The Clean Paper

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

Trikkið er ekki að sanna að leyndarmálið sé falið

Núllþekkingarsannanir eiga að sannfæra án þess að afhjúpa vitnið, en klassísk kenning leyfir ekki fulla útgáfu þess með einu skeyti, án trausts upphafs og með fullkomnum hljóðleika. Rahul Ilango breytir markmiðinu: í stað þess að krefjast þess að hermir sé til er krafist að valið sönnunarkerfi geti ekki með skilvirkum hætti sannað að enginn hermir sé til. Undir sterkum forsendum um sönnunarflækju og dulmálsfræði endurheimtir þetta prófanlegar afleiðingar núllþekkingar eiginleika fyrir eiginleika — fræðilegt öryggishugtak, ekki tilbúinn netstaðall.

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

Trikkið er ekki að sanna að leyndarmálið sé falið

Byrjum á einföldustu útgáfu af **núllþekkingu** (*zero-knowledge*).

Alice vill sannfæra Bob um að Sudoku-þraut hafi lausn. Ef hún sendir honum lausnina sannfærist hann, en þrautin er eyðilögð. Það sem hún vill er undarlegra: sönnun þess að lausn sé til, án þess að afhjúpa lausnina.

Það er loforð núllþekkingarsönnunar. Sannandinn, Alice, sannfærir sannprófarann, Bob, um að staðhæfing sé sönn án þess að afhjúpa neitt umfram sannleiksgildi staðhæfingarinnar.

Vandinn er að þetta loforð kostar eitthvað. Venjuleg stærðfræðileg sönnun hefur tvo þægilega eiginleika. Hún er **eitt skeyti**: þú skrifar hana niður, afhendir hana og ferð. Og hún er **fullkomlega hljóð** (*perfectly sound*): röng staðhæfing á enga gilda sönnun. Klassískar ómöguleikaniðurstöður segja að núllþekking geti ekki haldið báðum þessum eiginleikum — og raunar er hvor um sig vandamál einn og sér.

Í fyrsta lagi þarf klassísk núllþekking gagnvirk samskipti. Ef Alice sendir eitt skeyti, án trausts upphafsfyrirkomulags sem samið var um fyrirfram, hrynur núllþekkingartryggingin — sama hversu mikið þú ert reiðubúinn að gefa eftir af hljóðleika.

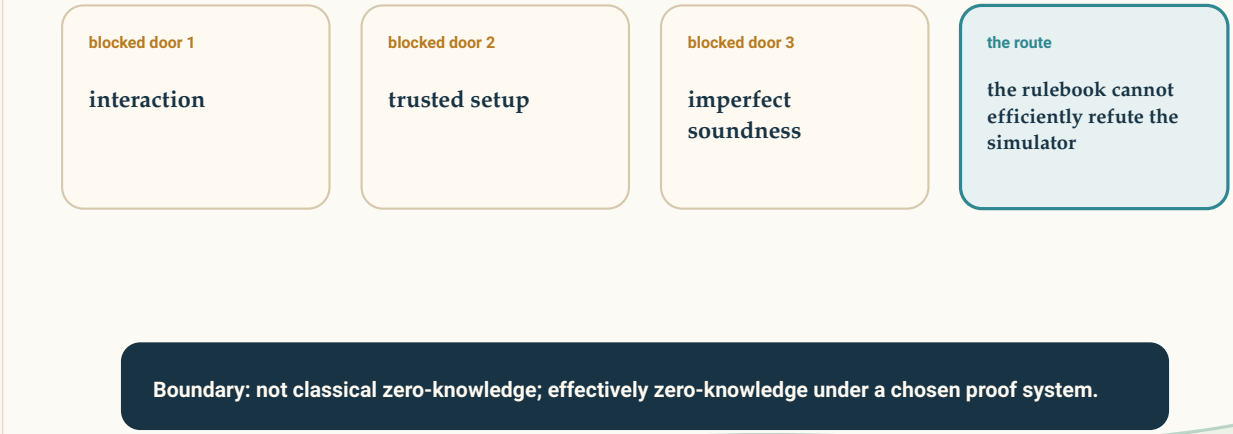
Í öðru lagi þarf núllþekking smá þol fyrir villu. Krafa um fullkominn hljóðleika reynist líka eyðileggja gagnvirknina: sannprófari sem er **aldrei** hægt að blekkja, óháð tilviljanakenndum valkostum hans, gæti í raun fest þessi val fyrirfram. Þegar sannprófarinn verður fyrirsjáanlegur getur Alice svarað öllu í einu skeyti — og þá erum við aftur í tilvikinu sem þegar var ómögulegt.

Grein Rahul Ilango fjallar um leið fram hjá þessum tvöfalda vegg. Ekki með því að láta sem veggurinn sé ekki til og ekki með því að smíða klassíska núllþekkingu í umhverfi þar sem hún er ómöguleg. Breytingin er fínni: veikja hvað „afhjúpar ekkert“ merkir, en gera það á þann hátt að öryggiseiginleikar sem dulmálsfræðingar geta **raunverulega prófað** haldist.

Nýja hugtakið heitir **effectively zero-knowledge** — hér kallað **virk núllþekking**.

