Kennissystemen selectief wieden

Diderik Batens

Ten geleide

Afdeling 1 bevat dingen die voor zienden nuttiger zijn dan zo'n hond. Als emeritus en onbezoldigd medewerker ben ik verbonden aan het Centrum voor Logica en Wetenschapsfilosofie van de Universiteit Gent; berichten aan `diderik.batens@ugent.be` beantwoord ik soms. Doordat de uitgever competentie of tijd miste om latex-bestanden te verwerken, waren inspanningen van Koen en Jan nodig om onze drie teksten zo lelijk te doen ogen als de andere. Als vrijmetselaar peterde ik onder anderen Jean Paul, wat ook in dit volume gevolgen blijkt te hebben.

1 Verontschuldigend voorwoord

Soms erf je wat. Van Leo Apostel erfde ik onder meer een doctoraatsstudent. Die hield zich met vrij buitenissige dingen bezig [1], zoals het voor een filosoof hoort. Daarom had hij zich omgeschoold van wiskunstenaar tot wijsgeer. Wiskundigen waren op slag ontwapend en in filosofisch gezelschap bleef hij zich doodleuk wiskundige noemen.

Hier hoort een kleine uitweiding. Dat wiskundigen met minder buitenissige dingen zouden bezig zijn dan filosofen heb ik niet beweerd en klopt ook niet. Wel ontwikkelden wiskundigen een jargon dat de buitenissigheid van hun onderzoek verdoezelt. Een cosinus of een kwadratische vergelijking is ruimschoots voldoende om de meeste medemensen te doen denken dat het wel aan hen zal liggen. Misschien zit daar wel de diepere reden waarom Jean Paul zich omschoolde: dat hij geloofde een en ander te hebben wat hij de moeite vond om uit te leggen aan zijn medemensen.

Sinds die erfenis zijn we beiden een lange weg gegaan, soms samen en dikwijls in contact. Daarover had ik het al in het artikel voor een vorige verjaardag van Jean Paul [2]. Toch zal ik, net als de vorige keer, geen 'persoonlijk' artikel schrijven. Ik zal het niet hebben over muziek, vrijmetselarij, vriendschap of de vele andere fascinerende voortbrengselen van mensen. Lees die "fascinerende" als inperking, want mensen brengen ook onzin, wreedheid en ander gruwelijks voort, en dingen, zoals het idee van een opperwezen, die interessant zijn maar niet fascinerend. Ik zal het ook niet hebben over de jonge Jean Paul, zoals editor Bart graag had gehad. Ik kende die jongeman daarvoor te onvolledig en bovendien zijn biografische opstellen mijn ding niet.

Hierna volgt een bijdrage over een logisch onderwerp. Geen artikel uiteraard. Want een redelijke hoeveelheid competente lezers bereik ik alleen in het Engels en Engelstaligen vinden dit hier niet. Daarom schrijf ik in het Nederlands—wie dat nog niet door had, kan best meteen naar de volgende bijdrage bladeren. De anderen zullen zich afvragen of ik verwacht dat zij gaan verder lezen. Zeker weten. Want ik zal tonen dat de technisch soms moeilijke problemen waarmee logici bezig zijn, perfect begrijpelijk kunnen worden gemaakt voor niet-specialisten. En ik zal meteen tonen dat hetzelfde geldt voor de oplossing van die problemen en voor de redenering die ertoe leidt. Verder beweer ik dat het probleem en de oplossing niet alleen begrijpelijk kunnen worden gemaakt, maar ook interessant zullen blijken. Al wat ik van de lezer vraag is te lezen: nagaan wat de zinnen betekenen en niet alleen de woorden begrijpen. Want filosofie is als goede seks: geen lol zonder aandacht.

Wanneer er iets mis loopt met ons kennissysteem—een beetje mis lopen kan al meer dan voldoende zijn—dan is het effect daarvan dat ons kennissysteem invalt. Nochtans ligt het probleem meestal bij een klein onderdeel of enkele kleine onderdelen. Er werden verschillende uitwegen voorgesteld om het effect te beperken, maar die werken niet zo goed—ik kom daar terloops op terug in afdelingen 2 en 3. Wat we zouden willen is dat de problematische onderdelen van ons kennissysteem 'invallen', maar niet het gehele kennissysteem.

De oplossing die ik hierna voorstel gaat terug op een fenomeen wat blijkbaar tot nog toe onopgemerkt bleef: lokale trivialiteit—wat dat is zeg ik in de volgende afdeling. Lokale trivialiteit dook voor het eerst op in een reeds vermeld artikel [2] geschreven voor die vorige verjaardag van Jean Paul omdat het ingaat op (niet-eindige varianten van) modellen van de rekenkunde ontworpen en gepropageerd door Jean Paul Van Bendegem en Graham Priest [3, 4, 5]. Daarna kwam ik lokale trivialiteit opnieuw tegen bij de voorbereiding van een nog niet verschenen artikel [6]. Het eerste artikel betreft rekenkunde, het tweede verzamelingenleer—in een volgende afdeling leg ik wel uit waar die wiskundige theorieën over gaan. Geen van beide artikels betreft lokale trivialiteit.

Sommigen zullen beweren dat beide artikels me ertoe brachten iets te vinden wat ik niet zocht, of ook iets te vinden terwijl ik iets anders zocht. Dertig jaar geleden was het in de mode dat serendipiteit te noemen—toelichting bij die gewichtige naam vind je wel op het internet. Belangrijker en erger is dat sommigen toen dachten dat het vinden van iets terwijl men iets anders zoekt, de sleutel vormt tot het begrijpen van wetenschappelijke ontdekking. Dit geloof is het symptoom van een maanziekte waartegen geen knolselder baat. Ten eerste wordt wetenschappelijke ontdekking niet verhelderd door 'het inzicht' dat je ze doet door naar iets anders te zoeken.[1] Ten tweede, ook als wat je niet zoekt voor je neus ligt, zal je het slechts opmerken wanneer (i) je daartoe de nodige competentie hebt ontwikkeld en (ii) er bovendien de juiste aandacht aan besteedt op het ogenblik waarop het voor je neus ligt.

