The Clean Paper

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

Kriptografinis įrodymas gali paslėpti paslaptį todėl, kad sunku įrodyti trūkstamo simuliatoriaus nebuvimą

Nulinio žinojimo įrodymai leidžia įrodyti teiginį neatskleidžiant liudytojo. Klasikinė teorija sako, kad pilna šios savybės versija neįmanoma su viena žinute, be patikimos pradinės sąrankos ir kartu su tobulu patikimumu. Rahulio Ilango straipsnis šios neįmanomybės nepanaikina. Jis pakeičia tikslą: vietoj reikalavimo, kad simuliatorius iš tikrųjų egzistuotų, reikalaujama, kad pasirinkta formali įrodymų sistema negalėtų efektyviai įrodyti, jog simuliatoriaus nėra. Esant svarbioms įrodymų sudėtingumo ir kriptografinėms prielaidoms, to pakanka po vieną atkurti falsifikuojamas, žaidimais grindžiamas nulinio žinojimo saugumo pasekmes. Tai nėra paruoštas interneto protokolas. Tai įrodymų teorijos būdas matematinį neįrodomumą paversti kriptografine priedanga.

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

Gudrybė nėra įrodyti, kad paslaptis tikrai paslėpta

Pradėkime nuo paprasčiausios nulinio žinojimo versijos.

Alice nori įtikinti Bobą, kad Sudoku uždavinys turi sprendinį. Jei ji nusiųs patį sprendinį, Bobas bus įtikintas, tačiau galvosūkis bus sugadintas. Alice nori keistesnio dalyko: įrodymo, kad sprendinys egzistuoja, neatskleidžiant paties sprendinio.

Būtent tai žada nulinio žinojimo įrodymas. Įrodantis dalyvis — Alice — įtikina tikrintoją Bobą, kad teiginys teisingas, neatskleisdamas nieko daugiau nei paties teiginio teisingumą.

Problema ta, kad ši garantija turi kainą. Paprastas matematinis įrodymas turi dvi patogias savybes. Jis yra **viena žinutė**: užrašai, perduodi ir nueini. Ir jis yra **tobulai patikimas** (*perfectly sound*): klaidingas teiginys apskritai neturi galiojančio įrodymo. Klasikiniai neįmanomumo rezultatai sako, kad nulinis žinojimas turi atsisakyti abiejų savybių — ir ne tik jų derinio; kiekviena atskirai jau kelia problemą.

Pirma, nulinio žinojimo įrodymui reikia dialogo. Jei Alice siunčia tik vieną žinutę ir iš anksto nėra jokios patikimos bendros sąrankos, nulinio žinojimo garantija sugriūva — nepriklausomai nuo to, kiek patikimumo sutiktumėte paaukoti mainais.

Antra, nulinio žinojimo įrodymui reikia bent menkos klaidos tikimybės. Pasirodo, reikalaujant tobulo patikimumo tyliai prarandama ir sąveika: tikrintojas, kurio neįmanoma apgauti nepriklausomai nuo jo atsitiktinių pasirinkimų, gali tuos pasirinkimus tiesiog nustatyti iš anksto. Kai tikrintojas tampa visiškai nuspėjamas, Alice gali į viską atsakyti viena žinute — o tai būtent tas atvejis, kuris jau buvo neįmanomas.

Rahulio Ilango straipsnyje siūlomas kelias aplink šią dvigubą sieną. Ne apsimetant, kad sienos nėra, ir ne sukuriant klasikinį nulinį žinojimą ten, kur jis neįmanomas. Žingsnis subtilesnis: susilpninti, ką reiškia „nieko neatskleidžia“, bet susilpninti taip, kad išliktų tos saugumo savybės, kurias kriptografai iš tikrųjų gali patikrinti.

Ši sąvoka vadinama **efektyviu nuliniu žinojimu** (*effectively zero-knowledge*).

A proof without leaking the witness

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

blocked door 1

interaction

blocked door 2

trusted setup

blocked door 3

imperfect soundness

the route

the rulebook cannot efficiently refute the simulator

Boundary: not classical zero-knowledge; effectively zero-knowledge under a chosen proof system.

Klasikinį nulinį žinojimą blokuoja trys durys — sąveika, patikima pradinė sąranka ir būtinybė toleruoti netobulą patikimumą. Ilango konstrukcija pasuka kitu keliu: pasirinkta taisyklių sistema negali efektyviai paneigti simuliatoriaus galimybės.

Original diagram — The Clean Paper CC BY 4.0

Classical privacy vs effective privacy

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

classical zero-knowledge

positive claim

A simulator exists and can fake the verifier's view without the witness.

effectively zero-knowledge

weaker claim

Your chosen proof system cannot efficiently prove that no simulator exists.

Boundary: the construction preserves testable consequences, not the full simulator guarantee.

Klasikinis nulinis žinojimas klausia, ar simuliatorius egzistuoja; „efektyvus nulinis žinojimas“ klausia tik, ar

pasirinkta taisyklių sistema gali efektyviai įrodyti, kad jis neegzistuoja. Šis silpnesnis klausimas ir leidžia konstrukcijai išlaikyti vieną žinutę, jokios sąrankos ir tobulą patikimumą.

Original diagram — The Clean Paper CC BY 4.0

Senasis testas: simuliatorius egzistuoja

Klasikinėje nulinio žinojimo formalizacijoje naudojamas išgalvotas pagalbininkas, vadinamas **simuliatoriumi**.

Idėja tokia: įsivaizduokime Jane, kuri **nežino** Alice paslapties. Jei Jane pati gali sugeneruoti įrodymus, atrodančius taip pat, kaip tie, kuriuos Bobas būtų gavęs iš Alice, vadinasi, Alice įrodymai Bobo neišmokė nieko naujo. Jane jau galėjo atkurti tą pačią patirtį neturėdama Alice paslapties.

Todėl klasikinis nulinis žinojimas reikalauja realaus simuliatoriaus. Turi egzistuoti efektyvus algoritmas, galintis generuoti į įrodymus panašius objektus nežinodamas paslapties — žargonu vadinamos *liudytoju* (*witness*). Sudoku atveju liudytojas tiesiog yra užpildytas sprendinio tinklelis.