A proof without leaking the witness

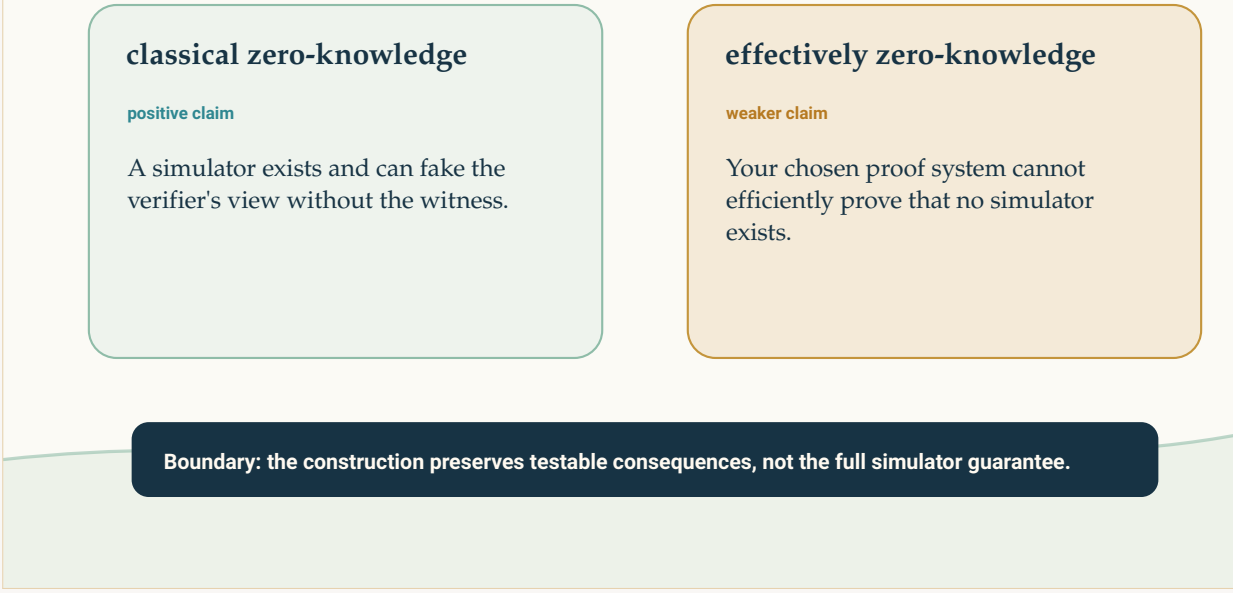
The paper keeps the ordinary-proof shape by weakening what privacy must prove.



Klassísk núllþekking rekst á þrjár dyr — gagnvirk samskipti, traust upphafsfyrirkomulag og ófullkominn hljóðleika. Smíði Ilango fer aðra leið: reglubókin getur ekki á skilvirkan hátt afsannað möguleika á hermi.

Classical privacy vs effective privacy

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



Klassísk núllþekking spyr hvort hermir sé í raun til; „virk núllþekking“ spyr aðeins hvort valin reglubók geti á skilvirkan hátt sannað að enginn slíkur hermir geti verið til. Veikari spurningin er það sem gerir kleift að halda

einu skeyti, engu upphafsfyrirkomulagi og fullkomnum hljóðleika.

Gamla prófið: hermir er til

Klassíska formlega skilgreiningin á núllþekkingu notar ímyndaðan hjálpara sem kallast **hermir** (*simulator*).

Hugmyndin er þessi: ímyndum okkur Jane sem **þekkir ekki leyndarmál Alice**. Ef Jane getur, alveg sjálf, framleitt sönnunargögn sem líta út eins og þau sem Bob hefði fengið frá Alice, þá kenndi sönnun Alice Bob ekkert nýtt um leyndarmálið. Jane gat þegar falsað alla reynsluna án þess að vita leyndarmálið.

Klassísk núllþekking krefst því raunverulegs hermis: skilvirks reiknirits sem getur búið til trúverðug sönnunargögn án þess að þekkja **vitnið** (*witness*), sem er hugtakið fyrir leynilega upplýsinguna sem staðfestir staðhæfinguna. Í Sudoku er vitnið einfaldlega fullleysta reiturinn.

Skilgreiningin er öflug, en einmitt þar bíta klassísku ómöguleikarnir. Einföld innsýn er sú að algerlega ógagnvirk sönnun er bara strengur. Þegar Bob hefur strenginn getur hann sýnt hann einhverjum öðrum; hann hefur þá öðlast getu til að sannfæra þriðja aðila, sem hljómar þegar eins og meira en „ekkert“. Klassísku setningarnar gera þessa innsýn nákvæma.

> **Þrír eiginleikar sem greinin neitar að gefa upp**
>
> Titill greinarinnar nefnir þrjár kröfur:
> **Engin gagnvirkni:** Alice sendir einn sönnunarstreng. Það er ekkert fram-og-til-baka samskiptakerfi.
> **Ekkert upphafsfyrirkomulag:** Alice og Bob treysta ekki á sameiginlegan traustan viðmiðunarstreng eða aðra opinbera tilviljun sem búin var til fyrirfram. Mörg kerfi sem kallast „non-interactive zero-knowledge“ nota samt upphafsfyrirkomulag; hér þýðir hugtakið bókstaflega ekkert slíkt.
> **Fullkominn hljóðleiki:** röng staðhæfing á enga gilda sönnun. Ekki „nánast aldrei samþykkt“, heldur **engin** gild sönnun.
> Þetta eru sömu þrír eiginleikar og venjuleg skrifuð stærðfræði hefur — og klassísk núllþekking getur ekki haldið þeim öllum.

