Hans van Ditmarsch
LORIA, Nancy, France
IMSc, Chennai, India
hans.van-ditmarsch@loria.fr

Barteld Kooi
Faculteit Wijsbegeerte
Rijksuniversiteit Groningen
b.p.kooi@rug.nl

Vakantiecursus

Honderd gevangenen en een gloeilamp

Op de vakantiecursus 2014 van het Platform Wiskunde Nederland bespreken Hans van Ditmarsch en Barteld Kooi het raadsel 'Honderd gevangenen en een gloeilamp'. Dit raadsel doet sinds het begin van deze eeuw de ronde in het wiskundepuzzelcircuit. Dit verhaal is een bewerking van het gelijknamige hoofdstuk uit het boek *Honderd gevangenen en een gloeilamp*, **deel 73 van Epsilon Uitgaven, waarin zij een aantal wiskunderaadsels beschrijven.**

Een groep van honderd gevangenen, gezamenlijk in de gevangeniskantine, wordt medegedeeld dat ze allemaal in isolatiecellen geplaatst zullen worden en daarna één voor één ondervraagd, in een kamer met een lamp met een aan-uitschakelaar. De gevangenen kunnen met elkaar communiceren door de lamp aan of uit te doen (en dat is de enige manier waarop ze kunnen communiceren). De lamp is aan het begin uit. Er is geen vaste volgorde van ondervraging, er is geen standaard tijdsduur tussen de ondervragingen, en dezelfde gevangene kan best meerdere keren achter elkaar ondervraagd worden. We mogen wel aannemen dat op ieder moment iedere gevangene ooit nog eens ondervraagd zal worden. Als een gevangene ondervraagd wordt, kan deze: niets doen, de lamp uitdoen, de lamp aandoen, of verklaren dat iedereen (ten minste één keer) ondervraagd is. Als dit waar is, worden alle gevangenen vrijgelaten. Anders worden ze allemaal opgehangen. Kunnen de gevangenen, zolang ze nog bij elkaar zijn in de kantine en niet naar de isolatiecellen gebracht zijn, een protocol overeenkomen waardoor ze vrijgelaten worden?

Tot honderd tellen met één bit?

Het raadsel lijkt onoplosbaar. Er is immers maar één bit beschikbaar voor de informatieoverdracht: de lamp kan aan, of uit. Maar er zijn honderd gevangenen. Het getal 100 ligt tussen 64 en 128 en vergt dus 7 bits om het te representeren. En dan hebben we het nog niet over het protocol om het raadsel op te lossen. Hoe kan één bit nu voldoende zijn om dat allemaal te doen?

Een struikelblok *lijkt* de wiskundige intuïtie te zijn. De natuurlijke inductie waarmee we vaak problemen te lijf gaan, suggereert een methode om de oplossing te zoeken, die helaas in dit geval niet ergens toe leidt. Laten we het even hardop doen. Oplossen voor één gevangene, oplossen voor twee gevangenen, oplossen voor meer gevangenen.