2 Paraconsistentie

Leken denken wel eens dat er maar een logica bestaat. Zij dolen. Er is wel een meest populaire logica, de zogenaamde klassieke logica **CL**—onthoud die naam (afkorting van het Engelse classical logic), want ik heb die nog herhaaldelijk nodig. Die logica werd aan het eind van de negentiende eeuw ontworpen door Gottlob Frege. Hij wordt meestal gehanteerd in artikels waarin de logische stappen uitdrukkelijk worden uitgeschreven en verder in handboeken. Daarnaast werden er om allerlei redenen heel wat andere logica's ontwikkeld. Dat volgens een bepaalde logica **L** de bewering A volgt uit de verzameling[2] van beweringen Γ zal ik schrijven als $\Gamma \vdash_{\mathbf{L}} A$.[3] Wanneer Γ een eindig aantal beweringen bevat, bijvoorbeeld B en C, dan kunnen we ze ook opsommen. Zo drukt $A \vee B, \neg B \vdash_{\mathbf{L}} A$ uit dat A volgt uit A-of-B en niet-B.[4]

In de klassieke logica **CL** en in een aantal andere logica's geldt dat gelijk welke bewering afleidbaar is uit beweringen die elkaar tegenspreken. In formulekens $A, \neg A \vdash_{\mathbf{CL}} B$. Terloops, de leden van een verzameling spreken elkaar tegen wanneer er (met de gegeven logica) zowel een zin A als zijn negatie $\neg A$ uit afleidbaar is; zo zijn p en $\neg p$ beide met **CL** afleidbaar uit $\{p \wedge q, p \supset r, \neg r\}$.[5] Een dergelijke verzameling van formules (of van zinnen) noemt men tegenstrijdig of *inconsistent*. Met **CL** kan je uit een inconsistente verzameling gelijk welke zin van de beschouwde taal afleiden. Men zegt daarom dat, volgens **CL**, elke inconsistente verzameling *triviaal* is.

Dat kan raar klinken, maar het is niet echt onbegrijpelijk hoe mensen daarop kwamen. Veel logici gaan er, in navolging van een bewering van Aristoteles, van uit dat A en $\neg A$ onmogelijk samen waar kunnen

[1] Tracht jij de kwantummechanica en de relativiteitstheorie te verzoenen, mevrouw? Dan kan ik je twee schitterende studieonderwerpen aanbevelen: de invloed van de aanwezigheid van steenklaver op de vruchtbaarheid van konijnen en het oorzakelijk verband tussen paranoïde schizofrenie en probleemoplossend vermogen bij tuinkabouters.

[2] Sommigen zeggen liever "set"; zoals die Brusselaar zei "De flèche van den tram quatorze is gedéraileerd aan de bifurcation van de Porte de Namur".

[3] Ik meen dat technische uitdrukkingen voldoende worden uitgelegd in dit essay. De wantrouwige lezer verwijs ik naar een goed handboek [7, 8].

[4] Uit "ik ben in Gent of in Sint-Blasius-Boekel" en "in ben niet in Gent" volgt "ik ben in Sint-Blasius-Boekel". Van daar kan je te voet naar 't Blaffend Konijn in Roborst.

[5] Het teken $\supset$ gebruik ik voor de 'materiële' implicatie, net zoals in artikels waarnaar ik verwijs. Dikwijls gebruikt men $\rightarrow$, maar dat teken staat in bedoelde artikels voor een andere implicatie.

zijn.[6] Je kan dat krachtig en beeldrijk uitdrukken door te zeggen: als A en $\neg A$ samen waar zijn, dan is gelijk wat waar. Dat is precies wat $A, \neg A \vdash_{\mathbf{CL}} B$ uitdrukt want B is een *willekeurige* bewering, ze hoeft geen enkel verband te hebben met A. Nu is vanzelfsprekend niemand zo zot om uit de elkaar tegensprekende beweringen A en $\neg A$ een bewering B af te leiden, laat staan *alle* beweringen B. Dat hoeft ook niet. Precies omdat $A, \neg A \vdash_{\mathbf{CL}} B$ uitdrukt dat A en $\neg A$ niet samen waar kunnen zijn, zodat je er gelijk wat uit kan afleiden, heeft het geen zin er wat dan ook uit af te leiden.[7] Maar wie dat wil kan gemakkelijk genoeg en met heel natuurlijke stapjes gelijk welke B afleiden. Om te beginnen volgt $A \vee B$ uit A: als A waar is, dan is immers A-of-B waar; de regel A / $A \vee B$ noemt men *Additie*. En uit $A \vee B$ en $\neg A$ volgt B: als A-of-B waar is en A is vals, dan is B waar; de regel $A \vee B, \neg A$ / B heet *Disjunctief Syllogisme*.

De bewering $A, \neg A \vdash_{\mathbf{CL}} B$ heeft een traditionele Latijnse naam: *Ex Falso Quodlibet*, afgekort als EFQ.[8] Recent noemt men het in het Engels ook wel Explosion—Ontploffing of Explosie? Het idee is wel duidelijk: wanneer een tegenstrijdigheid afleidbaar is uit een theorie en EFQ geldt volgens de onderliggende logica van de theorie, dan spat de theorie letterlijk in gruizelementen uiteen en kan je er niets meer mee aanvangen. De theorie voorspelt alles en verklaart alles en dus ben je er niks mee. Ik heb het al mijn hele leven tegen ontploffingen en ik wens mij ver te houden van de idioten die, van Syrisch puin tot dat van station Maalbeek, van het Witte Huis tot Noord-Korea, om maar te zwijgen over Amerikaanse scholen, graag boem doen of ermee dreigen. Daarom zal ik het bij EFQ houden.