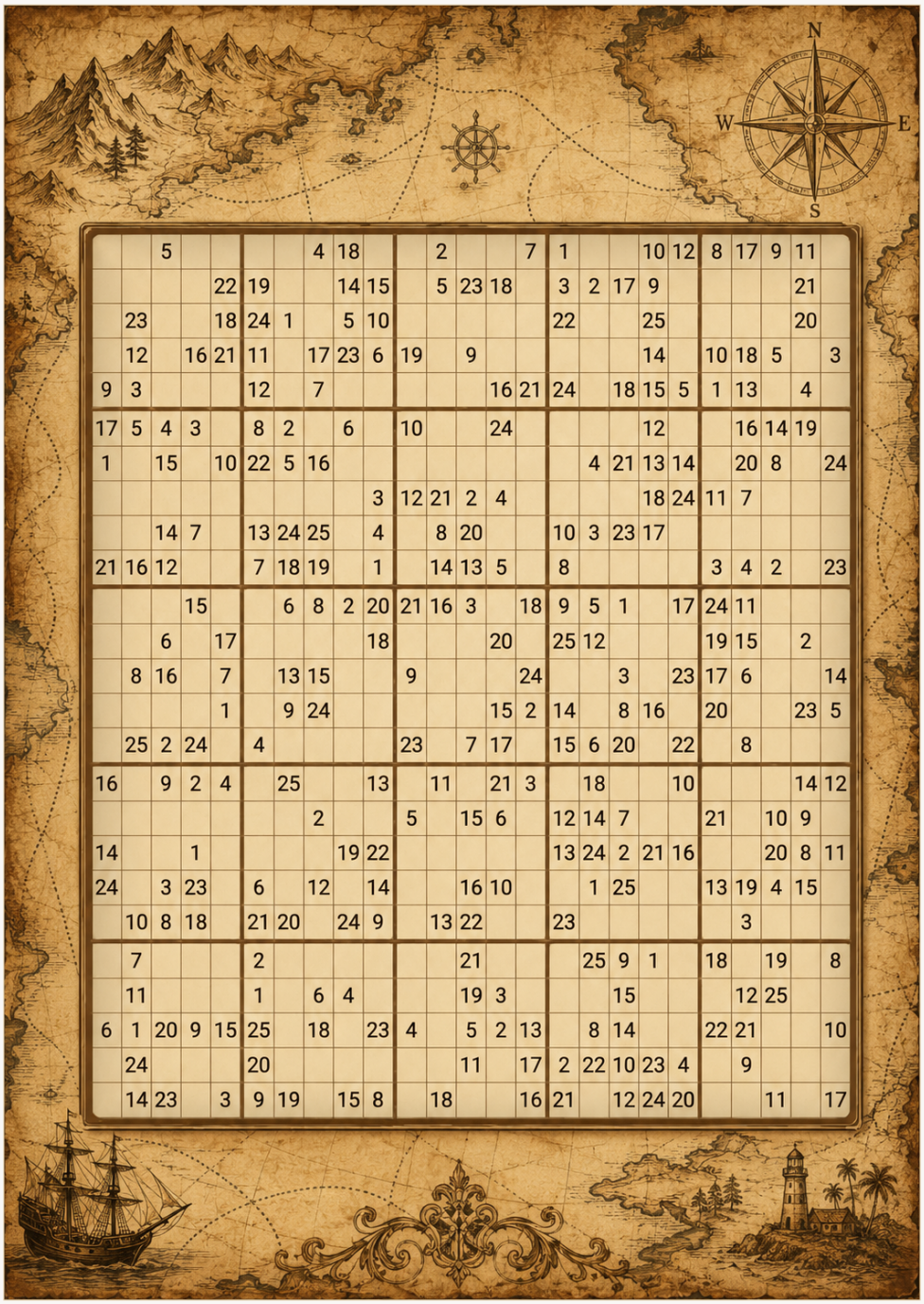
MegaSudoku sem tilfinning fyrir muninum

Hér er viljandi einföld mynd af muninum.

Ekki nota venjulegt 9×9 Sudoku fyrir alvarlega hlutann í líkingunni. Það er of lítið og endanlegt: tölva getur bara leyst það eða sýnt að engin lausn sé til. Ímyndum okkur í staðinn fjölskyldu af

MegaSudoku(n). Veljum blokkastærð n, látum $N = n^2$ og byggjum $N \times N$ reit sem skiptist í $n \times n$ blokkir og notar N tákn. Venjulegt Sudoku er örsmáa tilvikið $n = 3$, $N = 9$. Sönnunarflækjusagan byrjar fyrst þegar n má vaxa og þegar hægt er að bæta við sérsmíðuðum „græjum“ sem láta reitinn hegða sér eins og SAT-formúlu í Sudoku-búningi.

SAT er listi af já/nei-skilyrðum: er hægt að gefa breytunum gildin satt/ósatt þannig að öll skilyrðin verði uppfyllt?



25×25 Sudoku: hægt er að sannreyna reglurnar án þess að sýna fullgerða reitinn — sjónræn staðgengilsmynd fyrir sönnun sem staðfestir falið vitni.

Sudoku og SAT: sama þrautin í tveimur búningum

Fullyrðingin að Sudoku geti „hegðað sér eins og SAT-formúla“ er ekki myndlíking. Þýðingin gengur í báðar áttir og auðveldari áttina er hægt að skrifa niður alveg.

Frá Sudoku yfir í SAT. SAT talar aðeins satt/ósatt, svo við gefum því eina Boole-breytu fyrir hvert þrennt (röð, dálkur, gildi): *x(r,c,v)* þýðir „reitur í röð *r*, dálki *c* inniheldur gildið *v*“. 4×4 Sudoku með 2×2 blokkum og gildunum 1–4 þarf 4·4·4 = **64 breytur**; klassískt 9×9 þarf **729**. Hver Sudoku-regla verður svo að hópi klása. Klási er OR af breytum eða neitunum þeirra; öll formúlan er AND allra klása.

Hver reitur hefur að minnsta kosti eitt gildi — einn klási á reit:

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

Hver reitur hefur í mesta lagi eitt gildi — „ekki bæði“ klási fyrir hvert par:

$$\neg x(1,1,1) \vee \neg x(1,1,2) \quad \neg x(1,1,1) \vee \neg x(1,1,3) \quad \text{… og svo framvegis fyrir öll sex pörin.}$$

Hver röð inniheldur hvert gildi — fyrir röð 1 og gildið 3, að minnsta kosti einu sinni:

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

og í mesta lagi einu sinni: $\neg x(1,1,3) \vee \neg x(1,2,3)$, og sambærilegir klásar fyrir öll pör reita í röðinni.

Dálkar og blokkir — alveg sams konar skilyrði; aðeins reitahópurinn breytist. Fyrir efstu vinstri blokkina og gildið 2:

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

ásamt „ekki bæði“ klásum fyrir hvert par.

Prentuðu vísbendingarnar — einfaldasti hlutinn: hver vísbending verður klási með einni breytu. Prentuð 3 efst til vinstri verður

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

AND allra þessara skilyrða er uppfyllanlegt nákvæmlega þegar Sudoku hefur lausn — og fullnægjandi gildisúthlutun **er** lausnin: lesið hvaða x(r,c,v) eru sönn og fyllið reitinn. Fyrir 9×9 verða þetta 729 breytur og nokkur þúsund klásar, sem nútíma SAT-leysir afgreiðir á millisekúndum. Athugið vísbendingarklásann x(1,1,3): hann segir „þessi reitur er nákvæmlega 3“, ekki „þessir reitir eru allir ólíkir“. Þessi ósamhverfa er ástæðan fyrir aukatrikkinu með vísbendingareitina í næstu athugasemd.