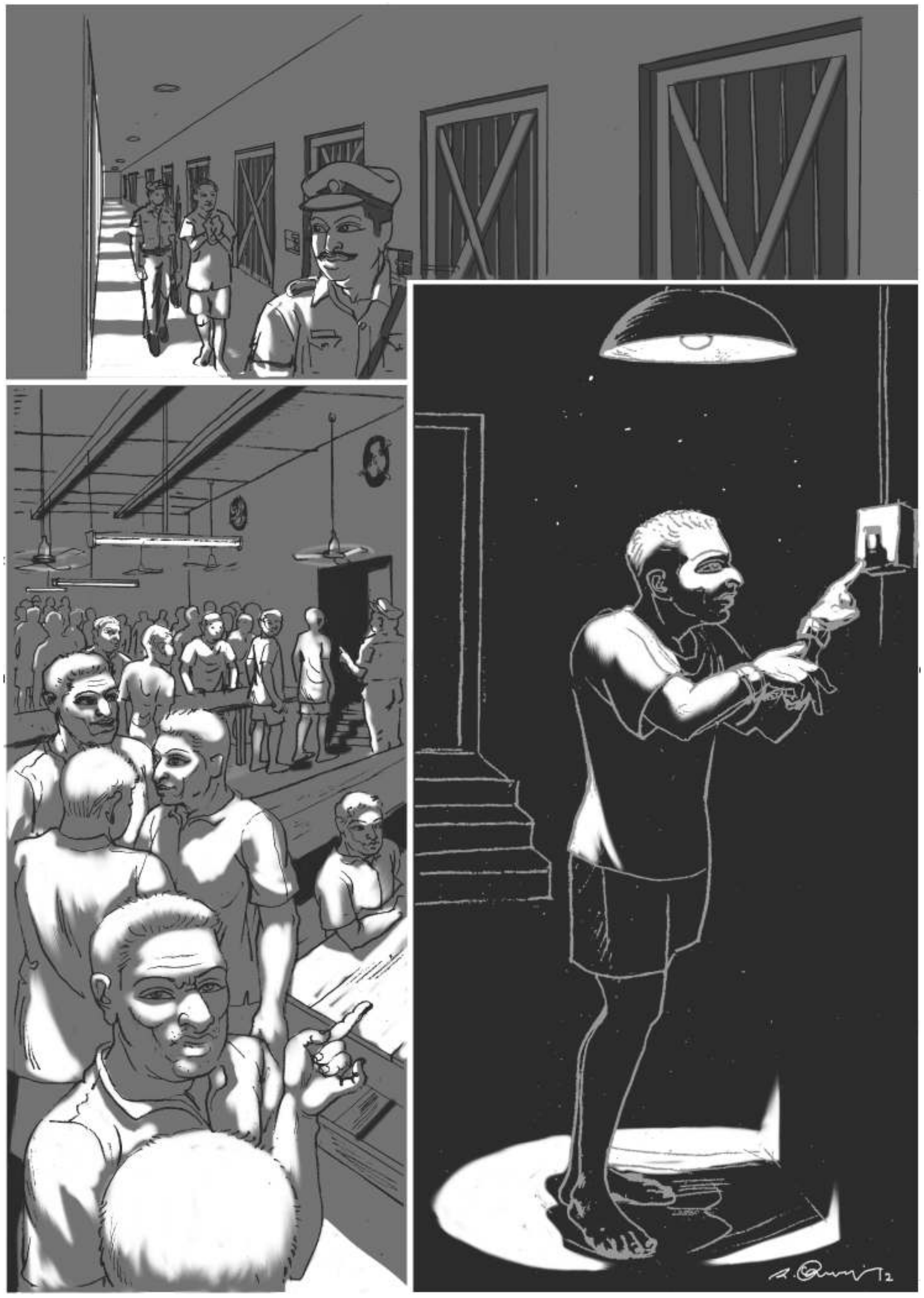
Eén gevangene

We beginnen met de basis. Stel er is één gevangene: Anna. De eerste keer dat Anna ondervraagd wordt, weet zij dat iedereen ondervraagd is. Daarvoor is de lamp niet eens nodig. Dus:

Protocol 1. *Als je ondervraagd wordt, is iedereen ondervraagd.*

Twee gevangenen

Protocol 1 werkt natuurlijk niet als er meer dan een gevangene is. Maar misschien kunnen we het aanpassen. Stel er zijn twee gevangenen: Anna en Bert. De eerste van de twee die ondervraagd wordt, doet de lamp aan. Stel, zonder verlies aan algemeenheid van de oplossing, dat dit wederom Anna is. Nu is het even afwachten wie er daarna ondervraagd wordt. Als het Bert is, dan ziet Bert dat de lamp aan is, en weet daarom dat Anna reeds ondervraagd is (alleen gevangenen mogen de lamp aan en uit doen). Bert kan dan dus naar waarheid zeggen dat iedereen ondervraagd is. En Anna en Bert gaan vrijuit. Als daarentegen de volgende ondervraagde opnieuw Anna is, is de lamp nog aan, en 'denkt ze dus' dat Bert nog niet ondervraagd is, anders was ze al vrijgelaten. We zijn nu wat vaag: Anna's argument is geba-



Illustratie: Elancheziyan, Chennai

Hier zien we nu zo'n gevangene de schakelaar omdraaien

seerd op wat ze in alle redelijkheid vermoedt dat Bert gedaan heeft. Dit kan in dit geval wel zonder afspraak, maar aan de andere kant is er niets op tegen dat Anna en Bert dit met elkaar afspreken: hier komt het protocol om de hoek kijken. Dit kunnen we als volgt verwoorden:

Protocol 2. *Als je ondervraagd wordt en de lamp is uit, doe hem dan aan; als je ondervraagd wordt, en jij hebt de lamp eerder aangedaan en de lamp is nog steeds aan, doe dan niets; als je ondervraagd wordt, en de lamp is aan maar jij hebt hem niet aangedaan, verklaar dan dat iedereen ondervraagd is.*

Een protocol voor drie gevangenen?

Stel er zijn drie gevangenen, Anna, Bert en Caroline... Nu wordt het moeilijker. Anna, die weer als eerste ondervraagd wordt, kan de lamp aandoen. Als daarna bijvoorbeeld Bert ondervraagd wordt, kan Bert de lamp weer uitdoen. Stel dat daarna Anna weer ondervraagd wordt, dan weet Anna dat ten minste twee gevangenen nu ondervraagd zijn. Dat schiet op! Als daarna Caroline ondervraagd wordt, weet Caroline echter niet of een andere gevangene al eerder ondervraagd is: het kan dan ook best zo zijn dat zij als eerste ondervraagd wordt (de tijdsduur tussen ondervragingen is niet bekend)! Wat moet Caroline nu doen? Dat is eigenlijk dezelfde vraag als wat Anna de eerste keer moet doen. Maar laten we nog even doorfantaseren — want het is namelijk fantaseren, we komen er zo echt niet uit: het zou handig zijn als Anna eerst de lamp aan kan doen, daarna zoals hiervoor de lamp weer uit kan zien, en *daarna* de lamp nog een keer aandoet. De volgende keer als ze de lamp weer uit ziet, zou dan dus iedereen ondervraagd moeten zijn? Laten we het formaliseren: het protocol wordt nu:

Protocol 3. *De eerste keer dat de lamp uit is als je ondervraagd wordt, doe hem dan aan; als de lamp aan is als je ondervraagd wordt en je weet dat jij hem niet hebt aangedaan, doe hem dan uit. In de andere gevallen, doe niets. Als je de lamp de tweede keer uitdoet, verklaar dat alle drie gevangenen ondervraagd zijn.*

Dit protocol werkt soms wel en soms niet. Het gaat wel goed bij deze uitvoering:

$$\text{Anna} - \text{Bert} - \text{Bert} - \text{Anna} - \text{Caroline} - \text{Anna}.$$

Stel dat 0 betekent dat de lamp uit is en 1 dat de lamp aan is. We kunnen deze uitvoering dan annoteren met een bovenindex voor de toestand van de lamp, met als resultaat

$${}^0\text{Anna}^1\text{Bert}^0\text{Bert}^1\text{Anna}^0\text{Caroline}^1\text{Anna}^0$$

en als we dan eveneens voor iedere gevangene bijhouden hoe vaak deze de lamp uitdoet, met een onderindex, dan krijgen we

$${}^0\text{Anna}^1_0\text{Bert}^0_1\text{Bert}^1_1\text{Anna}^0_1\text{Caroline}^1_0\text{Anna}^0_2.$$

Anna verklaart dus terecht dat iedereen is ondervraagd.