Toch zijn er bezwaren tegen EFQ. Ik vermeld er slechts een, wat mij het meest overtuigend voorkomt. Bijna alle theorieën die mensen ooit gebruikten om de wereld te begrijpen en erin te handelen—de wereld te transformeren—werden blijkbaar vroeger of later vervangen omdat ze fout bleken. Die fouten kwamen bijna steeds aan het licht omdat tegenstrijdigheden werden gevonden, hetzij binnen theorieën, hetzij tussen theorieën, hetzij tussen een theorie en waarnemingsgegevens.[9] Het geheel van onze kennis bevatte dus herhaaldelijk tegenstrijdigheden en die

[6] Elders schrijft Aristoteles het omgekeerde, maar laat ons niet vitten.

[7] Dat is ook de reden waarom mensen er in een discussie op uit zijn te tonen dat hun tegenstrever zichzelf tegenspreekt. Die tegenstrever vertelt onzin, wat als wat hij of zij beweert waar zou zijn, dan zou alles waar zijn. Maar dan zou "alles is vals" ook waar zijn, zodat wat de tegenstrever zegt vals is. Met enige moeite onderdruk ik de neiging op dit leuke pad door te gaan; uit mededogen met Jean Paul houd ik de voetnoten eindig.

[8] In bijna alle logica's waarin EFQ niet geldt, laat men ook Disjunctief Syllogisme wegvallen. In Additie komt namelijk niet eens een negatie voor.

[9] Naast bekende wiskundige theorieën, zoals de verzamelingenleer van Cantor, de verzamelingenleer van Frege en de infinitesimaalrekening van Newton, zijn er heel wat voorbeelden uit de empirische wetenschappen bestudeerd [9, 10, 11, 12, 13, 14, 15, 16].

dienden eruit te worden verwijderd. Let erop dat het soms decennia of zelfs eeuwen duurde voor de tegenstrijdigheden werden verwijderd—een paar notoire gevallen werden nooit verwijderd uit Newtons mechanica, maar komen niet meer voor in Einsteins relativiteitstheorie. Wetenschappers gingen soms lange tijd door met het gebruiken, toepassen en pogen te verbeteren van theorieën die tegenstrijdigheden bevatten. Maar ook wanneer de tegenstrijdigheden snel werden verwijderd nadat ze waren ontdekt, gebeurde de verwijdering door te *redeneren* vanuit de theorieën waarin de tegenstrijdigheden voorkwamen.

Hier zit een belangrijk inzicht. Ook als de wereld consistent kan worden beschreven, kunnen onze theorieën tegenstrijdig zijn. Onze theorieën zijn immers geen ware beschrijvingen van (aspecten van) de werkelijkheid, maar meestal onvolmaakte pogingen om de werkelijkheid te beschrijven. Precies de overtuiging dat de wereld consistent kan worden beschreven, geeft ons een reden om tegenstrijdige theorieën als vals te beschouwen, of om ze minstens als niet-optimaal en weinig efficiënt te beschouwen [17]. Wetenschappers redeneren echter vanuit de theorieën die tegenstrijdig bleken, wanneer ze op zoek gaan naar een niet-tegenstrijdige vervanger ervoor. Welnu, wanneer uit een tegenstrijdige theorie alles afleidbaar is, zoals wordt gesteld door logica's waarin EFQ geldt, dan heeft het geen zin te redeneren vanuit een tegenstrijdige theorie. Aangezien alle beweringen waar zijn volgens die theorie, biedt de theorie zelfs geen vage indicatie op de weg naar een vervangende theorie. EFQ gooit immers alle inconsistente 'theorieën' op een hoop. Ze zeggen allemaal hetzelfde, namelijk alles. Het verband met de oorspronkelijk bedoelde theorie is daardoor totaal zoek. Het besluit van deze alinea is dat ook indien tegenstrijdige theorieën vals zijn, en precies dan, het nodig is te kunnen nadenken over situaties waarin die theorieën waar zijn. Wat mij betreft mag je dat onwerkelijke of imaginaire situaties noemen, maar je hebt ze wel nodig om over inconsistente theorieën te redeneren.[10]

Een verwarrend en grondig fout argument betreffende EFQ werd ontwikkeld door Peter Vickers [18]. Hij meent dat men uit een theorie die zichzelf tegenspreekt alleen maar een willekeurige bewering kan afleiden door eerst een tegenstrijdigheid af te leiden en daarna EFQ toe te passen.[11] Dit klopt niet. De afleiding kan even gemakkelijk gebeuren langs een omweg die onvoorspelbaar lang kan zijn. Ik geef een heel sim-

[10] Voor de logici: inconsistente theorieën moeten bijgevolg modellen hebben; heeft Γ geen modellen, dan is elke formule A een semantisch gevolg van Γ (elk model is ofwel geen model van Γ ofwel een model van A).

