība pārvērš ideju no eksistences apgalvojuma par faktisku algoritmu, ko Alise var palaist: mānekļi D nāk no ražošanas līnijas, nevis no zila gaisa.

Kriptogrāfiskais gājiens ir šo pierādīšanas spējas trūkumu izmantot kā resursu.

Ko konstrukcija patiesībā dara

Lūk, raksta konstrukcija, atstājot tikai tās skeletu.

Fiksējam noteikumu grāmatu — piemēram, ZFC. Pieņemot pierādījumu sarežģītības hipotēzi, eksistē efektīvi ģenerējama formulu virkne, kuras patiesībā ir neizpildāmas, bet kurām šajā noteikumu sistēmā nav īsa neizpildāmības pierādījuma.

Tagad uzbūvējam vienas ziņas pierādījumu šādā formā:

> vai nu īstais apgalvojums ir izpildāms, vai arī šī īpašā grūtā formula ir izpildāma.

Īpašā grūtā formula nav izpildāma. Tāpēc, ja pamatā esošā pierādīšanas mehānika ir perfekti droša pret nepatiesiem apgalvojumiem, pieņemts ziņojums joprojām nozīmē, ka īstais apgalvojums ir patiess. Tas nodrošina perfektu drošumu.

Bet nulles zināšanām līdzīgajai drošībai iedomāsimies, ka īpašā grūtā formula *būtu* izpildāma. Tad tās liecinieku varētu izmantot, lai simulētu pierādījumus, nezinot īsto liecinieku. Realitātē formula nav izpildāma — taču noteikumu sistēma nespēj to efektīvi pierādīt. Tātad tā nespēj efektīvi pierādīt arī simulatora neiespējamību.

Tieši šeit ir eņģe. Sistēma nepaslēpj noslēpumu, uzbūvējot klasisku simulatoru. Plašai novērojamu drošības testu klasei tā paslēpj noslēpumu aiz formālās sistēmas nespējas apliecināt, ka simulatora nav.

Ko raksts apgalvo

Galvenā teorēma nāk vairākos slāņos. Kodola rezultāts ir šāds.

Pieņemot standarta kriptogrāfisku pieņēmumu — **neinteraktīvu pierādījumu ar liecinieku neatšķiramību** (*non-interactive witness indistinguishable proofs*) eksistenci; tie ir labi izpētīti objekti, kas seko no vairākiem zināmiem pieņēmumu komplektiem — un pierādījumu sarežģītības hipotēzi, ka **neeksistē (bezgalīgi bieži) optimāla pierādījumu sistēma**, raksts katrai izvēlētai noteikumu sistēmai konstruē vienas ziņas pierādītāju un pārbaudītāju NP/SAT problēmām ar **perfektu drošumu** un bez sākotnējas iestatīšanas, kas attiecībā pret šo sistēmu ir efektīvi nulles zināšanu. (NP/SAT ir standarta “grūtākais kopīgais saucējs” mīklām līdzīgām problēmām; MegaSudoku ir viens no tā tērpiem.)

Plašākam apgalvojumam par falsificējamu drošības īpašību saglabāšanu raksts pievieno vēl vienu standarta pieņēmumu — derandomizācijas hipotēzi **P = BPP** (aptuveni: nejaušība algoritmiem nedod būtisku papildu skaitļošanas spēku).

Pārtulkojot teorēmas valodu ikdienas teikumos:

- Pierādījums ir viena ziņa.

- Nav uzticamas sākotnējās iestatīšanas.
- Nepatiesus apgalvojumus nevar pierādīt.
- Pierādītājs nav klasiski nulles zināšanu — tam nav simulatora.
- Taču šajā vidē var iegūt katru falsificējamu, spēlē pārbaudāmu klasisko nulles zināšanu drošības seku.