Maar kijk nu eens naar deze uitvoering:

$$\text{Anna} - \text{Bert} - \text{Bert} - \text{Caroline} - \text{Caroline} - \text{Anna} -$$

Met dezelfde annotaties krijgen we nu

$${}^0\text{Anna}^1_0\text{Bert}^0_1\text{Bert}^1_1\text{Caroline}^0_1$$
$$\text{Caroline}^1_1\text{Anna}^0_1\ldots$$

Alle gevangenen hebben één keer de lamp uitgedaan. En ze hebben de lamp ook allemaal één keer aangedaan. En vaker zullen ze de lamp niet aandoen. Dus wie er nu verder ook ondervraagd wordt, en hoe lang al die ondervragingen nu ook verder door blijven gaan, de lamp blijft altijd uit. Er gebeurt niets meer. Het protocol termineert niet in dit geval.

Andere variaties op dit thema leiden ook tot ellende. Het probleem van dit protocol is dat het soms wel en soms niet werkt. Het is natuurlijk wel heel onwaarschijnlijk als we ervan uitgaan dat de ondervragers steeds een willekeurige gevangene uitzoeken voor ondervraging, en aangenomen dat we iets weten over de tijdsduur van en tussen ondervragingen, dat dan na 100 keer nog niet iedereen ondervraagd is. De gevangenen kunnen dan best een geïnformeerd gokje wagen dat iedereen ondervraagd is. Maar om kennis te verkrijgen dat iedereen ondervraagd is, is meer nodig.

Geen trucjes

Allerlei trucjes waar de lezer wellicht aan gedacht heeft, zijn niet toegestaan, en het raadsel is ook geen strikvraag in welke zin dan ook: voelen of de lamp warm of koud is, is niet toegestaan (als de lamp uit is maar warm, weet je dat er iemand voor jou ondervraagd is — maar wat schiet je ermee op?); je mag de lamp ook niet stukslaan als je deze voor de tweede keer aan ziet (waarna de gevangene die een stukgeslagen lamp ziet maar nog niet zelf een lamp heeft uitgedaan, kan verklaren dat iedereen ondervraagd is — dit geeft een oplossing voor drie gevangenen). En het interval tussen de ondervragingen is variabel: dus de tijd bijhouden heeft geen zin. De ondervragingskamer kan niet worden gezien vanuit de isolatiecellen (daarom zijn het ook *isolatiecellen*), en er is geen geheime klikverbinding tussen de schakelaar van de lamp en de isolatiecellen; je kunt er ook niet de schakelaar horen overgaan... De gevangenen mogen allemaal verschillende namen hebben, of met opeenvolgende getallen van 1 tot 100 geïdentificeerd worden, maar dat is allemaal lood om oud ijzer.

Oplossing voor honderd gevangenen

De oplossing komt snel naderbij als we ons realiseren dat een protocol niet aan iedere gevangene dezelfde *rol* hoeft toe te bedelen: aangezien de gevangenen het protocol kunnen afspreken zolang ze nog in de kantine zijn, kunnen ze elkaar verschillende taken geven. De telling die Anna bijhoudt in het voorbeeld hiervoor, werkt, zolang iedereen maar weet dat *alleen* Anna dit doet. Dit brengt ons tot het volgende protocol.

Protocol 4. *De gevangenen wijzen onder elkaar een teller aan. Voor de teller geldt: Als je ondervraagd wordt en de lamp is uit doe dan niets; als je ondervraagd wordt en de lamp is aan, doe hem dan uit — houdt bij hoe vaak je dat doet; als je de lamp* 99 *keer hebt uitgedaan, verklaar dan dat alle* 100 *gevangenen ondervraagd zijn. Voor de overige gevangenen geldt: Als je ondervraagd wordt en de lamp is uit en je hebt hem nog niet eerder aangedaan, doe hem dan aan; als je ondervraagd wordt en de lamp is uit en je hebt hem al eerder aangedaan, doe dan niets; als je ondervraagd wordt en de lamp is aan, doe dan niets.*

Het zal de lezer duidelijk zijn dat dit protocol inderdaad werkt. Voor drie gevangenen, waarbij Anna als teller is aangewezen, zijn de volgende drie uitvoeringen van het protocol succesvol. We gebruiken weer dezelfde annotaties waarbij de bovenindex staat voor de toestand (aan/uit) van de lamp en de onderindex voor de telling die in dit geval alleen Anna bijhoudt.