Šis apibrėžimas galingas, tačiau būtent čia ir smogia senasis neįmanomumo rezultatas. Intuicija tokia: tikrai nesąveikus įrodymas yra tiesiog eilutė. Gavęs ją Bobas gali parodyti ją kam nors kitam; jis įgyja gebėjimą įrodyti teiginį kitiems, o tai jau skamba kaip daugiau nei „nieko“. Klasikinės teoremos šią intuiciją paverčia aukščiau aprašytais neįmanomumo rezultatais.

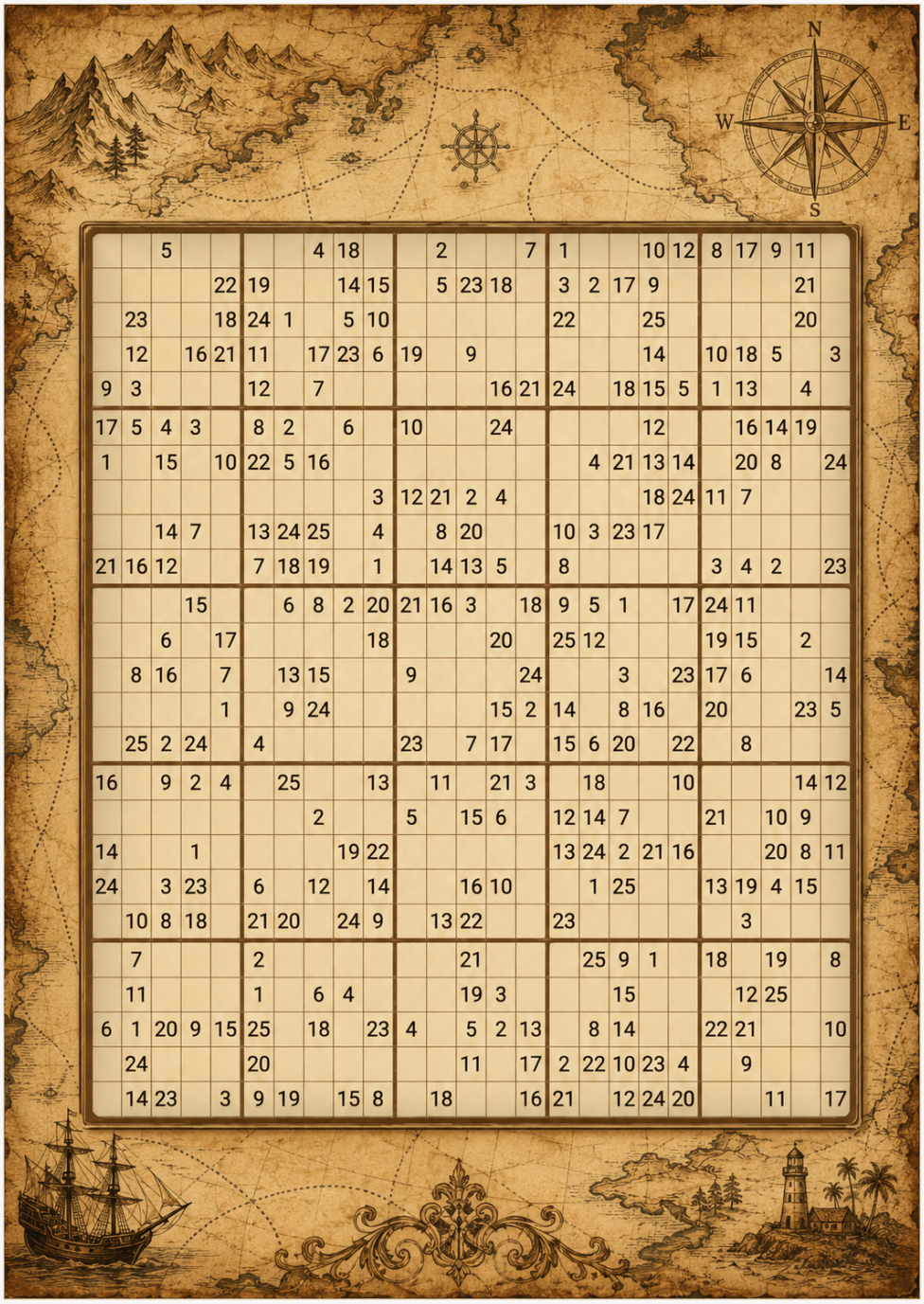
> **Trys savybės, kurių šis straipsnis neatsisako**
>
> Straipsnio pavadinime slypi trys apribojimai:
> **Be sąveikos:** Alice siunčia vieną įrodymo eilutę. Nėra dialogo pirmyn ir atgal.
> **Be sąrankos:** Alice ir Bobas nesiremia patikima bendra atskaitos eilute ar kita iš anksto parengta vieša atsitiktine informacija. Daugelis sistemų, vadinamų „nesąveikiu nuliniu žinojimu“, vis tiek turi sąranką; šiame straipsnyje sąrankos nėra visai.
> **Tobulas patikimumas:** klaidingas teiginys neturi jokio galiojančio įrodymo. Ne „beveik niekada nepriimamas“, o tiesiog nėra galiojančio įrodymo.
> Šios trys savybės yra būtent tos, kurias turi įprasta rašytinė matematika — ir, kaip paaiškinta aukščiau, klasikinis nulinis žinojimas negali jų visų išlaikyti.

Skirtumas per MegaSudoku pavyzdį

Štai sąmoningai supaprastintas būdas pajusti skirtumą.

Rimtoje analogijos dalyje nenaudokime įprasto 9×9 Sudoku. Jis per mažas ir pernelyg baigtinis: kompiuteris gali tiesiog jį išspręsti arba įrodyti, kad sprendinio nėra. Vietoj to įsivaizduokime **MegaSudoku(n)** uždavinių šeimą. Išplėskime taisyklę: pasirenkame bloko dydį n, nustatome $N = n^2$ ir sudarome $N \times N$ tinklelį, padalytą į $n \times n$ blokus ir naudojantį N simbolių. Įprastas Sudoku yra tik

mažytis $n = 3$, $N = 9$ atvejis: 9×9 tinklelis, 3×3 blokai ir devyni simboliai. Įrodymų sudėtingumo istorija prasideda tik tada, kai n leidžiama augti ir kai tinklelyje gali atsirasti papildomų „gadgetų“, verčiančių jį elgtis kaip SAT formulę, persirengusią Sudoku. SAT formulė tėra taip/ne apribojimų rinkinys: ar galima kintamiesiems priskirti true/false reikšmes taip, kad būtų patenkintas kiekvienas apribojimas?