[11] Indien Vickers het bij het rechte eind zou hebben, dan zou het volstaan **CL** toe te passen tot men een tegenstrijdigheid heeft afgeleid (deductieve fase). Vervolgens zou men de tegenstrijdigheid wegwerken (creatieve fase die een methode volgt eerder dan een deductieve logica) en daarna zou men weer **CL** toepassen. Vickers slaat ook hier de bal mis: hij ziet niet dat dit neerkomt op het toepassen van een weliswaar rudimentaire paraconsistente logica—de betekenis van "paraconsistente logica" wordt straks duidelijk.

pel voorbeeld. Uit de premissen $\{\neg p, q, \neg r, q \supset (r \vee p)\}$ kan je zowel p als $\neg p$ afleiden en daaruit dan bijvoorbeeld s. De formule s kan echter even goed worden afgeleid langs de volgende weg: uit $q \supset (r \vee p)$ volgt $q \supset ((r \vee p) \vee s)$; uit deze formule samen met q volgt $(r \vee p) \vee s$; uit dit samen met $\neg p$ volgt $r \vee s$; uit dit samen met $\neg r$ volgt s. Zo werd de willekeurige s afgeleid zonder dat p werd afgeleid en zonder dat EFQ werd toegepast. Wie een beetje logica kent ziet meteen dat s ook langs een veel langere omweg kan worden afgeleid uit de premissen.[12]

Laat ons terugkeren naar de lijn van het betoog. Ook wanneer alle inconsistente kennisgehelen vals zijn, hebben we logica's nodig waarin EFQ niet geldt, logica's waarin niet alle tegenstrijdige kennisgehelen op een hoop worden gegooid. Men noemt dat *paraconsistente* logica's, logica's waarin niet voor alle A en B geldt dat $A, \neg A \vdash B$, logica's dus volgens dewelke sommige inconsistente theorieën niet triviaal zijn.

Door een paraconsistente logica te hanteren elimineren we de nadelen van EFQ. Helaas heeft EFQ ook voordelen. Ik vermeld er snel twee. We zagen reeds dat Additie en Disjunctief Syllogisme samen toelaten EFQ af te leiden. Als EFQ wegvalt, moet ook Additie of Disjunctief Syllogisme wegvallen.[13] Er zijn ook nog een aantal andere combinaties van regels die er op zich alle 'natuurlijk' uit zien, maar die samen tot EFQ leiden. Voor elk van die combinaties moet telkens minstens een regel wegvallen. Een tweede voordeel van EFQ is dat het kennisonderdelen op een overtuigende manier elimineert. Eens men vaststelt dat alles eruit volgt, staat het vast dat het onderdeel dient te worden vervangen.

In wat voorafgaat zit een opvallende tegenstelling. EFQ reduceert een inconsistent kennisonderdeel tot de triviale verzameling zinnen. Dat daardoor het kennisonderdeel overtuigend wordt geëlimineerd noemde ik een voordeel. Dat het daardoor onmogelijk wordt over het onderdeel te redeneren om een niet-tegenstrijdige vervanging te vinden, noemde ik een nadeel. De kritische lezer zal hier wat uitleg willen, maar de anderen mogen meelezen. De uitleg is dat wetenschappers zowel van mening als van methode kunnen veranderen. In een eerste fase is het elimineren van het inconsistent kennisonderdeel belangrijk. In een tweede fase kan men overgaan tot het zoeken van een consistente vervanging. Ik kom hierop terug in afdeling 7. Eerst kunnen we ons immers beter concentreren op een andere moeilijkheid: Hoe het tegenstrijdige kennisonderdeel elimineren zonder ons hele kennissysteem te elimineren?

[12] Dit komt omdat **CL** een transitieve logica is, net zoals de meeste deductieve logica's. Een conclusie die in tweehonderd stappen uit de premissen werd afgeleid volgt evenzeer uit de premissen als een conclusie die er in een stap werd uit afgeleid.

[13] Men kan dit omzeilen door een logica te bepalen als een procedure, zoals ik elders aantoonde [19]. Voor de logici: de bekomen logica is dan niet-transitief.

3 Monoliet of lappendeken?

Er zijn grosso modo twee grote opvattingen over het geheel van onze kennis. De traditionele opvatting ziet dat geheel als een monolithisch blok. Het omvat alle kennis van goede kwaliteit, alles wat wij momenteel beschouwen als correct, of minstens als het beste wat we nu hebben. Voor sommige kennisonderdelen kunnen er alternatieven zijn die elkaar tegenspreken en waartussen men momenteel niet kan kiezen. In dat geval zal men verschillende kennissystemen naast elkaar beschouwen: telkens een van de alternatieven gecombineerd met de rest van onze kennis. Het geheel heeft een enkele onderliggende logica. Ook in dit verband kan men meerdere alternatieven beschouwen,[14] maar in elk beschouwd kennissysteem is er een enkele onderliggende logica.

De tweede grote opvatting ziet kennis als een lappendeken. Er zijn allerlei theorieën en daarnaast kunnen er meer fragmentaire brokjes kennis zijn: waarnemingsgegevens, empirische veralgemeningen, en dergelijke. Ik zal eenvoudigheidshalve het woord "kennisonderdeel" blijven gebruiken. Elke theorie en zelfs elk min of meer samenhangend kennisonderdeel heeft een onderliggende logica die bepaalt wat het kennisonderdeel precies inhoudt. Die onderliggende logica's kunnen verschillend zijn en kennisonderdelen kunnen elkaar in allerlei opzichten tegenspreken. Daardoor kan je wat volgt uit het ene onderdeel niet zonder meer bij het andere onderdeel voegen. Men zal moeten *regels* vastleggen die bepalen welke gevolgen van het ene kennisonderdeel kunnen worden gevoegd bij het andere.[15] Er zijn daarover een aantal vrij uitgewerkte voorstellen gepubliceerd [20, 21, 22, 23, 24].[16] Deze tweede visie op kennis is complexer dan de eerste, maar ook veel realistischer en veel soepeler. Zo maakt ze ruimte voor een pluralistische kijk op kennis. Die term wordt met verschillende betekenissen gebruikt. Wat ik vooral belangrijk vindt, hangt samen met wat in de laatste voetnoot staat. In een probleemoplossingssituatie zullen we een bepaalde selectie uit ons kennissysteem gebruiken en deze selecties kunnen elkaar onderling in allerlei opzichten tegenspreken [21, 22].