Frá SAT yfir í Sudoku. Greinin þarf hina, erfiðari áttina: gefin **handahófskennd SAT-formúla**, byggið MegaSudoku sem hefur lausn nákvæmlega þegar formúlan er uppfyllanleg. Innfæddar reglur Sudoku geta í grundvallaratriðum aðeins sagt „þessir reitir eru allir ólíkir“, þannig að almenn rökskilyrði verða að vera **smíðuð**. Þar koma „græjurnar“ inn. Græja er lítið fyrirfram hannað safn reita, eitt fyrir hvern klása, þar sem valdir reitir gegna hlutverki breyta — táknið sem þeir bera kóðar satt eða ósatt — og innri skilyrði græjunnar eru hönnuð þannig að einu löglegu fyllingarnar séu einmitt þær sem uppfylla klásann. Þetta er venjulegt handverk úr NP-fullkomnissönnunum; fyrir almennt Sudoku gerðu Yato og Seta þetta árið 2003.

Saman segja áttirnar tvær að N×N Sudoku og SAT séu sama vandamál í ólíkum búningum. Þess vegna má bæði greinin og þessi umfjöllun tala um allt NP með reitum og táknum.

Vitnið er enn auðvelt að sjá fyrir sér. Alice þekkir fulla löglega lausn MegaSudoku. Bob vill sannfærast um að slík lausn sé til en Alice vill ekki sýna hana. Ef hún sendir alla lausnina sannfærist Bob en leyndarmálið er farið.

Í klassískri núllþekkingu eiga Alice og Bob samskipti. Ein gömul hugarmynd notar hulda kubba. Alice felur lausnina, endurnefnir táknin leynilega fyrir hverja umferð og leyfir Bob að skoða eitt tilviljanakennt staðbundið skilyrði: röð, dálk, blokk eða græju. Ef opnu reitirnir sýna öll ólík tákn eykst traust Bob. Síðan er allt hulið aftur og táknin endurnefnd að nýju.

Ein flækja er að prentuðu vísbendingarnar þurfa aukatrikk, því endurnefningin felur þær líka. Athugasemdin hér að neðan sýnir hvernig klassískar aðferðir leysa það.

Hvernig klassísk kerfi meðhöndla vísbendingareitina í raun

Endurnefningartæknin hefur blindan blett. Reglur um raðir, dálka og blokkir segja allar „þessir reitir eru allir ólíkir“, og eiginleikinn **allir ólíkir** helst undir hvaða endurnefningu tákna sem er. En vísbending segir „þessi reitur inniheldur nákvæmlega 5“. Eftir endurnefningu sér Bob aðeins $\sigma(5)$, einhverja dulbúna táknmynd, án þess að vita vörpunina σ. Hann getur þá ekki athugað gildið. Ef ekkert er gert gæti Alice sannað að *einhver* lögleg Sudoku-fylling sé til á meðan hún hunsar prentuðu vísbendingarnar — sem segir ekkert um *þessa* þraut. Klassískar aðferðir hafa tvær algengar viðgerðir.

Litapallettan. Bættu einni aukaröð með N reitum við falda reitinn — pallettu sem Alice fyllir með táknunum 1…N í föstri opinberri röð og endurnefnir síðan með öllu hinu, þannig að hún inniheldur $\sigma(1)$…$\sigma(N)$. Handahófsáskorun Bob fær nú einn aukakost. Fyrir utan röð, dálk, blokk eða græju getur hann valið **pallettuna og einn vísbendingareit**. Alice afhjúpar bæði; pallettan sýnir endurnefningu þeirrar umferðar og Bob athugar að vísbendingareiturinn beri nákvæmlega endurnefnda útgáfu af prentuðu tölunni. Þetta helst núllþekking vegna þess að Bob lærir aðeins σ — nýja tilviljanakennda vörpun sem er gagnslaus ein og sér — og gildi reits sem hann vissi þegar. Leynireitirnir leka ekki og hermir getur líkt eftir sýninni með slembnu σ. Kerfið er hljóð því svindlandi Alice er gripin með föstum líkum í hverri umferð og umferðir eru endurteknar þar til efinn verður hverfandi.

Vísbendingarnar „þýddar“ yfir í skilyrði. Byggingalegri útgáfa fjarlægir sérstaka áskorun í stað þess að bæta henni við. Í stað þess að *staðfesta* gildi vísbendingar er það þvingað fram með ójöfnuskilyrðum: tengdu vísbendingareitinn við alla pallettureiti nema þann sem ber hans eigið gildi — „ólíkur $\sigma(1)$, ólíkur $\sigma(2)$, …, ólíkur öllu nema $\sigma(5)$“. Eina táknið sem reiturinn getur löglega haft er þá vísbendingin. Öll skilyrðin eru aftur af gerðinni „þessir tveir eru ólíkir“, sem lifir endurnefningu og er athuganlegt eins og röð. Þetta er sama bragð og fyrir fyrirfram litaða hnúta í klassískum sönnunum um grafalitun og er kjarninn í „græjunum“ í SAT↔MegaSudoku sögunni.

Raunverulegt spilakerfi. Spilasamskiptaregla fyrir Sudoku eftir Gradwohl, Naor, Pinkas og Rothblum (2007) notar enga endurnefningu og festir vísbendingarnar áður en feluleikurinn byrjar. Fyrir hvern reit leggur Alice þrjú eins spil með gildi reitsins — á hvolfi fyrir leynireiti en **upp fyrir vísbendingareiti**, þannig að Bob sér strax að prentuðu tölurnar eru virtar. Síðan fer eitt spil úr hverjum reit í röðarpakka, eitt í dálkapakka og eitt í blokkpakka; hver pakki er stokkaður og sýndur og Bob athugar að hann innihaldi öll N tákn. Stokkunin eyðir

> stöðuupplýsingunum — það er núllþekkingarhlutinn — en vísbendingarnar voru staðfestar strax við uppsetningu.
> Lærdómurinn er alltaf sá sami: núllþekkingarsamskiptaregla lýsir nákvæmlega **hvaða staðreyndir lifa feluleikinn af**. Endurnefning varðveitir „allir ólíkir“ en eyðir „jafnt 5“, svo „jafnt 5“ verður að koma aftur inn með annarri aðferð.