25×25 Sudoku: jo taisyklių laikymąsi galima tikrinti neatskleidžiant viso užpildyto tinklelio — vizualus įrodymo, tikrinančio paslėptą sprendinį, t. y. liudytoją, pakaitalas.

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

Sudoku ir SAT: tas pats galvosūkis dviem pavidalais

Teiginys, kad Sudoku gali „elgtis kaip SAT formulė“, nėra metafora. Transformacija veikia abiem kryptimis, o lengvesnę kryptį galima išrašyti visiškai.

Iš Sudoku į SAT. SAT kalba tik true/false reikšmėmis, todėl kiekvienai (eilutė, stulpelis, reikšmė) trijulei skirkime po loginį kintamąjį: *x(r,c,v)* reiškia „langelyje, esančiame eilutėje *r* ir stulpelyje *c*, yra reikšmė *v*“. 4×4 Sudoku (2×2 blokai, reikšmės 1–4) reikia 4·4·4 = 64 kintamųjų; klasikiniam 9×9 — 729. Tada kiekviena Sudoku taisyklė tampa klauzių rinkiniu. (Klauzė yra kintamųjų arba jų neiginių OR; visa formulė yra visų klauzių AND.)

Kiekviename langelyje yra bent viena reikšmė — po vieną klauzę kiekvienam langeliui:

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

Kiekviename langelyje yra daugiausia viena reikšmė — kiekvienai reikšmių porai po „ne abi“ klauzę:

$$\neg x(1,1,1) \lor \neg x(1,1,2) \quad \neg x(1,1,1) \lor \neg x(1,1,3) \quad \text{... ir taip visoms šešioms poroms.}$$

Kiekvienoje eilutėje yra kiekviena reikšmė — pavyzdžiui, 1 eilutėje reikšmė 3 bent kartą:

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

ir daugiausia kartą: $\neg x(1,1,3) \lor \neg x(1,2,3)$, ir taip kiekvienai tos eilutės langelių porai.

Stulpeliai ir blokai — tokios pat klauzių grupės, tik keičiasi nagrinėjamų langelių rinkinys. Viršutiniam kairiajam blokui ir reikšmei 2:

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

ir papildomos porinės „ne abi“ klauzės.

Atspausdintos užuominos — paprasčiausia dalis: kiekviena užuomina yra vieno kintamojo klauzė. Viršutiniame kairiajame kampe atspausdintas 3 tampa klauze

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

Visų šių sąlygų AND yra patenkinama tada ir tik tada, kai Sudoku turi sprendinį — o patenkinantis priskyrimas *ir yra* sprendinys: pažiūrėkite, kurie x(r,c,v) yra true, ir užpildykite tinklelį. 9×9 atveju gauname 729 kintamuosius ir kelis tūkstančius klauzių, su kuriais šiuolaikinis SAT sprendiklis susitvarko per milisekundes. Atkreipkite dėmesį į užuominos klauzę x(1,1,3): ji sako „šis langelis lygus tiksliai 3“, o ne „visi šie langeliai skirtingi“. Tas pats asimetriškumas vėliau privers panaudoti papildomą triuką užuominų langeliams.

Iš SAT į Sudoku. Straipsniui reikia priešingos, sunkesnės krypties: paėmus *bet kokią* SAT formulę sukurti MegaSudoku, kuris turi sprendinį tada ir tik tada, kai formulė patenkinama. Gimtosios Sudoku taisyklės gali pasakyti tik „šie langeliai visi skirtingi“, todėl savavališkus loginius apribojimus reikia *sukonstruoti* — būtent tam skirti gadgetai. Gadgetas yra nedidelė iš anksto suprojektuota langelių grupė, po vieną formulės klauzei, kurioje tam tikri langeliai atlieka kintamųjų vaidmenį (juose esantis simbolis koduoja true arba false), o vidiniai apribojimai sukonstruoti taip, kad teisėti užpildymai atitiktų tik tuos priskyrimus, kurie

patenkina klauzę. Tai standartinė NP-pilnumo įrodymų technika; generalizuotam Sudoku ją 2003 m. realizavo Yato ir Seta.
Abi kryptys kartu sako, kad N×N Sudoku ir SAT iš esmės yra ta pati problema skirtingais rūbais. Todėl ir šiame straipsnyje, ir pačiame darbe galima pasakoti apie visą NP klasę naudojant tinklelius ir simbolius.

Liudytoją vis dar lengva įsivaizduoti. Alice žino visiškai ir teisingai užpildytą MegaSudoku. Bobas nori įsitikinti, kad toks užpildymas egzistuoja, tačiau Alice nenori jo atskleisti. Jei ji nusiųs visą užpildymą, Bobas bus įtikintas, bet paslaptis dings.

Klasikinėje nulinio žinojimo versijoje Alice ir Bobas sąveikauja. Vienas senas intuityvus modelis naudoja uždengtas korteles. Alice paslepia išspręstą tinklelį, kiekviename etape slapta pervadina simbolius ir leidžia Bobui patikrinti vieną atsitiktinai pasirinktą vietinį apribojimą: eilutę, stulpelį, bloką ar gadgetą. Jei atversti langeliai rodo tarpusavyje skirtingus simbolius, Bobo pasitikėjimas padidėja. Tada viskas vėl uždengiama ir simboliai iš naujo atsitiktinai pervadinami. (Yra viena subtilybė: atspausdintoms užuominoms reikia papildomo triuko, nes pervadinus simbolius paslepiamos ir jos. Toliau esančioje pastaboje paaiškinta, kaip tai sprendžia klasikiniai protokolai; mūsų supaprastintai intuicijai šio vaizdo pakanka.)