Onder de aanhangers van de monolithische opvatting vindt men bijna alle logische monisten. Dat zijn mensen die geloven dat de Ene Ware Logica bestaat. De meeste klassieke logici horen daarbij, maar ook intuïtionisten, veel relevante logici en dialetheïsten zoals Priest. Er zijn inderdaad zoveel opvattingen over de Ene Ware Logica als over de

[14] Zo kan men in principe redeneren over het gehele kennissysteem met de ene keer **CL** als onderliggende logica en de andere keer de intuïtionistische logica.

[15] De formulering is wat slordig want het hoeft niet echt langs regels te gebeuren. Voor de huidige tekst maakt dat echter niet zoveel uit.

[16] Voor mij is het essentieel dat de kennisonderdelen verwijzen naar een concrete probleemoplossingssituatie terwijl ze dat voor Brown en Priest niet doen. Volgens mij kunnen aan elkaar tegengestelde varianten van dezelfde theorie worden gehanteerd in verschillende probleemoplossingssituaties.

Ene Ware God. De aanhangers zien elkaar als argument voor het bestaan van het Ding, als fout wat de natuur van het Ding betreft, en als tegenstanders bij het verspreiden van de leer van het Ding. Ermee discussiëren veroorzaakt zoveel hoofdpijn, dat je ze best stimuleert om elkaar te lijf te gaan, met woorden wel te verstaan.

In afdeling 2 heb ik betoogd dat we in bepaalde situaties een paraconsistente logica nodig hebben, een logica zonder EFQ. Wie het daarmee eens is en een monolithische opvatting over kennis aanhangt, belandt in de gracht. Dit hier aantonen is te technisch en te complex, zodat ik de geïnteresseerde lezer verwijs naar enkele recente artikels [25, 26]. Daarom zal ik het vanaf hier nog slechts hebben over kennis als lappendeken.

Wanneer het lappendeken onderdelen bevat waarvan men weet dat ze tegenstrijdig zijn, dan zal men ze als onderliggende logica een paraconsistente logica meegeven. Voor niet-tegenstrijdige onderdelen zal men een onderliggende logica kiezen waarin EFQ onverkort geldt.[17] Laten we de negatie van een dergelijke logica slordig weg een 'klassieke' negatie noemen. Merk op dat een paraconsistente logica naast de paraconsistente negatie ook een klassieke negatie kan hebben, waarvoor EFQ onverkort geldt; voorbeelden volgen nog. Vanzelfsprekend moet het kennisonderdeel zo geformuleerd zijn dat alle gekende inconsistenties geformuleerd zijn met een paraconsistente negatie.[18] Elk kennisonderdeel dat een klassieke negatie bevat, kan tegen alle verwachtingen in een tegenstrijdigheid voor die negatie blijken te bevatten. Wanneer dit het geval is, is uit dat onderdeel elke bewering afleidbaar en is er geen garantie dat de tevoren vermelde regels verhinderen dat het gehele kennissysteem triviaal wordt. Kennis zien als lappendeken helpt hiertegen spijtig genoeg niet.

Sommige lezers hebben misschien de indruk dat dingen hier nodeloos complex worden voorgesteld. Is de moeilijkheid niet opgelost wanneer we eisen dat elk onderdeel dat bij ons kennissysteem wordt gevoegd, consistent moet zijn? Wel, je kan dat wel eisen, maar je kan niet nagaan of het zo is. Het is net alsof men in het reglement van een katholiek seminarie zou decreteren dat mensen zich niet mogen inschrijven indien de poging in celibaat te leven op een bepaald punt in hun leven tot pedofilie zal leiden. Van uiterst simpele kennisonderdelen kan je aantonen dat ze consistent zijn, van de meeste onderdelen van onze hedendaagse kennis kan je dat beslist niet. Dat komt niet omdat ze inconsistent zijn, maar omdat er geen beslissingsmethode is om aan te tonen dat ze consistent zijn indien ze het zijn. Klinkt ingewikkeld, maar is het niet. Je kan voor **CL** en voor veel andere logica's een machine (of een stel

[17] Wanneer volgens een consistente theorie zowel $A \vee B$ als $\neg A$ waar zijn, dan moet ook B waar zijn volgens die theorie. Daarom kunnen we beter meteen een logica kiezen waarin Disjunctief Syllogisme geldt.

[18] Er staat terecht "een", want er kunnen meerdere paraconsistente negaties voorkomen in dezelfde taal.

logici) aan het werk zetten om de gevolgen van een kennisonderdeel op te sommen. Elk gevolg zal vroeg of laat worden gevonden, maar er zijn er wel oneindig veel. En er is geen (algemene) methode om aan te tonen dat een bepaalde bewering geen gevolg is van het kennisonderdeel. Stel dat we vinden dat A een gevolg is. Wat ons dan interesseert is of $\neg A$ ook een gevolg is. Als het er een is, dan zal het ooit worden gevonden; anders uiteraard niet. Dat het, na willekeurig veel uren rekenen, niet is gevonden, kan komen omdat het geen gevolg is, maar ook omdat het een gevolg is dat pas veel later zal worden gevonden. Er zijn leuke en begrijpelijke boeken [27] voor wie meer over dit soort problemen wenst te weten.

Laat ons even een alinea teruggaan. Wat was daar precies spijtig? Wat we graag zouden willen, is dat het kennisonderdeel wel wordt 'uitgeschakeld', maar dat de rest van ons kennissysteem overeind blijft. Dat het afleiden van de inconsistentie (voor de klassieke negatie) het kennisonderdeel zelf geheel onzinnig maakt, is niet zo erg. We kunnen de klassieke negatie in het onderdeel achteraf lezen als een paraconsistente negatie en met het resultaat daarvan op zoek gaan naar een consistente vervanging—zie ook afdeling 7.[19] Maar ook al wordt het kennisonderdeel op zich triviaal, het is belangrijk dat die trivialiteit niet wordt gespreid over het gehele kennissysteem. Alleen was het, toch tot voor kort, niet in te zien hoe iets dergelijks kan worden gerealiseerd.