Þetta er **ekki** samskiptaregla greinarinnar. Það er aðeins hugarmynd fyrir klassíska núllþekkingu:

- Alice og Bob eiga fram-og-til-baka samskipti.
- Bob velur tilviljanakenndar athuganir.
- Alice sýnir aðeins staðbundið samræmi, ekki alla lausnina.
- Persónuverndarsönnunin sýnir að sýn Bob hefði getað verið búin til án leynilausnar Alice.

Klassísk núllþekking byggist því á jákvæðri staðreynd:

> Hermir er raunverulega til.

Nú fjarlægjum við þægilegu hlutina. Alice sendir einn sönnunarstreng og fer. Það er ekkert traust upphafsfyrirkomulag, enginn sameiginlegur tilviljanakenndur strengur undirbúinn fyrirfram og Bob má **aldrei** samþykkja ranga þraut. Þetta er stillingin þar sem klassísk núllþekking lifir ekki af.

Við þurfum einn nýjan „persónu“ áður en trikkið kemur. Festum **reglubók**: formlegt sönnunarkerfi í skilningi rökfræðings — fasta mengið af frumsendum og vélrænum reglum til að athuga skrifaðar stærðfræðilegar sannanir. **ZFC**, venjulegar frumsendur nútímastærðfræði, er dæmigerða dæmið. Allt sem fylgir er skilgreint miðað við eina reglubók sem valin er fyrirfram; smíðin virkar fyrir hvaða fasta reglubók sem er, ZFC þar með talið.

(Orðanóta úr greininni: „proof system“ þýðir hér alltaf þessa reglubók, formlega kerfið sem sannreynir stærðfræðilegar sannanir, **ekki** skeytin sem Alice sendir. Tæki Alice og Bob eru kölluð sannandi og sannprófari.)

Gödel-stíls útgáfan heldur MegaSudoku-sögunni en breytir sönnuninni.

Veljum annað skilyrðakerfi af sömu sýndu stærð og köllum það **D**. Í sögunni eru S og D tvær MegaSudoku(n)-þrautir í sama sniði. Undir húddinu gæti D hafa byrjað sem erfið rökformúla af annarri stærð; ef þarf má fylla hana upp með saklausum aukaskilyrðum. D er byggt úr formúlu sem er **í raun óuppfyllanleg**: ekkert gildismat getur uppfyllt öll skilyrði hennar, rétt eins og brotin Sudoku-þraut hefur enga löglega fyllingu. Barnalegt dæmi væri formúla sem krefst bæði „X er satt“ og „X er ósatt“. D hefur því enga gilda lausn.

En D má ekki vera brotin þraut sem er **auðvelt að afhjúpa**. Barnalega dæmið er gagnslaust því hvaða reglubók sem er afsannar „X og ekki-X“ í einni línu. D þarf að vera ósatt á þann hátt að valda reglubókin geti ekki vottað ósannleikann með stuttri sönnun. Ef reglubókin gæti afsannað D fljótt myndi sagan hrynja: þá mætti formlega útiloka valleiðina sem hefði getað búið til sannanir án leyndarmáls Alice, og með henni hyrfi persónuverndartryggingin. D er því valið úr fjölskyldu sem

fasta reglubókin getur ekki **afsað á skilvirkan hátt**: innan reglubókarinnar er engin stutt sönnun þess að D hafi enga lausn.

Einskiptissönnun Alice snýst svo um annaðhvort/eða-staðhæfingu:

> annaðhvort hefur raunverulega MegaSudoku S lausn, eða blekkingarþrautin D hefur lausn.

Þetta er rökfræðilega tengingin. D er **ekki** framleitt með töfrum til að gera S satt. Sönnunin segir ekki „D hefur enga lausn, því hefur S lausn“. Hún sannar **S eða D**. Fullkominn hljóðleiki segir að röng disjunction geti ekki haft gilda sönnun. Þar sem D er í raun ósatt — hefur enga lausn — getur „S eða D“ aðeins verið satt ef **S er satt**. Ef sönnunin er samþykkt verður S því að hafa lausn. Blekkingarþrautin getur ekki gert ranga S að sannri staðhæfingu.

En fyrir núllþekkingarhlutann skulum við spyrja hvað myndi gerast ef D **hefði** lausn. Sú lausn gæti virkað sem annað vitni og gert einhverjum kleift að framleiða sönnunargögn án þess að vita raunverulega MegaSudoku-lausn Alice — það er, hún gæti orðið grunnur að hermi. Í veruleikanum hefur D enga lausn, þannig að þessi leið er lokuð. Kjarni hugmyndarinnar er að **reglubókin getur ekki á skilvirkan hátt sannað að hún sé lokuð**.

D gegnir því tveimur hlutverkum. Fyrir **hljóðleika** er D ósatt, þannig að gild sönnun á „S eða D“ þvingar S til að vera satt. Fyrir **virka núllþekkingu** er D erfitt að afsanna, þannig að reglubókin getur ekki fljótt útilokað leiðina sem hefði gert hermun mögulega.

Öryggisspurningin er því ekki lengur:

> Getum við sannað að hermir sé raunverulega til?

Heldur:

> Getur reglubókin þín á skilvirkan hátt sannað að hermir sé ómögulegur?

Ef svarið er nei gerist eitthvað sterkt. Sérhver öryggiseiginleiki sem (a) má mæla með keyrslu prófs og (b) má sanna, **innan reglubókarinnar**, út frá tilvist hermis, gildir í raun. Vel heppnuð árás á slíkan eiginleika myndi sjálf gefa stuttu afsönnunina sem reglubókin á ekki. Það er „virki“ hlutinn í virkri núllþekkingu.