Kaip klasikiniai protokolai iš tikrųjų tvarko užuominų langelius

Simbolių pervadinimo triukas turi akląją vietą. Eilutės, stulpelio ir bloko taisyklės visos sako „šie langeliai tarpusavyje skirtingi“, o savybė *visi skirtingi* išlieka bet kaip pervadinus simbolius. Tačiau užuomina sako „šiame langelyje yra lygiai 5“. Po pervadinimo Bobas mato tik $\sigma(5)$ — kažkokį užmaskuotą simbolį — bet nežino pervadinimo σ, todėl negali nieko patikrinti. Jei problema liktų neišspręsta, Alice galėtų įrodyti, kad egzistuoja *kažkoks* galiojantis tinklelis, visiškai ignoruodama atspausdintas šio konkretaus galvosūkio užuominas. Klasikinėje literatūroje naudojami du standartiniai pataisymai.
Paletė. Prie paslėpto tinklelio pridedama viena papildoma N langelių eilutė — paletė, kurią Alice užpildo simboliais 1…N fiksuota vieša tvarka, o tada pervadina kartu su visu kitu, taigi joje atsiranda $\sigma(1)$…$\sigma(N)$. Bobo atsitiktinis iššūkis dabar turi dar vieną variantą. Be eilutės, stulpelio, bloko ar gadgeto jis gali pasirinkti **paletę ir vieną užuominos langelį**. Alice atverčia abu; paletė parodo to raundo pervadinimą, o Bobas patikrina, ar užuominos langelyje yra būtent pervadinta atspausdintos reikšmės versija. Nulinis žinojimas išlieka, nes Bobas sužino tik σ — kiekviename raunde jis parenkamas naujai ir pats savaime bevertis — bei langelio reikšmę, kurią jis jau žinojo iš užduoties. Apie slaptus langelius niekas neatskleidžiama, o simuliatorius gali suklastoti tokį vaizdą tiesiog parinkdamas atsitiktinį σ. Patikimumas išlieka todėl, kad sukčiaujanti Alice kiekviename raunde su fiksuota tikimybe bus pagauta; raundai kartojami, kol abejonė tampa nereikšminga.
Užuominų „sukompiliavimas“. Struktūriškesnė versija pašalina specialų iššūkį vietoj to, kad jį pridėtų. Užuot *tikrinus* užuominos reikšmę, ji priverstinai nustatoma nelygybės apribojimais: užuominos langelis susiejamas su kiekvienu paletės langeliu, išskyrus tą, kuriame yra jo paties reikšmė — „skiriasi nuo $\sigma(1)$, skiriasi nuo $\sigma(2)$, …, skiriasi nuo visko, išskyrus $\sigma(5)$“. Vienintelis simbolis, kurį tas langelis gali teisėtai turėti, yra užuominos reikšmė. Dabar visi apribojimai

> vėl yra tipo „šios dvi reikšmės skirtingos“ — pervadinimui invariantūs ir tikrinami taip pat kaip eilutė. Tas pats metodas naudojamas iš anksto nuspalvintoms viršūnėms klasikiniame grafų spalvinimo protokole; tai ir yra aukščiau minėtų *gadgetų* dvasia: MegaSudoku-kaip-SAT vaizde užuominos, kaip ir kiti apribojimai, sukompiliuojamos į nelygybės gadgetus.
> **Fizinis protokolas.** Realus kortelių protokolas Sudoku uždaviniams (Gradwohl, Naor, Pinkas ir Rothblum, 2007) išvis nenaudoja pervadinimo ir užuominas sutvarko dar prieš pradedant slėpti. Kiekvienam langeliui Alice padeda tris vienodas korteles su langelio reikšme — slaptiems langeliams užverstas, o **užuominų langeliams atverstas**, todėl Bobas savo akimis mato, kad užuominų laikomasi prieš korteles užverčiant. Tada po vieną kortelę iš kiekvieno langelio dedama į jo eilutės paketą, po vieną — į stulpelio, po vieną — į bloko paketą; kiekvienas paketas sumaišomas ir atverčiamas, o Bobas tikrina, ar jame yra visi N simbolių. Maišymas sunaikina informaciją apie pozicijas — tai ir suteikia nulinį žinojimą — o užuominos jau buvo patikrintos dalijimo metu.
> Abiem atvejais pamoka ta pati ir nuolat kartojasi šiame straipsnyje: nulinio žinojimo protokolas yra kruopšti apskaita, *kurie* faktai išlieka po slėpimo. Pervadinimas išsaugo „visi skirtingi“ ir ištrina „lygu 5“, todėl „lygu 5“ turi būti grąžinta kitu būdu.

Tai nėra šiame straipsnyje naudojamas protokolas. Tai tik mentalinis klasikinio nulinio žinojimo modelis:

- Alice ir Bobas keičiasi žinutėmis pirmyn ir atgal.
- Bobas pasirenka atsitiktinius patikrinimus.
- Alice atskleidžia tik vietinį nuoseklumą, o ne visą sprendinį.
- Privatumo įrodymas remiasi tuo, kad Bobo matomą vaizdą būtų buvę galima sugeneruoti nežinant slapto Alice sprendinio.

Taigi klasikinis nulinis žinojimas remiasi teigiamu faktu:

> Simuliatorius iš tikrųjų egzistuoja.

Dabar pašalinkime patogias dalis. Alice siunčia vieną įrodymo eilutę ir pasitraukia. Nėra patikimos pradinės sąrankos, nėra iš anksto parengtos bendros atsitiktinės eilutės, o Bobas niekada negali priimti klaidingo galvosūkio. Tai aplinka, kurioje klasikinis nulinis žinojimas neišgyvena.

Prieš pereinant prie gudrybės reikia dar vieno veikėjo. Pasirinkime **taisyklių knygą**: formalią įrodymų sistemą logikų prasme — fiksuotą aksiomų rinkinį ir mechanines taisykles, pagal kurias tikrinami užrašyti matematiniai įrodymai. Kanoninis pavyzdys yra ZFC, standartinės matematikos aksiomos. Nuo čia visi teiginiai formuluojami pasirinktos taisyklių knygos atžvilgiu, o pasirinkimas lankstus: konstrukcija veikia su bet kuria iš anksto fiksuota sistema, įskaitant ZFC.