Vārdam "falsificējama" te ir nozīme. Tas nozīmē, ka drošības kļūmi var pārbaudīt, palaižot pretinieku noteiktā eksperimentā jeb spēlē. Daudzām kriptogrāfijas drošības definīcijām ir tieši šāda forma: vai pretinieks spēj atšķirt divus šifrējumus, invertēt funkciju, atgūt liecinieku vai uzvarēt kādā konkrēti definētā eksperimentā? Teorēma katrai šādai falsificējamai īpašībai dod pierādītāju atsevišķi. Viens pierādītājs, kam vienlaikus piemistu *visas* falsificējamās īpašības, visticamāk, nav iespējams — vecais atkārtotas izmantošanas uzbrukums ("Bobs var parādīt pierādījumu citiem") pats ir falsificējama īpašība, un šeit tā patiešām netiek saglabāta. Raksts izvirza domu, ka viens pierādītājs varētu aptvert visas *dabiskās* falsificējamās īpašības — tās, kas patiešām sastopamas kriptogrāfijas praksē —, taču šī daļa ir nosacīta teorēma, kas balstās gan neformālā jēdzienā "dabiska", gan skaidri izteiktā papildu hipotēzē. Garantija ir vērsta uz novērojamām drošības kļūmēm, nevis uz katru filozofisku vai simulācijā formulējamu slepenības nozīmi.

Vienu konkrētu secinājumu ir vērts nosaukt atsevišķi: konstrukcija dod pirmos neinteraktīvos **liecinieku slēpjošos** (*witness hiding*) pierādījumus ar vienmērīgu pierādītāja algoritmu — "mīklas pierādījums nepalīdz atrast tās atrisinājumu", bez mijiedarbības un bez iestatīšanas. Tas izklausās pieticīgi, bet šāda konstrukcija bija pretojusies mēģinājumiem gadu desmitiem.

Ko tas neapgalvo

Šī sadaļa ir svarīga, lai nepārlasītu rezultātu.

Tas **nenozīmē**, ka vecās neiespējamības teorēmas bija kļūdainas. Konstrukcija tās apiet, mainot definīciju.

Tas **nedod** parastas, klasiskas nulles zināšanas bez mijiedarbības, bez iestatīšanas un ar perfektu drošumu. Raksts skaidri pasaka, ka uzbūvētajam pierādītājam simulatora nav.

Tas **nenozīmē**, ka pierādījumu nevar izmantot atkārtoti. Vienas ziņas pierādījumu joprojām var parādīt kādam citam; raksts nesaglabā noliedzamības (*deniability*) tipa īpašības. (Tāds pats ierobežojums ir arī neinteraktīviem nulles zināšanu pierādījumiem ar uzticamu iestatīšanu.)

Tas **nenozīmē**, ka iegūts praktisks, ieviešanai gatavs protokols. Tas ir sarežģītības teorijas un kriptogrāfijas pamatu darbs. Rezultāts balstās lielos pieņēmumos no pierādījumu sarežģītības un kriptogrāfijas, un konstrukcija runā par to, kas principā ir iespējams.

Tas **nepadara "Gēdeli" par maģisku drošības primitīvu**. Saikne ar Gēdeli iet caur pierādījumu sistēmām, optimālām pierādījumu sistēmām un galīgiem nepilnības analogiem. Noderīgā intuīcija nav "nepilnības teorēma sargā jūsu paroli". Tā ir: ja noteikumu sistēma nespēj efektīvi pierādīt, ka simulators nav iespējams, tad uzbrukumus, kuru pierādīšanai būtu vajadzīgs šāds atspēkojums, var

bloķēt drošības definīciju līmenī.

Kāpēc tas tomēr ir interesanti

Kriptogrāfija bieži pārvērš grūtību drošībā. Faktorizācija ir grūta, tāpēc RSA tipa pieņēmumi kļūst noderīgi. Režģu problēmas ir grūtas, tāpēc režģu kriptogrāfija kļūst noderīga. Šeit grūtība ir dīvaināka: nevis "grūti aprēķināt noslēpumu", bet "grūti pierādīt, ka konkrēts pierādījumu objekts nevar eksistēt".

Tieši tāpēc raksts šķiet neparasts. Tas gandrīz izturas pret aksiomām un formālām noteikumu sistēmām kā pret kriptogrāfiskiem resursiem. Parastā neiespējamība saka, ka starp drošumu un simulāciju ir spriedze. Ilango gājiens šo spriedzi paslēpj aiz pierādījumu teorijas aizkara: simulatora nav, bet formālā sistēma nespēj efektīvi atklāt tā neesamību.