4 Trivialiteit uitdrukken

In twee studies kwam ik een eigenaardig fenomeen tegen. De ene betreft de theorie **APA** (adaptieve Peano rekenkunde) [2], de andere betreft AFS, niet een theorie maar een familie van theorieën voor adaptieve Fregeaanse verzamelingenleer [6]. In beide gevallen kan ik het punt maken zonder adaptieve logica's te hoeven bespreken, wat de lezer een hele complicatie spaart. Voor de onderliggende paraconsistente logica worden verschillende namen gebruikt—mijn geprefereerde naam is **CLuNs** [28]. Veel hoeft de lezer niet te onthouden over **CLuNs**: deze logica gedraagt zich zoals verwacht (en zoals **CL**) voor alle logische tekens behalve de negatie. De negatie is paraconsistent en negaties van complexe formules worden gereduceerd tot eenvoudigere formules: $\neg\neg A$ betekent hetzelfde als A, $\neg(A \wedge B)$ betekent hetzelfde als $\neg A \vee \neg B$, $\neg \forall x A$ betekent hetzelfde als $\exists x \neg A$, en zo verder.

De taal van **APA** bevat, naast de haakjes (en), de logische tekens $\vee$, $\wedge$, $\supset$, $\equiv$, $\neg$, $\forall$, $\exists$ en $=$, en ook variabelen x, y, z, x_1, …, die

[19] De bedoelde transformatie wordt niet uitgevoerd op de verzameling gevolgen, want dat is de triviale verzameling. Elk kennisonderdeel is steeds gegeven op een 'eindige' manier en daarop gebeurt de transformatie. Dat verschillende 'eindige' verzamelingen equivalent zijn, doet niets af aan het feit dat er een motivering kan zijn voor sommige van deze verzamelingen, terwijl de andere louter met een logisch trucje bekomen zijn.

eveneens voorkomen in de onderliggende logica. Daarnaast bevat die taal ook nog specifieke tekens voor de rekenkunde: de constante 0, de opvolgersfunctie $'$, die één argument heeft,[20] en de functies $+$ en $\times$, die beide twee argumenten hebben. De axioma's van **APA** komen neer op de traditionele Peano axioma's en laten toe alles aan te tonen wat in de klassieke rekenkunde kan worden aangetoond.

Bekijk nu even de formule $\forall x \forall y\, x = y$. Vanzelfsprekend is dat geen stelling van **APA**. Maar vraag je even af wat uit die formule volgt, met andere woorden, wat er allemaal waar is indien deze formule waar zou zijn in de klassieke rekenkunde. Dat $x = y$ betekent dat x en y dezelfde entiteit aanduiden, in dit geval hetzelfde getal—met "getal" wordt in geheel deze tekst een natuurlijk getal bedoeld; wie dit niet kent mag denken aan de positieve gehele getallen: 0, 1, 2, enz., 'officieel' genoteerd als 0, $0'$, $0''$, enz. De formule $\forall x \forall y\, x = y$ betekent met andere woorden dat er slechts een enkel getal bestaat, dat alle getalnamen hetzelfde ding benoemen, namelijk dat ene getal. Nu is het in de rekenkunde aantoonbaar dat 1 verschillend is van 0, $0 \neq 1$.[21] Maar als alle getallen hetzelfde zijn, dan zijn ook 0 en 1 hetzelfde, zodat we 1 door 0 mogen vervangen in de stelling $0 \neq 1$ en zo bekomen dat $0 \neq 0$ een stelling is. En als alle getallen gelijk zijn aan elkaar en dus ook aan 0, volgt uit $0 \neq 0$ dat $\forall x\, x \neq x$. Uit $\forall x \forall y\, x = y$ volgt ook dat x zijn eigen opvolger is, $\forall x\, x = x'$, en ook dat x gelijk is aan de opvolger van gelijk welk getal, $\forall x \forall y\, x = y'$. En aangezien sommen en producten ook getallen opleveren, volgt ook $\forall x \forall y \forall z\, x = y + z$ en $\forall x \forall y \forall z\, x = y \times z$. Maar, zoals we zagen, volgt $\forall x \forall y\, x \neq y$ uit $\forall x \forall y\, x = y$; en uit $\forall x \forall y\, x \neq y$ volgen, met dezelfde redenering als hierboven, $\forall x\, x \neq x'$, $\forall x \forall y\, x \neq y'$, $\forall x \forall y \forall z\, x \neq y + z$ en $\forall x \forall y \forall z\, x \neq y \times z$. Gegeven de betekenis van de universele kwantor volgen uit de bovenstaande formules ook de formules die je bekomt door de kwantoren weg te laten en x, y en z telkens te vervangen door een bepaald getal. Dit alles volgt uit de formule $\forall x \forall y\, x = y$.

Je kan gemakkelijk nazien dat *alle* formules van de rekenkundige taal volgen uit de formule $\forall x \forall y\, x = y$. Want ten eerste volgen uit die formule alle uitdrukkingen waarin alleen haakjes, getalnamen, $=$ en $\neg$ voorkomen—doorloop maar even alle gevallen. En ten tweede kan je over de logica **CLuNs** aantonen dat uit de verzameling van al die formules meteen ook alle andere formules volgen—die waarin nog andere tekens voorkomen dan haakjes, getalnamen, $=$ en $\neg$. Met andere woorden, binnen de rekenkunde volgen alle formules van de taal van de rekenkunde uit de formule $\forall x \forall y\, x = y$. Maar dit betekent dat, binnen de rekenkunde, $\forall x \forall y\, x = y$ trivialiteit uitdrukt. We kunnen dat exact

[20] De opvolger van 0 is 1; de opvolger van 1 is 2; en zo verder.