(Paties straipsnio terminologijos pastaba: „įrodymų sistema“ čia visada reiškia šią taisyklių knygą — formalų mechanizmą, tikrinantį matematinius įrodymus — o ne Alice siunčiamas žinutes. Alice ir Bobo mechanizmai vadinami „įrodymo kūrėju“ (*prover*) ir „tikrintoju“ (*verifier*).)

Gödelio stiliaus versija išlaiko MegaSudoku istoriją, bet pakeičia patį įrodymą.

Pasirinkime antrą tokio pat rodomo dydžio apribojimų sistemą ir pavadinkime ją **D**. Istorijoje S ir D yra du tokio pat formato MegaSudoku(n) uždaviniai. Užkulisiuose D galėjo prasidėti kaip kito dydžio sudėtinga loginė formulė; jei reikia, ją galima papildyti nekenksmingais fiktyviais apribojimais, kad tilptų į tokį pat tinklelį. D sukurta iš loginės formulės, kuri iš tikrųjų yra **nepatenkinama** (*unsatisfiable*): nėra jokio reikšmių priskyrimo, tenkinančio visus jos apribojimus, kaip sugedęs galvosūkis neturi teisėto užpildymo. Žaislinis pavyzdys būtų formulė, vienu metu reikalaujanti „X yra true“ ir „X yra false“. Taigi D sprendinio nėra.

Tačiau D negali būti „sugedęs galvosūkis“, kurio gedimą *lengva įrodyti*. Žaislinis pavyzdys čia netinka: bet kuri taisyklių knyga „X ir ne-X“ paneigtų viena eilute. D turi būti klaidinga taip, kad pasirinkta taisyklių sistema negalėtų to patvirtinti trumpu įrodymu. Jei sistema galėtų greitai paneigti D, žemiau aprašoma konstrukcija sugriūtų: būtų formaliai atmestas alternatyvus kelias, kuris galėjo generuoti įrodymus nežinant Alice paslapties, o kartu pradingtų ir privatumo garantija. Todėl D pasirenkama iš šeimos, kurios fiksuota įrodymų sistema negali efektyviai paneigti: toje sistemoje nėra trumpo įrodymo, kad D neturi sprendinio.

Tada Alice vienos žinutės įrodymas yra apie „arba/arba“ teiginį:

> arba tikrasis MegaSudoku S turi sprendinį, arba masalas D turi sprendinį.

Tai loginė jungtis. D **nėra** kažkaip magiškai sugeneruojama taip, kad S taptų teisinga. Įrodymas nesako „D neturi sprendinio, vadinasi, S turi sprendinį“. Jis įrodo disjunkciją **S arba D**. Tobulas patikimumas reiškia, kad klaidinga disjunkcija negali turėti galiojančio įrodymo. Kadangi realybėje D yra klaidinga — sprendinio neturi — vienintelis būdas disjunkcijai būti teisingai yra S teisingumas. Taigi jei įrodymas priimamas, S privalo turėti sprendinį. Masalas negali klaidingo S paversti teisingu.

Tačiau nulinio žinojimo tipo daliai paklauskime, kas būtų, *jei D turėtų sprendinį*. Tas masalo sprendinys būtų alternatyvus liudytojas. Jis leistų generuoti įrodymus nežinant tikrojo Alice MegaSudoku sprendinio — kitaip tariant, veiktų kaip simuliatorius. Realybėje D sprendinio nėra, tad šis simuliavimo kelias uždarytas. Esmė ta, kad taisyklių knyga negali efektyviai įrodyti, jog jis uždarytas.

Taigi D turi du darbus. **Patikimumui** D yra klaidinga, todėl galiojantis „S arba D“ įrodymas priverčia S būti teisingą. **Efektyviam nuliniam žinojimui** D sunku paneigti, todėl taisyklių sistema negali greitai atmesti masalo kelio, kuris, jei būtų atviras, leistų simuliuoti.

Todėl saugumo klausimas jau nebe toks:

> Ar galime įrodyti, kad simuliatorius iš tikrųjų egzistuoja?

Jis tampa:

> Ar jūsų taisyklių knyga gali efektyviai įrodyti, kad simuliatorius neįmanomas?

Jei atsakymas ne, seka stebėtinai stipri pasekmė: kiekviena saugumo garantija, kuri (a) gali būti stebima paleidus testą ir (b) pasirinktoje taisyklių sistemoje įrodomai seka iš simuliatoriaus egzistavimo,

iš tikrųjų galioja. Sėkminga ataka prieš bet kurią iš jų pati reikštų tą trūkstamą trumpą paneigimą, o tokio trumpo paneigimo nėra. Būtent tai ir reiškia „efektyvus“ efektyviame nuliniame žinojime.

Taigi supaprastintas kontrastas toks:

Klasikinis nulinis žinojimas: įrodymai saugūs todėl, kad simuliatorius egzistuoja.

Gödelio stiliaus efektyvus nulinis žinojimas: stebimų saugumo testų atžvilgiu įrodymai laikomi saugiais todėl, kad taisyklių sistema negali efektyviai įrodyti simuliatoriaus neįmanomumo.

Antrasis teiginys silpnesnis. Tačiau būtent todėl straipsnis gali išlaikyti tris savybes, kurios sugriovė klasikinę versiją: vieną žinutę, jokios sąrankos ir tobulą patikimumą.

Naujasis testas: negalite įrodyti, kad simuliatoriaus nėra

Ilango siūlomas susilpninimas pakeičia klausimą.

Klasikinis nulinis žinojimas klausia:

> Ar simuliatorius egzistuoja?

Efektyvus nulinis žinojimas klausia silpniau:

> Ar pasirinkta taisyklių sistema gali efektyviai įrodyti, kad simuliatoriaus nėra?

Tai gali skambėti kaip techninis išsisukinėjimas, tačiau tai pagrindinė idėja. Konstrukcija yra keistoje būsenoje: simuliatorius iš tikrųjų neegzistuoja — straipsnis tai aiškiai sako — bet pasirinkta formali sistema negali efektyviai įrodyti jo nebuvimo. Jei kiekviena jums svarbi bloga pasekmė reikalautų tokio paneigimo, sistema vis tiek elgiasi kaip nulinio žinojimo sistema tų pasekmių atžvilgiu.