1. ${}^0\text{Bert}\,{}^1\text{Anna}^0_1\text{Caroline}\,{}^1\text{Anna}^0_2$;

2. ${}^0\text{Anna}\,{}^0\text{Bert}\,{}^1\text{Caroline}\,{}^1\text{Anna}^0_1\text{Bert}\,{}^0\text{Anna}^0_1$ $\text{Caroline}\,{}^1\text{Caroline}\,{}^1\text{Bert}\,{}^1\text{Bert}\,{}^1\text{Anna}^0_2$;

3. ${}^0\text{Bert}\,{}^1\text{Anna}^0_1\text{Bert}\,{}^0\text{Caroline}\,{}^1\text{Bert}\,{}^1\text{Anna}^0_2$.

Stel dat Anna de teller is en dat van de 100 gevangenen alleen om de beurt Anna en Bert ondervraagd worden. Dan zal het nooit gebeuren dat Anna kan zeggen dat iedereen ondervraagd is. Maar dan werkt het protocol toch niet? Wel, in dat geval komt de oplossing inderdaad nooit dichterbij. Maar dat is dan omdat nooit alle gevangenen ondervraagd worden! Een voorwaarde voor de terminatie van Protocol 4 is de zogenaamde 'fairness' (eerlijkheid) of 'liveness' van het ondervragingsrooster: we hebben dit in het raadsel verwoord als de eis: "We mogen wel aannemen dat op ieder moment iedere gevangene ooit nog eens ondervraagd zal worden." Onder die aanname zal de teller ooit naar waarheid kunnen beweren dat iedereen ondervraagd is.

We kunnen ons ook nog voorstellen dat bij de eerste honderd ondervragingen alle gevangenen na elkaar ondervraagd worden, en daarna tot in de eeuwigheid alleen Anna en Bert om de beurt. Dan is wel iedereen ondervraagd, maar komt de oplossing toch nooit naderbij. Maar ook dan wordt niet voldaan aan "We mogen wel aannemen dat op ieder moment iedere gevangene ooit nog eens ondervraagd zal worden": er is immers een moment waarna alleen nog Anna en Bert ondervraagd worden. Uit de eis van eerlijk roosteren volgt dat alle gevangenen oneindig vaak ondervraagd zullen worden: Neem een zeker moment. Daarna wordt Anna nog eens ondervraagd. Neem dat moment. Daarna wordt Anna nog eens ondervraagd. Neem dat moment. Enzovoort.

Lamp aan het begin aan of uit?

Stel dat niet bekend is of het licht in eerste instantie aan of uit is. En, als voorheen, niets bekend is over de tijdsduur tussen de ondervragingen. Wat is dan een protocol om het probleem op te lossen?

Protocol 4 werkt nu niet, en ook niet als we de teller één verder door laten tellen. Bijvoorbeeld, voor drie gevangenen, kunnen we de volgende uitvoeringen niet van elkaar onderscheiden:

1. (licht aan)
${}^1\text{Anna}^0_1\text{Caroline}\,{}^1\text{Anna}^0_2\text{Bert}\,{}^1\text{Anna}^0_3$;

2. (licht uit)
${}^0\text{Bert}\,{}^1\text{Anna}^0_1\text{Caroline}\,{}^1\text{Anna}^0_2\text{Bert}\,{}^0\text{Anna}^0_2$.

Als Anna slechts tot twee telt, verklaart ze bij uitvoering 1 na de tweede keer ondervraagd te zijn ten onrechte dat iedereen ondervraagd is. Als we Anna daarentegen tot drie laten tellen, dan komt ze bij uitvoering 2 nooit aan die verklaring toe, hoe vaak ook Anna, Bert en Caroline ondervraagd zullen blijven worden in de toekomst: zowel Bert als Caroline hebben het licht precies één keer aangedaan, en Anna doet daarna dat licht weer uit, en daar blijft het bij. We kunnen Protocol 4 echter aanpassen — de truc is dat we moeten compenseren voor de onzekerheid dat Anna één keer teveel het licht moet uitdoen, omdat het initieel aan had kunnen zijn.

Protocol 5. *De gevangenen wijzen onder elkaar een teller aan. Voor de teller geldt: Als je ondervraagd wordt en de lamp is uit, doe dan niets; als je ondervraagd wordt en de lamp is aan, doe hem dan uit — houd bij hoe vaak je dat doet; als je de lamp* 198 *keer hebt uitgedaan, verklaar dan dat alle n gevangenen ondervraagd zijn. Voor de overige gevangenen geldt: Als je ondervraagd wordt en de lamp is uit en je hebt hem nog niet twee keer aangedaan, doe hem dan aan; als je ondervraagd wordt en de lamp is uit en je hebt hem al twee keer eerder aangedaan, doe dan niets; als je* *ondervraagd wordt en de lamp is aan, doe dan niets.*

Voor n gevangenen is het $2n-2$. We leggen uit waarom dit werkt voor het geval van drie gevangenen. Er moet dan vier keer ondervraagd worden. In het ergste geval was het licht aan het begin aan, 1; heeft een van Bert of Caroline het licht al twee keer aangedaan, stel dit is Bert, 2; en heeft Caroline het licht pas één keer aangedaan, 1. Anna hoeft dus niet te wachten totdat Caroline het licht twee keer aangedaan heeft. Maar drie keer is te weinig: want dat krijgen we al als het licht initieel aan is en Bert het twee keer heeft aangedaan, zonder dat Caroline ooit ondervraagd is.

In het algemeen hoeft de teller niet te wachten totdat de laatste niet-teller die het licht nog niet twee keer heeft aangedaan, dat voor de tweede keer heeft gedaan.