Lasītājam pārsteidzošākais nav tas, ka šī pieeja aizstās mūsdienu nulles zināšanu sistēmas. Vistic-amāk, vismaz tieši, tā nenotiks. Pārsteidzoši ir tas, ka matemātiskās loģikas ierobežojumu var iz-mantot konstruktīvi — ne tikai kā sienu, bet arī kā aizsegu.

Cik pārliecinošs ir rezultāts?

Šis ir teorēmu raksts, tāpēc "pierādījumi" šeit nozīmē kaut ko citu nekā bioloģijā vai astronomijā. Jautājums nav, vai eksperiments atkārtojās. Jautājums ir, vai definīcijas, pieņēmumi un pierādījumu ķēde atbalsta apgalvojumu.

Pierādījums ir formāls, un raksts skaidri uzskaita pieņēmumus. Tie nav nejauši. Neinteraktīvi pierādījumi ar liecinieku neatšķiramību ir standarta kriptogrāfiski objekti un seko no vairākiem labi zināmiem pieņēmumu komplektiem. Hipotēze par optimālas pierādījumu sistēmas neesamību ir centrāla pierādījumu sarežģītības hipotēze. **P = BPP** ir standarta derandomizācijas pieņēmums, kas vajadzīgs tikai plašākajai teorēmai par falsificējamām īpašībām.

Raksts arī argumentē, ka šie pieņēmumi nav patvaļīgi uzbūvētas sastatnes, bet būtībā ir pareizā cena. Tas pierāda pretējo virzienu: ja šāda veida konstrukcijas vispār eksistē, tad jāeksistē neinteraktīviem pierādījumiem ar liecinieku neatšķiramību un — pieņemot standarta vienvirziena funkciju eksistenci — optimāla pierādījumu sistēma nevar eksistēt. Turklāt pieņēmumi ir "abpusēji interesanti": jebkura no tiem atspēkošana pati par sevi būtu nozīmīgs atklājums pierādījumu sarežģītībā, kriptogrāfijā vai sarežģītības teorijā.

Taču, tā kā rezultāts ir nosacīts, nosacīta ir arī pārliecība par tā interpretāciju. Ja pieņēmumi izrādās nepareizi, teorēmas nozīme mainās. Un pat tad, ja tie ir pareizi, garantija nav pilnas klasiskās nulles zināšanas; tā ir rakstā definētā vājākā, pierādījumu teorijas versija.

Tāpēc pareizā pārliecības pakāpe ir augsta, ka raksts izveido konsekventu nosacītas iespējamības rezultātu; mērena, ka pieņēmumi apraksta kriptogrāfisko pasauli, kurā patiešām dzīvojam; un zema

attiecībā uz tūlītējām praktiskām sekām.

Kāpēc tas ir svarīgi

Raksts atver ceļu, kam klasiskajā kartē vajadzēja būt slēgtam.

Klasiskā teorija saka: pilnas nulles zināšanas bez iestatīšanas nevar saspiest vienā ziņojumā, un tās nevar būt perfekti drošas pret nepatiesiem apgalvojumiem. Ilango raksts saka: ja prasām tikai tās nulles zināšanu sekas, kuras var pārbaudīt drošības spēlēs, un ja ļaujam drošības definīcijai būt atkarīgai no tā, ko izvēlētā formālā sistēma spēj vai nespēj efektīvi atspēkot, tad lielu daļu noderīgās uzvedības var atgūt — ar vienu ziņojumu, bez iestatīšanas un ar perfektu drošumu.

Tas nav mazs definīcijas labojums. Tas ir cits veids, kā domāt par kriptogrāfiskām garantijām. Neprasīt tikai, kas eksistē, bet arī — ko jūsu noteikumu sistēma spēj izslēgt. Nepierādāmību neuztvert tikai kā filozofisku neērtību, bet izmantot kā struktūru.

Praktiskā pasaule rīt, iespējams, nemainīsies. Taču konceptuālā karte mainās. Tagad ir formāla nozīme, kurā apgalvojums "neviens nespēj efektīvi pierādīt, ka noslēpums ir noplūdis" var būt pietiekami spēcīgs, lai atgūtu daudzas spēlēs pārbaudāmas aizsardzības īpašības, ko mēs vēlējāmies no stiprākā apgalvojuma "noslēpums nav noplūdis".