Čia į istoriją ateina Gödelis. Ne kaip puošmena ir ne kaip teiginys „Gödelis daro kriptografiją saugią“. Ryšys yra įrodymų teorinis. Taisyklių sistema vadinama **optimalia**, jei ji tiksliąja prasme yra geriausia įmanoma: kai bet kuri kita taisyklių sistema gali atitinkamo tipo formulę paneigti trumpu įrodymu, optimali sistema irgi gali tai padaryti, o jos įrodymas bus daugiausia polinomiškai ilgesnis. Krajíčekas ir Pudlákas 1989 m. iškėlė hipotezę, kad **optimalių įrodymų sistemų nėra**: kokią sistemą befiksuotumėte, visada atsiras kita sistema, kuri kai kurią teisingų teiginių šeimą įrodo gerokai trumpiau. Tai viena centrinių atvirų įrodymų sudėtingumo hipotezių ir baigtinė, sudėtingumo teorijos prasme suformuluota Gödelio nepilnumo teoremos giminaitė: kai kurie teisingi teiginiai fiksuotoje taisyklių sistemoje neturi trumpo įrodymo — ne todėl, kad jų apskritai negalima įrodyti, o todėl, kad kiekviena fiksuota sistema palieka dalį trumpai suformuluojamų tiesų be trumpų įrodymų.

Straipsnyje ši hipotezė priimama šiek tiek stipresne, kriptografijoje įprasta „be galo dažnai“ (*infinitely often*) forma. Krajíčeko ir Pudláko teorema tada duoda konkretų rezultatą: kiekvienai taisyklių sistemai egzistuoja sekos formulių, kurios iš tikrųjų nepatenkinamos, tačiau ta sistema negali jų paneigti trumpais įrodymais — ir, svarbiausia, efektyvus algoritmas gali tas formules *generuoti*.

Pastaroji savybė, vienodumas (*uniformity*), paverčia idėją iš gryno egzistavimo teiginio algoritmu, kurį Alice gali realiai vykdyti: masalai D gaunami sistemingai, o ne ištraukiami iš oro.

Kriptografinis žingsnis — panaudoti šį įrodymo galios trūkumą kaip saugumo išteklių.

Ką daro konstrukcija

Štai straipsnio konstrukcijos forma be techninių detalių.

Fiksuojame taisyklių sistemą — tarkime, ZFC. Pagal įrodymų sudėtingumo prielaidą egzistuoja efektyviai generuojama formulių seka, kurios realybėje nepatenkinamos, bet pasirinktoje sistemoje nėra trumpų jų nepatenkinamumo įrodymų.

Tada sukuriamas vienos žinutės įrodymas tokios formos:

> arba tikrasis teiginys yra patenkinamas, arba ši speciali sunki formulė yra patenkinama.

Speciali sunki formulė iš tikrųjų nepatenkinama. Todėl, jei pagrindinė įrodymų mašinerija yra tobulai patikima, priimta žinutė vis tiek reiškia, kad tikrasis teiginys teisingas. Taip gaunamas tobulas patikimumas.

Tačiau nulinį žinojimą primenančiam saugumui įsivaizduokime, kad speciali sunki formulė *būtų* patenkinama. Jos liudytojas tada galėtų būti naudojamas simuliuoti įrodymus nežinant tikrojo liudytojo. Realybėje formulė nepatenkinama, tačiau taisyklių sistema negali to efektyviai įrodyti. Vadinasi, ji negali efektyviai įrodyti ir simuliatoriaus neįmanomumo.

Tai pagrindinis lankstas. Sistema nepaslepia paslapties sukurdama klasikinį simuliatorių. Didelei stebimų saugumo testų klasei ji paslepia paslaptį už formalios sistemos negebėjimo patvirtinti, kad simuliatoriaus nėra.

Ką teigia straipsnis

Pagrindinė teorema turi kelis sluoksnius. Esminis rezultatas toks:

Esant standartinei kriptografinei prielaidai — kad egzistuoja **nesąveikūs liudytojo neatskiriamumo įrodymai** (*non-interactive witness indistinguishable proofs*), gerai ištirti objektai, gaunami iš kelių nusistovėjusių prielaidų rinkinių — ir įrodymų sudėtingumo hipotezei, kad **nėra („be galo dažnai“) optimalios įrodymų sistemos**, straipsnyje kiekvienai pasirinktai taisyklių sistemai sukonstruojami vienos žinutės NP/SAT įrodymo kūrėjas ir tikrintojas, turintys **tobulą patikimumą**, nereikalaujantys sąrankos ir esantys efektyviai nulinio žinojimo tos taisyklių sistemos atžvilgiu. (NP/SAT yra standartinis „sunkiausias bendras vardiklis“ galvosūkių tipo problemoms; MegaSudoku — tik vienas jos pavidalas.)

Platesniam teiginiui apie falsifikuojamų saugumo savybių išsaugojimą straipsnyje pridedama dar viena standartinė prielaida — derandomizavimo hipotezė **P = BPP** (apytikriai: atsitiktinumas algoritmams nesuteikia esminės papildomos galios).

Išvertus iš teoremos kalbos:

- Įrodymas yra viena žinutė.
- Patikimos pradinės sąrankos nėra.
- Klaidingų teiginių įrodyti negalima.
- Įrodymo kūrėjas nėra klasikinio nulinio žinojimo — simuliatoriaus nėra.
- Tačiau šioje aplinkoje po vieną galima pasiekti kiekvieną falsifikuojamą, saugumo žaidimais išreiškiamą klasikinio nulinio žinojimo pasekmę.