Soms weet je het, voordat de teller het weet

We kijken nog eens naar de drie verschillende uitvoeringen van Protocol 4, voor het geval van drie gevangenen. In uitvoering 1 is iedereen ondervraagd na drie ondervragingen, en weet Anna dat bij de vierde ondervraging. In uitvoering 2 is iedereen eveneens ondervraagd na drie ondervragingen, maar duurt het nog acht ondervragingen voordat Anna, als teller, dat ook weet en bekend kan maken. Kan dit sneller? In het eerste geval niet, maar in de andere gevallen wel. We geven de uitvoeringen nog eens weer, waarbij de vetgedrukte gevangene naar waarheid kan verklaren dat iedereen ondervraagd is:

1. ^{0}Bert^{1}Anna${}^{0}_{1}$Caroline^{1}**Anna**${}^{0}_{2}$

2. ^{0}Anna^{0}Bert^{1}Caroline^{1}Anna${}^{0}_{1}$Bert^{0}Anna${}^{0}_{1}$ **Caroline**^{1}Caroline^{1}Bert^{1}Bert^{1}Anna${}^{0}_{2}$

3. ^{0}Bert^{1}Anna${}^{0}_{1}$Bert^{0}Caroline^{1}**Bob**^{1}Anna${}^{0}_{2}$

In uitvoering 2 weet Caroline bij de tweede keer dat ze ondervraagd wordt, dat iedereen ondervraagd is. De eerste keer dat ze wordt ondervraagd ziet ze dat de lamp aan is: dat moet dus door Bert gedaan zijn. De tweede keer dat ze wordt ondervraagd ziet ze dat de lamp uit is: dat moet dus door Anna gedaan zijn. Dus iedereen is ondervraagd.

In uitvoering 3 weet Bert vóór Anna dat iedereen ondervraagd is, ook al heeft Bert niet de rol van teller. De eerste keer dat hij ondervraagd wordt doet hij de lamp aan. De tweede keer dat hij ondervraagd wordt is de lamp uit en doet hij niets, maar hij weet daarom dat Anna nu ook ondervraagd is en de lamp weer heeft uitgedaan. De derde keer dat hij ondervraagd wordt ziet hij dat de lamp weer aan is: dat kan dus alleen door Caroline gedaan zijn. En hij kan dan naar waarheid verklaren dat iedereen ondervraagd is, *voordat* teller Anna dat weet en dat kan verklaren.

Een probabilistisch protocol

In de protocols die we tot nog toe gezien hebben vervullen de gevangenen verschillende rollen: teller, of niet-teller. Ook de variaties daarop (en de variaties in het vervolg van dit verhaal) vallen in dit stramien. Het is daarom verrassend dat er toch een protocol is waarin alle gevangenen dezelfde rol hebben. We kunnen niet zeggen: “waarin alle gevangenen hetzelfde *doen*”, want wat ze doen hangt ook af van het toeval. Voor sommige acties bepaalt een dobbelsteen (of een munt) wat ze gaan doen. Maar dit toeval heeft dan wel bij iedere gevangene dezelfde functie. Om dit protocol te beschrijven introduceren we de notie van *pot* (zoals in jampot of spaarpot, in het Engels, ‘token’) en van *knikker*.

Iedere gevangene heeft een pot waar een aantal knikkers in zit, om te beginnen één knikker. Verder stellen we ons de lamp voor als een klein hooggerand schaaltje waar precies één knikker op past (net zoals een ‘tee’ voor golfballen). Als het schaaltje leeg is, is de lamp uit, als het schaaltje vol is, is de lamp aan. Als de gevangene het licht aandoet wanneer het licht uit is, vertalen we dat nu in: hij legt een knikker op de schaal. Het licht aan laten als het al aan is, betekent: de gevangene kan geen knikker op de schaal leggen (want de schaal is al vol). Tot nog toe waren deze handelingen voorbehouden aan de niet-tellers. We noemen dit nu het gedrag van een *legger*. Het licht uitdoen als het aan is, betekent: knikker van de schaal pakken en in je pot doen. Het licht uitlaten als het al uit was, betekent: geen knikker van de schaal kunnen pakken. Tot nog toe waren deze handelingen aan de teller voorbehouden. We noemen dit het gedrag van een *pakker*. Iedere gevangene kan zowel legger als pakker zijn. In deze nieuwe opzet termineert het protocol als een gevangene n knikkers in zijn pot heeft. We vervolgen nu met de beschrijving van het protocol, waarbij, uiteraard, $[0,1]$ staat voor het interval van de reële getallen tussen 0 en 1 en $Pr(x)$ voor de waarschijnlijkheid om legger of pakker te zijn als de som van het aantal knikkers in je pot en op de schaal x is.

Protocol 6. *Als je ondervraagd wordt, laat m de som zijn van het aantal knikkers in je pot en op de schaal. Gegeven is een functie $Pr:\{0,\dots,n\}\to[0,1]$ met $Pr(0)=Pr(1)=1$, $0<Pr(x)<1$ voor $x\neq 0,1,n$, en $Pr(n)=0$. Je bent legger met waarschijnlijkheid $Pr(m)$, en anders ben je pakker. Verklaar dat iedereen ondervraagd is als je n knikkers hebt in je pot.*

Een knikker leggen als je er geen hebt, betekent niets doen. Daarom is $Pr(0)=1$. Dit protocol termineert altijd (dat wil zeggen, de verwachtingswaarde is strikt positief), maar je kunt je kansen vergroten door de waarschijnlijkheid Pr dalend te laten zijn op het interval tussen 0 en (het kleinste aantal groter dan of gelijk aan) de helft van het aantal knikkers, en 0 te laten zijn als de gevangene meer dan de helft van de knikkers heeft (in feite is dit dus een *variant* op Protocol 6, want we eisten dat $Pr(m)$ strikt positief was behalve voor $m=n$). Met andere woorden: de eerste gevangene die meer dan de helft van de knikkers weet te vergaren, is alleen nog een pakker vanaf dat moment (waarschijnlijkheid 0 om legger te zijn) en nooit meer legger. We illustreren deze variant van dit protocol aan de hand van een voorbeeld.