Andstæðan í kennslustofunni er því:

Klassísk núllþekking: sönnunin er örugg því hermir er til.

Gödel-stíls virk núllþekking: sönnunin er örugg gagnvart mælanlegum öryggisprófum vegna þess að reglubókin getur ekki á skilvirkan hátt sannað að hermir sé ómögulegur.

Seinni fullyrðingin er veikari. Einmitt þess vegna getur smíðin haldið eiginleikunum þremur sem brutu klassísku útgáfuna: eitt skeyti, ekkert upphafsfyrirkomulag og fullkominn hljóðleiki.

Nýja prófið: þú getur ekki sannað að herminn vanti

Slökun Ilango breytir spurningunni.

Klassísk núllþekking spyr:

> Er hermir til?

Virk núllþekking spyr veikari spurningu:

> Getur valin reglubók á skilvirkan hátt sannað að enginn hermir sé til?

Þetta hljómar eins og tæknilegt undanfæri, en er kjarninn. Smíðin býr í undarlegu ástandi: hermir er **ekki** raunverulega til — greinin segir það skýrt — en fasta reglubókin getur ekki á skilvirkan hátt sannað að hann sé ekki til. Ef allir slæmu öryggisatburðirnir sem þú óttast myndu krefjast slíkrar afsönnunar hegðar kerfið sér samt eins og núllþekking gagnvart þeim atburðum.

Hér kemur Gödel inn. Ekki sem skraut og ekki sem „Gödel gerir dulmál öruggt“. Tengingin er sönnunarfræðileg. Reglubók er kölluð **bestmöguleg** (*optimal*) ef hún er í nákvæmum skilningi sú besta: hvenær sem einhver reglubók getur afsannað formúlu af viðkomandi gerð með stuttri sönnun getur bestmögulega reglubókin líka gert það, með í mesta lagi margliðulega lengri sönnun.

Krajíček og Pudlák settu árið 1989 fram tilgátuna að **ekkert bestmögulegt sönnunarkerfi sé til**. Hvaða reglubók sem þú festir er til önnur sem getur sannað einhverja fjölskyldu sannra staðhæfinga mun styttra. Þetta er ein af meginopnu tilgátum sönnunarflækju og er endanlegur, flækjufræðilegur frændi ófullkomleikasetningar Gödel: sumar sannar staðhæfingar hafa enga stutta sönnun í reglubókinni sem þú valdir — ekki vegna þess að þær séu ósannanlegar yfir höfuð, heldur vegna þess að hvert fast kerfi skilur einhverjar stuttar sannleiksstaðhæfingar eftir án stuttra sannana.

Greinin gerir ráð fyrir þessari tilgátu í örlítið sterkari útgáfu sem gildir „óendanlega oft“ (*infinitely often*) sem er algeng þegar tilgátur eru notaðar dulmálsfræðilega. Með setningu Krajíček og Pudlák fæst þá nákvæm niðurstaða: fyrir hverja reglubók er til röð formúla sem eru í raun óuppfyllanlegar en reglubókin getur ekki afsannað með stuttri sönnun — og, sem skiptir öllu, **skilvirkt reiknirit getur framleitt þær**. Sú staðreynd að eitt skilvirkt reiknirit getur framleitt dæmin breytir hugmyndinni úr hreinni tilvist í aðferð sem Alice getur raunverulega keyrt. D-blekkingarnar koma af færibandi, ekki úr tómarúmi.

Dulmálsbragðið er að nota þennan skort á sönnunarkrafti sem auðlind.

Hvað smíðin gerir

Hér er form greinarinnar, svipt niður í beinagrindina.

Festu reglubók — segjum ZFC. Samkvæmt sönnunarflækjutilgátunni er hægt að framleiða á skilvirkan hátt röð formúla sem eru **í raun óuppfyllanlegar**, en reglubókin hefur enga stutta sönnun á óuppfyllanleikanum.

Byggðu síðan eins-skeytis sönnun af gerðinni:

> annaðhvort er raunverulega staðhæfingin uppfyllanleg, eða þessi sérstaka erfiða formúla er uppfyllanleg.

Sérstaka formúlan er ekki uppfyllanleg. Ef undirliggjandi sönnunarkerfi er fullkomlega hljóð þýðir samþykkt skeytisins því enn að raunverulega staðhæfingin er sönn. Þannig fæst fullkominn hljóðleiki.

Fyrir núllþekkingarlíka öryggið ímyndum við okkur hins vegar að erfiða formúlan **væri** uppfyllanleg. Þá mætti nota vitni hennar til að herma sönnunargögn án þess að vita raunverulega vitnið. Í veruleikanum er hún óuppfyllanleg — en reglubókin getur ekki á skilvirkan hátt sannað það. Hún getur því ekki á skilvirkan hátt sannað að hermirinn sé ómögulegur.

Þarna er lömin. Kerfið felur ekki leyndarmálið með því að hafa klassískan hermi. Fyrir stóran flokk mælanlegra öryggisprófa felur það leyndarmálið **bak við vanmátt reglubókarinnar til að votta fjarveru hermis**.

Hvað heldur greinin fram?

Meginsetningin kemur í nokkrum lögum.

Undir staðlaðri dulmálsforsendu — tilvist **non-interactive witness indistinguishable proofs**, vel rannsakaðra hluta sem fylgja úr nokkrum rótgrónum forsendupökkum — og undir sönnunarflækjutilgátunni að **ekkert (infinitely often) bestmögulegt sönnunarkerfi sé til**, smíðar greinin fyrir **hverja fasta reglubók** eins-skeytis sannanda og sannprófara fyrir NP/SAT með **fullkomnum hljóðleika** og engu upphafsfyrirkomulagi sem er virkt núllþekkingarlegt miðað við reglubókina. NP/SAT er hinn staðlaði „harðasti sameiginlegi nefnari“ þrautavandamála; MegaSudoku er bara einn búningur þess.

Fyrir víðari fullyrðingu um að varðveita prófanlega öryggiseiginleika bætir greinin við einni staðlaðri forsendu til viðbótar: af-tilviljunartilgátunni **P = BPP**, í grófum dráttum að tilviljun gefi reikniritum ekki grundvallaraukakraft.