[21] De uitdrukking $0 \neq 1$ is een afkorting voor $\neg\, 0 = 1$. Het is belangrijk dit op te merken omdat we zo dadelijk uitdrukkingen als $\sim 0 = 1$ zullen tegenkomen, waarin $\sim$ de klassieke negatie is.

en proper formuleren door te schrijven: $\vdash_{\mathbf{APA}} \forall x \forall y\, x = y \supset A$, in woorden: het is een stelling van **APA** dat $\forall x \forall y\, x = y$ elke formule van de taal van de rekenkunde impliceert.

Let erop dat ik hierboven schrijf “binnen de rekenkunde”. We hebben inderdaad stellingen van de rekenkunde nodig om aan te tonen dat $\forall x \forall y\, x \neq y$ volgt uit $\forall x \forall y\, x = y$. Maar we hebben geen stellingen van de rekenkunde nodig om te tonen dat enige andere formule eveneens afleidbaar is uit $\forall x \forall y\, x = y$. Daarvoor volstaat de logica **CLuNs**. We kunnen dan ook besluiten dat elke formule van de taal van de rekenkunde afleidbaar is uit de formule

$$\forall x \forall y (x = y \wedge x \neq y)\,,$$

zodat deze formule trivialiteit uitdrukt binnen de taal van de rekenkunde (ook al zouden we andere axioma’s kiezen om de betekenis vast te leggen van de typisch rekenkundige symbolen: de constante 0, de functie met één argument $'$ en de functies met twee argumenten $+$ en $\times$).

Wat is hier zo merkwaardig aan? De onderliggende logica van **APA** is de paraconsistente logica **CLuNs**. Die logica heeft slechts één negatie en die negatie is paraconsistent.[22] En toch is elke formule van de taal van de rekenkunde afleidbaar uit $\forall x \forall y (x = y \wedge x \neq y)$, zodat deze formule trivialiteit uitdrukt: $\forall x \forall y (x = y \wedge x \neq y) \vdash_{\mathbf{CLuNs}} B$ voor gelijk welke B.

Even merkwaardig is een onmiddellijk gevolg hiervan. We kunnen de klassieke negatie *definiëren* binnen de taal van de rekenkunde, bijvoorbeeld door de definitie ${\sim}A =_{df} A \supset \forall x \forall y (x = y \wedge x \neq y)$. Het zo gedefinieerde nieuwe teken $\sim$ heeft inderdaad alle eigenschappen van de klassieke negatie. Dat EFQ ervoor geldt zal wel meteen duidelijk zijn. Dit komt immers neer op $A, A \supset \forall x \forall y (x = y \wedge x \neq y) \vdash B$. Welnu, (i) uit A en $A \supset \forall x \forall y (x = y \wedge x \neq y)$ volgt $\forall x \forall y (x = y \wedge x \neq y)$ (door Modus Ponens) en, zoals we net zagen, volgt gelijk welke formule van de taal van de rekenkunde uit $\forall x \forall y (x = y \wedge x \neq y)$.[23]

De oplettende lezer vraagt zich misschien hoe het zit met relaties als “groter dan”. Deze relatie kan worden gedefinieerd in **APA** en aangezien $\forall x \forall y (x = y \wedge x \neq y)$ trivialiteit uitdrukt voor de taal zonder gedefinieerde symbolen, doet het dat ook voor de taal met gedefinieerde symbolen. Concreet zullen we $x > y$ bepalen als $\exists z (y + z = x \wedge z \neq 0)$ (er is een z zodat de som van y en z gelijk is aan x en z verschilt van 0). De formule (eens proper gesloten met kwantoren) is inderdaad afleidbaar uit $\forall x \forall y (x = y \wedge x \neq y)$ onder meer omdat $0 = 0 + 0 \wedge 0 \neq 0$ eruit afleidbaar is.

[22] Voor de logici: **CLuNs** is *strikt paraconsistent*, wat betekent dat er in het (propositioneel of predikatief) *taalschema* geen enkele formule A is waarvoor geldt dat $A, \neg A \vdash_{\mathbf{CLuNs}} B$.

[23] Het enige wat we nog meer nodig hebben om te tonen dat $\sim$ de klassieke negatie is, is $\vdash_{\mathbf{CLuNs}} A \vee {\sim}A$, wat neerkomt op $\vdash_{\mathbf{CLuNs}} A \vee (A \supset \forall x \forall y (x = y \wedge x \neq y))$. Dit volgt meteen uit $\vdash_{\mathbf{CLuNs}} A \vee (A \supset B)$.

Laat ons nu veel sneller de situatie in AFS bekijken. De taal van deze theorieën voor Fregeaanse verzamelingenleer bevat, naast de haakjes (en), de logische tekens $\vee$, $\wedge$, $\supset$, $\equiv$, $\neg$, $\forall$, $\exists$ en $=$, en de variabelen x, y, z, x_1, ..., die eveneens voorkomen in de onderliggende logica. Het verschil met de taal van **APA** zit hem in de specifieke tekens voor de verzamelingenleer. Het gaat om een enkel teken, het relationeel predikaat $\in$, waar $x \in y$ uitdrukt dat x lid is van y. Gemakkelijkheidshalve neemt men de taal iets ruimer en wel zo dat $\{x \mid A(x)\}$ staat voor de verzameling van alle x waarvoor geldt dat $A(x)$. Individuele constanten komen niet voor in de taal, maar men kan er wel definiëren. Zo wordt de inconsistente Russell-verzameling gedefinieerd als $R =_{df} \{x \mid x \notin x\}$ en de (generieke) verzameling die een gegeven verzameling y als enig lid heeft, definieert men als: $\{y\} = \{x \mid x = y\}$.[24] De onderliggende logica is weer **CLuNs** en de axioma's hebben we niet echt nodig, zodat ik ze onvermeld laat.