Er zijn vier gevangenen Anna, Bert, Caroline en Dick. Ze gebruiken Protocol 6 met $Pr(0)=Pr(1)=1$, $Pr(2)=0{,}5$, $Pr(3)=0$, en $Pr(4)=0$, en de volgende uitvoering resulteert. We laten nu zien dat deze uitvoering termineert en dat Bert aan het eind naar waarheid kan zeggen dat iedereen ondervraagd is:

Anna, Bert, Caroline, Dick, Bert, Caroline, Caroline, Bert, Caroline, Bert.

We gebruiken nogmaals annotatie om de toestandsovergangen te verduidelijken. In dit geval staat de onderindex voor de inhoud van de pot *na* het uitvoeren van de handeling. (Dus dit is als de onderindex in het voorgaande, waarbij het stond voor het aantal gevangenen dat de teller al geteld had; maar in dat geval zonder zichzelf mee te tellen.) De bovenindex staat voor de toestand van het licht, net als hiervoor. De naam van de gevangene is niet vetgedrukt als hij of zij legger is, en vetgedrukt als hij of zij pakker is. Gezien de eenvoudige $Pr(2)=0{,}5$ kunnen we simpelweg voorstellen dat de gevangene een muntje opwerpt en laat neervallen: kruis staat voor leggen en munt staat voor pakken. Nogmaals voor alle duidelijkheid: de functie Pr heeft als argument de *som* van het aantal knikkers in de pot van de gevangene en de inhoud van de schaal:

$$^{0}\text{Anna}^{1}_{0}\text{Bert}^{1}_{1}\textbf{Caroline}^{0}_{2}\text{Dick}^{1}_{0}\textbf{Bert}^{0}_{2}\textbf{Caroline}^{0}_{2}\,\text{Caroline}^{1}_{1}\textbf{Bert}^{0}_{3}\text{Caroline}^{1}_{0}\textbf{Bert}^{0}_{4}.$$

Anna wordt als eerste ondervraagd en doet het licht aan (= legt haar knikker neer). Dan wordt Bert ondervraagd, gooit een muntje op, kruis, en doet daarom niet het licht uit (= hij is legger en niet pakker), Caroline komt binnen, munt, en doet het licht uit, dan Dick, licht aan, dan nogmaals Bert, munt, dus deze keer licht uit. Bert heeft nu 2 knikkers in zijn pot. Dan nogmaals Caroline, kruis, ze doet dus niets; en daarna nogmaals, munt, ze doet het licht aan. Bert wordt nu ondervraagd, munt, en doet het licht uit. Dit moment is cruciaal. Bert heeft nu meer dan de helft van het aantal knikkers dus blijft nu alleen nog pakken. Dit moment breekt al snel aan. Caroline laat nog een knikker vallen (ze heeft geen keus, $Pr(1) = 1$), waarop Bob (die ook geen keus heeft, $Pr(3) = 0$) de knikker pakt en er nu vier heeft. Victorie! Slechts tien dagen in de gevangenis!

Het is van groot belang hoe de functie Pr gekozen wordt. In het bijzonder moet $Pr(2) = 0$ uitgesloten zijn, want dan kan een situatie optreden waarin twee gevangenen allebei precies de helft van het aantal knikkers hebben en ze koppig vast blijven houden. Dan termineert het protocol niet, en blijft iedereen tot alle eeuwigheid in de gevangenis.

Optimalisering en synchronisatie

De gevangenen kunnen de ondervragingen kennelijk met slimme trucs wellicht beperken, en zich eerder in vrijheid stellen. Als niets bekend is over de tijdsduur tussen ondervragingen, weten we geen slimmere trucs (ook niet voor honderd gevangenen) dan dat iedere niet-teller voor zichzelf bijhoudt hoeveel keer hij het licht uit en aan ziet gaan, net als Bert en Caroline hiervoor. (Van het probabilistische protocol weten we niet hoe de verwachte uitvoeringsduur zich verhoudt tot de ‘gewone’ protocollen.)

Hoe lang duurt het gemiddeld voordat Anna bekend kan maken dat iedereen ondervraagd is? Die vraag is zinloos, want de tijdsduur tussen ondervragingen is onbekend. Hoeveel ondervragingen duurt het dan gemiddeld voordat Anna dat bekend kan maken? Die vraag heeft wel zin. Dit hangt natuurlijk af van de verroostering van de ondervragingen door de gevangenisbewaarders. Als de verroostering puur op basis van toeval is (‘random’) kunnen we dit bepalen. Het antwoord is leuker als we eveneens aannemen dat er iedere dag een ondervraging is.

Stel dat er iedere dag precies één ondervraging is. Hoe lang duurt het gemiddeld voordat de honderd gevangenen vrijgelaten worden?