Þýtt úr setningamáli:

- Sönnunin er eitt skeyti.
- Það er ekkert traust upphafsfyrirkomulag.
- Rangar staðhæfingar geta ekki haft gilda sönnun.
- Sannandinn er **ekki** klassísk núllþekking — enginn raunverulegur hermir er til.
- En hver **afsannanlegur, leikjamiðaður öryggisafleiðing** klassískrar núllþekkingar má ná í þessu umhverfi.

Orðið **afsannanlegur** skiptir máli. Það merkir að hægt sé að prófa öryggisbrest með því að keyra andstæðing í leik. Mörg dulmálsöryggisskilgreiningar eru þannig: getur andstæðingur greint tvær dulkóðanir í sundur, snúið við falli, fundið vitnið eða unnið ákveðna tilraun? Setningin gefur sannanda fyrir **hvern slíkan eiginleika fyrir sig**.

Einn sannandi sem hefði **alla** afsannanlega eiginleika samtímis er líklega ómögulegur. Gamla endurnýtingarárásin — „Bob getur sýnt sönnunina öðrum“ — er sjálf afsannanlegur eiginleiki og hann bregst raunverulega hér. Greinin leggur til að einn sannandi gæti mögulega uppfyllt alla **náttúrulega** afsannanlega eiginleika, þ.e. þá sem koma raunverulega fyrir í dulmálsfræði, en sá hluti er skilyrt niðurstaða sem byggir bæði á óformlegu hugtakinu „náttúrulegt“ og sérstakri tilgátu. Tryggingin beinist því að mælanlegum öryggisbrestum, ekki að sérhverri heimspekilegri eða hermimiðaðri merkingu leyndar.

Ein afleiðing er sérstaklega auðskilin: smíðin gefur fyrstu ógagnvirku **witness-hiding** sannanirnar með uniform sannanda — „sönnun þess að þraut hafi lausn hjálpar þér ekki að finna lausnina“ — án gagnvirkni og án upphafsfyrirkomulags. Það hljómar hóflega en hafði staðist smíði í áratugi.

Hvað segir greinin ekki?

Þessi hluti heldur sögunni heiðarlegri.

Hún segir **ekki** að gömlu ómöguleikasetningarnar hafi verið rangar. Smíðin fer fram hjá þeim með því að breyta skilgreiningunni.

Hún gefur **ekki** venjulega, klassíska núllþekkingu með einu skeyti, engu upphafsfyrirkomulagi og fullkomnum hljóðleika. Greinin segir beinlínis að smíðaði sannandinn hafi engan raunverulegan hermi.

Hún þýðir **ekki** að sönnunin sé ekki endurnýtanleg. Eins-skeytis sönnun má enn sýna öðrum; greinin varðveitir ekki deniability-eiginleika eða óendurnýtanleika. Sama takmörkun gildir raunar fyrir mörg ógagnvirk núllþekkingarkerfi með traustu setup-i.

Hún þýðir **ekki** að þetta sé hagnýt samskiptaregla tilbúin til notkunar. Þetta er flækjufræði og grunnkenningar dulmáls. Niðurstaðan byggir á stórum forsendum í sönnunarflækju og dulmálsfræði og snýst fyrst og fremst um hvað sé mögulegt í grundvallaratriðum.

Og hún gerir **ekki Gödel að töfraöryggisfrumstæðu**. Gödel-tengingin liggur í sönnunarkerfum, bestmögulegum sönnunarkerfum og endanlegum hliðstæðum ófullkomleika. Gagnlega innsýnin er ekki „ófullkomleiki verndar lykilorðið þitt“. Hún er: ef reglubók getur ekki á skilvirkan hátt sannað að hermir sé ómögulegur, þá má hindra ákveðna flokka árása á stigi öryggisskilgreininga sem annars myndu krefjast slíkrar sönnunar.

Af hverju er þetta samt áhugavert?

Dulmálsfræði breytir oft **erfiðleika í öryggi**. Þáttun stórra talna er erfið, svo RSA-líkar forsendur verða gagnlegar. Grindarvandamál eru erfið, svo grindardulritun verður gagnleg. Hér er erfiðleikinn undarlegri: ekki „erfitt að reikna leyndarmálið“, heldur **„erfitt að sanna að ákveðinn sönnunarhlutur geti ekki verið til.“**

Það er það sem gerir greinina óvenjulega. Hún meðhöndlar frumsendur og reglubækur næstum eins og dulmálsauðlindir. Venjulegi ómöguleikinn segir að spennan sé milli hljóðleika og hermunar. Bragð Ilango er að færa spennuna bak við sönnunarfræðilegt tjald: hermirinn er ekki til, en formlega kerfið getur ekki á skilvirkan hátt afhjúpað fjarveru hans.

Það óvænta fyrir lesandann er ekki að þetta muni leysa af hólmi núverandi núllþekkingarkerfi. Það mun líklega ekki gera það beint. Hitt er óvæntara: takmörkun úr stærðfræðilegri rökfræði má nota **uppbyggilega** — ekki bara sem vegg heldur sem eins konar skjól.

Hversu sterk eru gögnin?

Þetta er **setningagrein**, þannig að „sönnunargögn“ merkja annað en í líffræði eða stjörnufræði. Spurningin er ekki hvort tilraun hafi endurtekið sig heldur hvort skilgreiningar, forsendur og sönnunarkeðja styðji fullyrðinguna.

Sönnunin er formleg og greinin er skýr um forsendur. Þær eru ekki léttvægar. Non-interactive witness indistinguishable proofs eru staðlaðir hlutir í dulmálsfræði og fylgja úr nokkrum rótgrónum forsendupökkum. Tilgátan um að ekkert bestmögulegt sönnunarkerfi sé til er miðlæg opin tilgáta í sönnunarflækju. **P = BPP** er staðlað derandomization-viðhorf og er aðeins notað fyrir víðari setninguna um afsannanlega eiginleika.