Tieši tāpēc Gēdelis ir raksta nosaukumā.

Īss kopsavilkums

Nulles zināšanu pierādījumi ļauj pierādītājam pārliecināt pārbaudītāju, ka apgalvojums ir patiess, neatklājot liecinieku. Klasiskie neiespējamības rezultāti rāda, ka pilnas nulles zināšanas nevar saspiest vienā ziņojumā bez sākotnējas iestatīšanas un nevar vienlaikus prasīt perfektu drošumu. Rahul Ilango raksts šīs neiespējamības neatspēko. Tas definē vājāku jēdzienu — efektīvas nulles zināšanas: tā vietā, lai prasītu, ka simulators patiešām eksistē, tiek prasīts, lai izvēlētā pierādījumu sistēma — formāla noteikumu grāmata, piemēram, ZFC — nespētu efektīvi pierādīt, ka simulatora nav. Pie būtiskiem kriptogrāfiskiem pieņēmumiem (neinteraktīvi pierādījumi ar liecinieku neatšķiramību) un pierādījumu sarežģītības hipotēzes (optimāla pierādījumu sistēma neeksistē) raksts konstruē NP/SAT vienas ziņas pierādītājus bez iestatīšanas un ar perfektu drošumu, kas pa vienai iegūst falsificējamās, spēlēs pārbaudāmās klasisko nulles zināšanu sekas. Viens pierādītājs, kas aptvertu visas "dabiskās" šādas īpašības, ir tālāks, daļēji hipotētisks paplašinājums; pilnīgi visu falsificējamo īpašību vienlaicīga saglabāšana, visticamāk, nav iespējama, jo pierādījumus joprojām var izmantot atkārtoti. Rezultāts ir teorētisks un nosacīts, nevis gatavs kriptogrāfisks primitīvs, taču tas parāda jaunu veidu, kā pierādījumu teorijas nepierādāmību izmantot par kriptogrāfisku resursu.

Bez pārspīlējumiem

Ko raksts parāda: Pie skaidri norādītiem pieņēmumiem NP/SAT var uzbūvēt vienas ziņas pierādītājus bez sākotnējas iestatīšanas un ar perfektu drošumu, kuri ir efektīvi nulles zināšanu attiecībā pret jebkuru izvēlētu pierādījumu sistēmu un pa vienai sasniedz katru falsificējamu, spēlē pārbaudāmu klasisko nulles zināšanu drošības seku.

Kas ir ticams, bet nav beznosacījumu pierādīts: Ka nepieciešamie pierādījumu sarežģītības un kriptogrāfiskie pieņēmumi ir patiesi. Tie ir nopietni un plaši pētīti pieņēmumi — un raksts rāda, ka tie būtībā ir ne tikai pietiekami, bet arī nepieciešami —, tomēr tie joprojām ir pieņēmumi.

Ko tas neparāda: Klasiskas nulles zināšanas bez mijiedarbības, bez iestatīšanas un ar perfektu drošumu; praktisku ieviešanai gatavu sistēmu; pierādījumu noliedzamību vai neatkārtotu izmantošanu; vai to, ka Gēdeļa nepilnības teorēma pati par sevi nodrošina kriptogrāfisku drošību.

Galvenie ierobežojumi: Garantija ir nulles zināšanu relaksācija; plašākā versija balstās vairākos pieņēmumos; apgalvojumi par vienu universālu pierādītāju joprojām daļēji ir hipotētiski; un rezultāts galvenokārt ir fundamentāls, nevis praktisks.

Cik lielai pārliecībai vajadzētu būt vispārīgam lasītājam? Augstai, ka, pieņemot definīcijas un pieņēmumus, tas ir nozīmīgs nosacīts teorētisks rezultāts. Mērenai, ka pieņēmumi patiešām raksturo mūsu kriptogrāfisko pasauli. Zemai attiecībā uz tūlītēju praktisku ieviešanu. Drošākais secinājums: raksts nesalauž klasiskās nulles zināšanu neiespējamības teorēmas; tas atrod jaunu, pierādījumu teorijā balstītu ceļu apkārt tām daļām, kas ir būtiskas daudzām drošības spēlēm.

Avoti

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`