Eerst moet een niet-teller ondervraagd worden, om het licht aan te doen. De kans hierop is $99/100$. Dan moet de teller ondervraagd worden, om het licht weer uit te doen. De kans daarop is $1/100$. Daarna moet een niet-teller die nog niet ondervraagd is, het licht weer aandoen. Die kans is nu $98/100$. Daarna de teller weer, dat blijft altijd $1/100$. Enzovoorts. De verwachtingswaarde in dagen is net andersom. Het verwachte aantal dagen voordat de eerste niet-teller ondervraagd wordt is $100/99$, ongeveer een dag dus, de verwachting dat daarna de teller ondervraagd wordt is $100/1$, honderd dagen dus, enzovoorts. De gemiddelde tijd die verstrijkt voordat de teller kan verklaren dat iedereen ondervraagd is, is daarom:

$$\frac{100}{99} + 100 + \frac{100}{98} + 100 + \cdots + \frac{100}{2} + 100 + \frac{100}{1} + 100.$$

De deelreeks $\frac{100}{99} + \frac{100}{98} + \cdots + \frac{100}{2} + \frac{100}{1}$ telt op tot ongeveer 518 dagen en voor de rest hebben we 99 keer 100 dagen: 9.900 dagen. Dit sommeert tot 10.418 dagen wat ruim $28{,}5$ jaar is. Dat valt toch wat tegen. Het is maar goed dat vast niet alle gevangenen wiskundigen zijn, anders waren ze waarschijnlijk niet aan de uitvoering van het protocol begonnen.

Nogmaals: als we verder niets over de regelmaat van ondervragingen weten, kunnen we de verwachte 9.900 ondervragingen niet naar beneden krijgen. Maar als de gevangenen weten dat er precies één ondervraging per dag is, kan het echter (gemiddeld) ook veel sneller. We onderzoeken eerst weer de situatie voor drie gevangenen: stel dat teller Anna niet op de eerste dag ondervraagd wordt. Dan weet ze meteen al dat daarom Bert of Caroline ondervraagd moeten zijn en de lamp nu dus aan is, zonder zelf de lamp aan gezien te hebben. Kunnen we dit uitbuiten?

Protocol 7. *Het protocol bestaat uit twee stadia. Het eerste stadium duurt* 100 *dagen en verloopt als volgt. De gevangene die voor het eerst twee keer ondervraagd wordt, doet het licht aan. Stel dit is op dag* m*. Op dag* 100 *van het eerste stadium: als het licht uit is verklaart de gevangene die dan ondervraagd wordt dat iedereen ondervraagd is; anders, dus als het licht aan is, doet die gevangene het licht uit. Het tweede stadium verloopt als volgt; er zijn nu drie verschillende rollen.* (i) *De teller is de gevangene die in stadium één voor het eerst twee keer ondervraagd is. Voor de teller geldt: Als je ondervraagd wordt en de lamp is uit doe dan niets; als je ondervraagd wordt en de lamp is aan, doe hem dan uit — houdt bij hoe vaak je dat doet; als je de lamp* $101 - m$ *keer hebt uitgedaan, verklaar dan dat alle* 100 *gevangenen ondervraagd zijn.* (ii) *De gevangenen die in het eerste stadium het licht uit gezien hebben, doen niets in het tweede stadium.* (iii) *Voor de overige gevangenen geldt: Als je ondervraagd wordt en de lamp is uit en je hebt hem nog niet eerder aangedaan, doe hem dan aan; als je ondervraagd wordt en de lamp is uit en je hebt hem al eerder aangedaan, doe dan niets; als je ondervraagd wordt en de lamp is aan, doe dan niets.*

Als $m = 2$, krijgen we de oorspronkelijke telling tot 99 terug. Dan levert dit geen tijdwinst op (maar zelfs verlies, namelijk van de 100 dagen in het eerste stadium). Als m maar groter dan 2 is, levert dit protocol tijdwinst op ten opzichte van Protocol 4. Het gemiddelde aantal dagen voordat iemand voor de tweede keer ondervraagd wordt is 13. Dit betekent dat de hiermee aangewezen teller al van 11 andere gevangenen weet dat ze ondervraagd zijn voordat hij met tellen begint. (En die 11 doen in het tweede stadium niets meer!) Hij hoeft dus nog maar tot 88 te tellen, in plaats van tot 99. De tijdwinst die hiermee bereikt wordt is een jaar of vier. Dit komt voornamelijk omdat de teller 11 keer minder dan 100 dagen hoeft te wachten totdat hij het licht weer uit kan doen.

Herkomst en meer informatie

Op http://domino.watson.ibm.com/Comm/wwwr_ponder.nsf/challenges/July2002.html van IBM Research wordt in 2002 vermeld dat “this puzzle has been making the rounds of Hungarian mathematicians’ parties” in een versie voor 23 gevangenen. (En we zouden erg graag een Hongaarse wiskundige spreken die hier meer over weet.) Het raadsel deed de ronde in de Verenigde Staten vanaf 2001. William Wu, destijds van Stanford University, was vanaf dit vermoedelijke begin bij de ontwikkelingen en oplossingen van dit raadsel betrokken, zie wuriddles.com. Het raadsel wordt uitvoerig behandeld in het tijdschrift *Mathematical Intelligencer* door Paul Dehaye, Daniel Ford en Henry Segerman in [1], in *Mathematical Puzzles: A Connoisseur’s Collection* door Paul Winkler in [10], in de *Nieuwe Wiskrant* door Hans van Ditmarsch in [3], en in ‘One hundred prisoners and a lightbulb – logic and computation' door Hans van Ditmarsch, Jan van Eijck en William Wu in [7] en in [8]. Het is verder een hoofdstuk in het logische-puzzelboek *Honderd gevangenen en een gloeilamp* [5] (deze bijdrage is grotendeels gebaseerd op dat hoofdstuk). Een Engelse en geheel herziene vertaling van [5], waarin ook nog meerdere andere puzzels staan, verschijnt in 2015 bij Springer/Birkhauser.