Žodis „falsifikuojama“ čia svarbus. Jis reiškia, kad saugumo nesėkmę galima patikrinti paleidus priešininką apibrėžtame žaidime. Daugybė kriptografinių saugumo apibrėžimų turi tokią formą: ar priešininkas gali atskirti du šifruotus tekstus, invertuoti funkciją, atkurti liudytoją ar laimėti nurodytą eksperimentą? Teorema kiekvienai tokiai falsifikuojamai savybei suteikia atitinkamą įrodymo kūrėją — po vieną savybę vienu metu. Vienas įrodymo kūrėjas, vienu metu turintis *visas* falsifikuojamas savybes, greičiausiai neįmanomas: senoji pakartotinio panaudojimo ataka („Bobas gali parodyti įrodymą kitiems“) pati yra falsifikuojama savybė ir šioje konstrukcijoje iš tiesų žlunga. Straipsnyje siūloma, kad vienas įrodymo kūrėjas galėtų apimti visas *natūralias* falsifikuojamas savybes — tas, kurios realiai pasitaiko kriptografinėje praktikoje — tačiau ši dalis yra sąlyginė teorema, besiremianti neformalia „natūralumo“ sąvoka ir papildoma aiškia hipoteze. Garantija taikoma stebimoms nesėkmėms, o ne kiekvienai filosofinei ar simuliacija grįstai slaptumo prasmei.

Verta įvardyti vieną konkretų padarinį: konstrukcija suteikia pirmuosius nesąveikius **liudytoją slepiančius** (*witness hiding*) įrodymus su vienodu (*uniform*) įrodymo kūrėju — „galvosūkio įrodymas nepadeda rasti jo sprendinio“, be sąveikos ir be sąrankos. Skamba kukliai, tačiau tokios konstrukcijos nepavyko gauti dešimtmečius.

Ko straipsnis nesako

Ši dalis svarbiausia, kad pasakojimas liktų tikslus.

Straipsnis **nesako**, kad seni neįmanomumo teoremos buvo klaidingos. Konstrukcija jas apeina pakeisdama apibrėžimą.

Jis **nesuteikia** įprasto klasikinio nulinio žinojimo be sąveikos, be sąrankos ir su tobulu patikimumu. Straipsnyje aiškiai pasakyta, kad sukonstruotas įrodymo kūrėjas neturi simuliatoriaus.

Tai **nereiškia**, kad įrodymo negalima pakartotinai panaudoti. Vienos žinutės įrodymą vis tiek galima parodyti kam nors kitam; straipsnio konstrukcija neišsaugo paneigiamumo (*deniability*) tipo savybių. (Nesąveikus nulinis žinojimas su patikima sąranka turi tą patį apribojimą.)

Tai **nereiškia**, kad turime praktinį, diegti paruoštą protokolą. Tai sudėtingumo teorijos ir kriptografi-

jos pamatų darbas. Rezultatas priklauso nuo svarbių įrodymų sudėtingumo ir kriptografijos prielaidų ir kalba apie principinę galimybę.

Tai **nepaverčia „Gödelio“ magiška saugumo primityva**. Ryšys su Gödelio idėjomis eina per įrodymų sistemas, optimalias įrodymų sistemas ir baigtinius nepilnumo analogus. Naudinga intuicija nėra „nepilnumas apsaugo jūsų slaptažodį“. Ji tokia: jei taisyklių sistema negali efektyviai įrodyti, kad simuliatorius neįmanomas, atakos, kurioms reikėtų tokio įrodymo, gali būti blokuojamos jau saugumo apibrėžimų lygmenyje.

Kodėl tai vis tiek įdomu

Kriptografija dažnai paverčia sunkumą saugumu. Faktorizuoti sunku, todėl RSA tipo prielaidos tampa naudingos. Gardelių uždaviniai sunkūs, todėl naudinga gardelių kriptografija. Čia sunkumas keistesnis: ne „sunku apskaičiuoti paslaptį“, o „sunku įrodyti, kad tam tikras įrodymo objektas negali egzistuoti“.

Būtent todėl straipsnis neįprastas. Aksiomas ir formalias taisyklių sistemas jis traktuoja beveik kaip kriptografinius išteklius. Įprastas neįmanomumas sako, kad tarp patikimumo ir simuliavimo yra įtampa. Ilango žingsnis — paslėpti šią įtampą už įrodymų teorijos uždangos: simuliatoriaus nėra, tačiau formali sistema negali efektyviai parodyti jo nebuvimo.

Skaitytojui netikėčiausia ne tai, kad ši konstrukcija pakeis dabartines nulinio žinojimo sistemas. Bent tiesiogiai greičiausiai nepakeis. Netikėta tai, kad matematinės logikos apribojimą galima panaudoti konstruktyviai — ne tik kaip sieną, bet ir kaip savotišką priedangą.

Kiek tvirti yra įrodymai?

Tai teoremų straipsnis, todėl „įrodymai“ čia reiškia ką kita nei biologijoje ar astronomijoje. Klausimas ne tas, ar eksperimentas pakartotas. Klausimas — ar apibrėžimai, prielaidos ir įrodymų grandinė pagrindžia teiginį.

Įrodymas formalus, o prielaidos straipsnyje aiškiai įvardytos. Jos nėra atsitiktinės. Nesąveikūs liudytojo neatskiriamumo įrodymai yra standartiniai kriptografijos objektai ir seka iš kelių nusistovėjusių prielaidų rinkinių. Hipotezė, kad nėra optimalios įrodymų sistemos, yra centrinė įrodymų sudėtingumo hipotezė. P = BPP yra standartinė derandomizavimo prielaida, naudojama tik platesnei falsifikuojamų savybių teoremai.

Straipsnyje taip pat argumentuojama, kad tai yra tinkama kaina, o ne savavališkos atramos. Įrodomas ir atvirkštinis rezultatas, rodantis, kad prielaidos iš esmės būtinos: jei tokios konstrukcijos apskritai egzistuoja, tada turi egzistuoti nesąveikūs liudytojo neatskiriamumo įrodymai ir — darant standartinę vienkrypčių funkcijų prielaidą — negali egzistuoti optimali įrodymų sistema. Be to, prielaidos yra „win-win“ pobūdžio: bet kurios jų paneigimas pats savaime būtų reikšmingas atradimas įrodymų sudėtingumo, kriptografijos ar sudėtingumo teorijoje.