Greinin færir líka rök fyrir því að forsendurnar séu **rétt verð**, ekki handahófskennd stoðgrind. Hún sýnir converse-niðurstöðu: ef slíkar smíðar eru til yfir höfuð verða non-interactive witness indistinguishable proofs að vera til og — að teknu tilliti til staðlaðra einátta falla — getur ekkert bestmögulegt sönnunarkerfi verið til. Forsendurnar eru líka „win-win“: að hrekja einhverja þeirra væri sjálft stórmerkileg niðurstaða í sönnunarflækju, dulmálsfræði eða flækjufræði.

En þar sem niðurstaðan er **skilyrt** er traustið líka skilyrt. Ef forsendurnar reynast rangar breytist túlkun setningarinnar. Og jafnvel ef þær halda er tryggingin ekki full klassísk núllþekking; hún er slakari, sönnunarfræðileg útgáfa greinarinnar.

Rétt traust er því hátt á að greinin sýni samhangandi skilyrta möguleikaniðurstöðu, miðlungs á að forsendurnar lýsi dulmálsheiminum sem við lifum í og lítið á beina hagnýta notkun á næstunni.

Af hverju skiptir þetta máli?

Greinin opnar leið sem átti að vera lokuð.

Klassísk kenning segir: full núllþekking getur ekki verið eitt skeyti án upphafsfyrirkomulags og getur ekki haft fullkominn hljóðleika. Ilango segir: ef við spyrjum í staðinn um **afleiðingar núllþekkingar sem hægt er að prófa í öryggisleikjum**, og leyfum öryggisskilgreiningunni að ráðast af því hvað reglubók getur eða getur ekki á skilvirkan hátt afsannað, þá má endurheimta stóran hluta gagnlegu hegðunarinnar — með einu skeyti, engu setup-i og fullkomnum hljóðleika.

Þetta er ekki lítil orðabreyting. Þetta er önnur leið til að hugsa um dulmálsábyrgðir. Í stað þess að spyrja aðeins **hvað er til**, spyrðu **hvað reglubókin þín getur útilokað**. Í stað þess að líta á ósannanleika sem heimspekilegt ónæði má nota hann sem byggingarhluta.

Hagnýti heimurinn breytist kannski ekki á morgun. En hugtakakortið breytist. Það er nú formlegur skilningur þar sem „enginn getur á skilvirkan hátt sannað að leyndarmálið hafi lekið“ getur verið nógu sterkt til að endurheimta marga af leikjamiðuðu öryggiseiginleikunum sem við vildum fá úr „leyndarmálið lak ekki“.

Þess vegna á Gödel heima í titlinum.

Í stuttu máli

Núllþekkingarsannanir gera sannanda kleift að sannfæra sannprófara um að staðhæfing sé sönn án þess að sýna vitnið. Klassískar ómöguleikaniðurstöður segja að núllþekking verði hvorki eitt skeyti án setup-s né fullkomlega hljóð. Grein Rahul Ilango hrekur þetta ekki. Hún skilgreinir veikari hugmynd, **virka núllþekkingu**: í stað þess að krefjast þess að raunverulegur hermir sé til er krafist að valið formlegt sönnunarkerfi — reglubók eins og ZFC — geti ekki á skilvirkan hátt sannað að enginn hermir sé til. Undir stórum forsendum í dulmálsfræði (non-interactive witness indistinguishable proofs) og sönnunarflækju (ekkert bestmögulegt sönnunarkerfi) smíðar greinin eins-skeytis sannanir fyrir NP/SAT, án setup-s og með fullkomnum hljóðleika, sem ná afsannanlegum leikjamiðuðum afleiðingum klassískrar núllþekkingar, eiginleika fyrir eiginleika. Einn sannandi sem nær öllum „náttúrulegum“ slíkum eiginleikum er frekari og að hluta til getgátukennd útvíkkun — og að ná bókstaflega öllum afsannanlegum eiginleikum er líklega ómögulegt vegna þess að sannanir má endurnýta. Niðurstaðan er fræðileg og skilyrt, ekki tilbúið dulmálsverkfæri, en hún sýnir nýja leið til að nota ósannanleika innan sönnunarkerfa sem dulmálsauðlind.

Raunhæf yfirferð

Hvað sýnir greinin? Undir tilgreindum forsendum er hægt að byggja eins-skeytis, setup-lausa og fullkomlega hljóða sannendur fyrir NP/SAT sem eru virkt núllþekkingarlegir miðað við hvaða fasta reglubók sem er og geta náð hverri afsannanlegri leikjamiðaðri afleiðingu klassískrar núllþekkingar.

Hvað er trúverðugt en ekki sannað skilyrðislaust? Að nauðsynlegar forsendur í sönnunarflækju og dulmálsfræði haldi. Þetta eru alvarlegar, vel rannsakaðar forsendur — og greinin sýnir að þær eru að verulegu leyti nauðsynlegar sem og nægjanlegar — en þær eru samt forsendur.

Hvað sýnir greinin ekki? Klassíska núllþekkingu með engu samspili, engu setup-i og fullkomnum hljóðleika; hagnýtt kerfi tilbúið til notkunar; deniability eða óendurnýtanleika sannana; eða að ófullkomleikasetning Gödel ein og sér geri dulmál öruggt.

Helstu takmarkanir: Tryggingin er slökun á núllþekkingu; víðasta útgáfan byggir á mörgum forsendum; fullyrðingar um einn universal sannanda eru að hluta getgátukenndar; og niðurstaðan er fyrst og fremst grunnkenning.

Hversu mikið traust ætti almennur lesandi að leggja á niðurstöðuna? Mikið traust á að þetta sé mikilvæg skilyrt kenningarniðurstaða ef skilgreiningarnar eru samþykktar. Miðlungs traust á að forsendurnar lýsi raunverulegum dulmálsheimi. Lítið traust á beina notkun strax. Örugga takeaway-ið er: greinin brýtur ekki klassísku ómöguleikaniðurstöðurnar; hún finnur nýja sönnunarfræðilega leið fram hjá þeim hlutum sem skipta máli fyrir marga öryggisleiki.

Heimildir

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