In de variant van het protocol waarin een niet-teller het weet voordat dat teller het weet, door het aantal keren uit/aan te tellen (zie paragraaf 'Soms weet je het, voordat de teller het weet'), is het niet zo waarschijnlijk dat een niet-teller voor de teller weet dat iedereen ondervraagd is (aangenomen dat toeval bepaalt welke gevangene ondervraagd wordt). Bij drie gevangenen is die kans nog 50%, maar bij 100 gevangenen is die kans ten hoogste $5{,}63 \cdot 10^{-72}$ (zie [7]). Dit protocol wordt verder beschreven in [4].

Veel academisch werk over deze protocollen valt onder de globale noemer van 'dynamisch epistemische logica', waarin ook andere logische puzzels zijn te analyseren. Andere publicaties in het *Nieuw Archief voor Wiskunde* hierover zijn: [2, 6, 9] — waarbij we natuurlijk niet kunnen nalaten het 'Som-en-Product'-probleem van Freudenthal te noemen ([6] is een analyse van [9] in dynamisch epistemische logica). Dit probleem wordt ook, uitvoerig, behandeld in [5].

Onder de aanname van synchronisatie, dat wil zeggen: iedere dag vindt een ondervraging plaats en alle gevangenen weten dat (zie paragraaf 'Optimalisering en synchronisatie'), valt nog op allerlei manieren te optimaliseren. Het is niet bekend wat de kleinste verwachtingswaarde is voor de terminatie van een protocol om het probleem op te lossen, het record staat op ongeveer negen jaar. Dat is al een stevige verbetering ten opzichte van de 28 en een half jaar die we in detail uitgelegd hebben. De snellere algoritmes onderscheiden nog meer dan twee rollen voor de gevangenen, en meerdere fasen in het protocol waarin gevangenen van rol kunnen wisselen (net als hiervoor, in Protocol 5, maar dan nog ingewikkelder). Zie [7] voor details.

Het probabilistische protocol is van Paul Dehaye (ooit uitgewisseld bij email-correspondentie met Hans van Ditmarsch). Het wordt ook in het kort beschreven in [7] en we weten niet of dit ergens grondiger onderzocht is. Het zou interessant zijn om onder de aanname van synchronisatie te weten wat de verwachtingswaarde voor terminatie is van zulke probabilistische protocollen.

Elancheziyan

De illustratie bij deze bijdrage is gemaakt door Elancheziyan. Het puzzelboek [5] heeft zo'n illustratie bij ieder hoofdstuk. Elancheziyan is een Tamil-illustrator en woont in Chennai. Hans heeft een band met IMSc (Institute of Mathematical Sciences) in Chennai, en Elancheziyan heeft in opdracht deze illustraties gemaakt. Elancheziyan spreekt alleen Tamil. Een kort verhaal over de ontwikkeling van de illustraties staat op http://personal.us.es/hvd/lightbulb.html. Elancheziyan is te bereiken per e-mail op het adres elancheziyans@gmail.com – vrienden die Engels spreken vertalen zijn e-mails.

Referenties

1 P. Dehaye, D. Ford en H. Segerman, One hundred prisoners and a lightbulb, *Mathematical Intelligencer* 25(4) (2003), 53–61.

2 H. van Ditmarsch, Het zeven-kaartenprobleem (the seven cards problem), *Nieuw Archief voor Wiskunde* 5/3(4) (2002), 326–333.

3 H. van Ditmarsch, Honderd gevangenen en een gloeilamp, *Nieuwe Wiskrant* 27(1) (2007), 15–18.

4 H. van Ditmarsch, S. Ghosh, R. Verbrugge en Y. Wang, Hidden protocols: Modifying our expectations in an evolving world, *Artificial Intelligence* 208 (2014), 28–40.

5 H. van Ditmarsch en B. Kooi, *Honderd gevangenen en een gloeilamp*, Epsilon Uitgaven, deel 73, 2013.

6 H. van Ditmarsch, J. van Eijck en R. Verbrugge, Publieke werken: Freudenthal's som-en-productraadsel, *Nieuw Archief voor Wiskunde* 5/10(2) (2009), 126–131.

7 H. van Ditmarsch, J. van Eijck en W. Wu, One hundred prisoners and a lightbulb — logic and computation, in F. Lin, U. Sattler en M. Truszczynski, eds., *Proc. of KR 2010 Toronto*, 2010, pp. 90–100.

8 H. van Ditmarsch, J. van Eijck en W. Wu, Verifying one hundred prisoners and a lightbulb, *Journal of Applied Non-Classical Logics* 20(3) (2010), 172–191.

9 H. Freudenthal, Formulering van het som-en-productprobleem, *Nieuw Archief voor Wiskunde* 3/17 (1969), 152.

10 P. Winkler, *Mathematical Puzzles: A Connoisseur's Collection*, AK Peters, 2004.