Tačiau rezultatas sąlyginis, todėl sąlyginis ir pasitikėjimas jo interpretacija. Jei prielaidos neteisingos, teoremos reikšmė keičiasi. Ir net jei jos teisingos, garantija nėra pilnas klasikinis nulinis žinojimas; tai susilpninta, įrodymų teorija paremta straipsnio versija.

Todėl tinkamas pasitikėjimo lygis toks: didelis, kad straipsnis pateikia nuoseklų sąlyginį galimybės rezultatą; vidutinis, kad jo prielaidos aprašo kriptografinį pasaulį, kuriame realiai gyvename; mažas, kad iš to greitai atsiras praktinis taikymas.

Kodėl tai svarbu

Straipsnis atveria kelią, kuris buvo laikomas uždarytu.

Klasikinė teorija sako: pilnas nulinis žinojimas negali būti viena žinutė be sąrankos ir negali turėti tobulo patikimumo. Ilango straipsnis sako: jei klausiame apie nulinio žinojimo pasekmes, kurias galima išbandyti saugumo žaidimuose, ir leidžiame saugumo apibrėžimui priklausyti nuo to, ką pasirinkta taisyklių sistema gali arba negali efektyviai paneigti, galima atkurti didelę dalį naudingo elgesio — su viena žinute, be sąrankos ir su tobulu patikimumu.

Tai nėra mažas apibrėžimo kosmetinis pakeitimas. Tai kitoks būdas mąstyti apie kriptografines garantijas. Užuot klausę tik, kas egzistuoja, galime klausti, ką mūsų taisyklių sistema pajėgi atmesti. Užuot laikę neįrodomumą filosofiniu nepatogumu, galime panaudoti jį kaip struktūrą.

Praktinis pasaulis rytoj gal ir nepasikeis. Tačiau sąvokų žemėlapis jau pasikeičia. Dabar turime formalų kontekstą, kuriame „niekas negali efektyviai įrodyti, kad paslaptis nutekėjo“ gali būti pakankamai stipru, kad būtų atkurtos daugelis žaidimais grindžiamų apsaugų, kurių norėjome iš teiginio „paslaptis nenutekėjo“.

Todėl Gödelis ir yra pavadinime.

Trumpa santrauka

Nulinio žinojimo įrodymai leidžia įrodymo kūrėjui įtikinti tikrintoją, kad teiginys teisingas, neatskleidžiant liudytojo. Klasikiniai neįmanomumo rezultatai sako, kad nulinio žinojimo negalima sutalpinti į vieną žinutę be sąrankos ir kartu išlaikyti tobulo patikimumo. Rahulio Ilango straipsnis šių neįmanomumų nepaneigia. Jis apibrėžia silpnesnę sąvoką — efektyvų nulinį žinojimą: vietoj reikalavimo, kad simuliatorius iš tikrųjų egzistuotų, reikalaujama, kad pasirinkta įrodymų sistema — formali taisyklių knyga, pavyzdžiui, ZFC — negalėtų efektyviai įrodyti, jog simuliatoriaus nėra. Esant svarbioms kriptografijos prielaidoms (nesąveikiems liudytojo neatskiriamumo įrodymams) ir įrodymų sudėtingumo prielaidai (kad optimali įrodymų sistema neegzistuoja), straipsnis konstruoja vienos žinutės NP/SAT įrodymo kūrėjus be sąrankos ir su tobulu patikimumu, kurie po vieną pasiekia falsifikuojamas, saugumo žaidimais grindžiamas klasikinio nulinio žinojimo pasekmes. Vienas universalus įrodymo kūrėjas visoms „natūralioms“ tokioms savybėms yra tolesnis, iš dalies hipotezėmis paremtas išplėtimas, o pažodžiui visas falsifikuojamas savybes vienu metu turėti greičiausiai neį-

manoma, nes įrodymai išlieka pakartotinai panaudojami. Rezultatas teorinis ir sąlyginis, ne paruošta kriptografinė primityva, tačiau jis parodo naują būdą įrodymų teorinį neįrodomumą naudoti kaip kriptografinį išteklių.

Be pagražinimų

Ką straipsnis parodo: esant nurodytoms prielaidoms, galima sukurti vienos žinutės, sąrankos nereikalaujančius ir tobulai patikimus NP/SAT įrodymo kūrėjus, kurie pasirinktos įrodymų sistemos atžvilgiu yra efektyviai nulinio žinojimo ir po vieną pasiekia kiekvieną falsifikuojamą, saugumo žaidimais grindžiamą klasikinio nulinio žinojimo pasekmę.

Kas tikėtina, bet besąlygiškai neįrodyta: kad reikalingos įrodymų sudėtingumo ir kriptografinės prielaidos iš tiesų teisingos. Tai rimtos, plačiai tyrinėtos prielaidos, o straipsnis parodo, kad jos iš esmės ne tik pakankamos, bet ir būtinos — tačiau jos vis tiek lieka prielaidomis.

Ko tai nerodo: klasikinio nulinio žinojimo be sąveikos, be sąrankos ir su tobulu patikimumu; diegti paruoštos praktinės sistemos; įrodymų paneigiamumo ar nepakartotinio panaudojimo; arba kad vien Gödelio nepilnumo teorema savaime užtikrina kriptografinį saugumą.

Svarbiausi apribojimai: garantija yra susilpninta nulinio žinojimo forma; plačiausia versija remiasi keliomis prielaidomis; vieno universalaus įrodymo kūrėjo teiginiai lieka iš dalies hipoteziniai; rezultatas pirmiausia pamatinis ir teorinis.

Kiek tuo turėtų pasitikėti bendras skaitytojas? Didelis pasitikėjimas, kad, priėmus apibrėžimus ir prielaidas, tai svarbus sąlyginis teorinis rezultatas. Vidutinis pasitikėjimas, kad prielaidos atspindi tikrąjį skaičiavimo ir kriptografijos pasaulį. Mažas pasitikėjimas tiesioginiu praktiniu pritaikymu artimiausiu metu. Saugi išvada: straipsnis nepanaikina nulinio žinojimo neįmanomumo teoremų; jis randa naują, įrodymų teorija paremtą būdą apeiti tas jų dalis, kurios svarbios daugeliui saugumo žaidimų.

Šaltiniai

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