Ook in deze taal zijn er formules waaruit alle andere formules afleidbaar zijn, bijvoorbeeld $\forall x \forall x(x = y \wedge x \neq y \wedge x \in y \wedge x \notin y)$, waarin $x \notin y$ weer staat voor $\neg\, x \in y$.[25] Anders gezegd, in deze verzamelingstheoretische taal drukt

$$\forall x \forall x(x = y \wedge x \neq y \wedge x \in y \wedge x \notin y)$$

trivialiteit uit. En opnieuw kunnen we de klassieke negatie definiëren: $\sim\! A =_{df} A \supset \forall x \forall x(x = y \wedge x \neq y \wedge x \in y \wedge x \notin y)$.

Dat een formule van een predikatieve taal trivialiteit uitdrukt is uitzonderlijk.[26] Daarom voert men soms het symbool $\perp$ in om trivialiteit uit te drukken—men definieert het impliciet langs het axiomaschema $\perp \supset A$. Logici weten dat je de klassieke negatie kan definiëren met de materiële implicatie $\supset$ en een formule die trivialiteit uitdrukt. Voor leken is de meest spectaculaire formulering van het resultaat wellicht dat de klassieke negatie definieerbaar is in de vermelde talen, ondanks het feit dat de onderliggende logica strikt paraconsistent is.

5 Lokale trivialiteit

Veronderstel even dat we de taal van de rekenkunde uitbreiden. We zouden er een zin Z kunnen aan toevoegen, bijvoorbeeld "Het regent vandaag op de Koppenberg". En zie, de hele zaak valt op zijn gat. Want al volgt elke zin van de taal van de rekenkunde uit $\forall x \forall y(x = y \wedge x \neq y)$, zodra je die taal uitbreid met Z volgen een oneindig aantal zinnen er niet

[24] Ook allerlei operatoren kunnen worden gedefinieerd, zoals $x \subseteq y =_{df} \forall z(z \in x \supset z \in y)$ (x is een deelverzameling van y indien alle leden van x ook leden van y zijn).

[25] Sommigen zullen het verhelderend vinden dat trivialiteit ook wordt uitgedrukt door $\exists x(\forall y\, x = y \wedge x \neq x \wedge x \in x \wedge x \notin x)$.

[26] Voor de logici: in het predikatief taalschema, wat traditioneel wordt gebruikt om **CL** te bepalen, lukt het niet.

meer uit, Z, $Z \wedge 0 = 0$, en zo verder. Hetzelfde gebeurt wanneer men de taal van de rekenkunde uitbreid met een nieuw predikaat, bijvoorbeeld I, waar Ix staat voor "Diderik vindt x een interessant getal".[27] Welnu, $I11$ is *niet* afleidbaar uit $\forall x \forall y(x = y \wedge x \neq y)$ en hetzelfde geldt voor een oneindig aantal andere formules, zoals $\neg I11$, $0 = 0 \wedge I257$ enz.

Opdat de lezer niet zou denken dat hier iets complex aan de hand is, zeg ik even waarom Z niet afleidbaar is uit $\forall x \forall y(x = y \wedge x \neq y)$. De atomaire (lees simpelste) uitdrukkingen van de taal van de rekenkunde hebben alle de vorm $t_1 = t_2$, waarin t_1 en t_2 getalnamen zijn (zoals 0, 17, $6 + (11 \times 4)$, ...). Alle complexe uitdrukkingen van de taal van de rekenkunde worden bekomen door atomaire uitdrukkingen met logische tekens (behalve de gelijkheid) te verbinden of door een getalnaam te vervangen door een variabele en een kwantor over die variabele voor de gehele uitdrukking te schrijven. We hebben gezien dat al dergelijke formules volgen uit $\forall x \forall y(x = y \wedge x \neq y)$. Maar Z en It (voor t een getalnaam) zijn geen dergelijke formules en volgen er niet uit. Wat er ook niet uit volgt zijn formules die slechts waar zijn wanneer Z of een formule van de vorm It waar is.

Dat is de reden waarom ik het heb over *lokale* trivialiteit: er zijn formules die trivialiteit uitdrukken in de zin dat alle formules van de taal van de rekenkunde eruit volgen. In sommige rijkere talen drukken die formules echter niet meer trivialiteit uit. Wanneer het om een zware uitbreiding gaat, dan zijn er geen formules waaruit alle formules van de uitgebreide taal afleidbaar zijn. Voor sommige lichtere uitbreidingen zijn er soms formules *van de uitgebreide taal* die trivialiteit uitdrukken in de uitgebreide taal. Dat heeft echter geen belang. Wat het punt is, staat al hoger: een formule van de oorspronkelijke taal drukt slechts lokale trivialiteit uit wanneer alle formules van de oorspronkelijke taal eruit volgen, maar niet alle formules van alle uitbreidingen van de oorspronkelijke taal. Zowel in de taal van de rekenkunde als in de taal van de Fregeaanse verzamelingenleer zijn er formules die trivialiteit uitdrukken, maar ze drukken slechts lokale trivialiteit uit. Wanneer een dergelijke formule volgt uit een kennisonderdeel dan noemen we dat kennisonderdeel lokaal triviaal.

Wat gebeurd er wanneer **APA** wordt toegevoegd aan ons kennissysteem. Vooreerst, in **APA** zijn er stellingen van de vorm $A \supset \forall x \forall y(x = y \wedge x \neq y)$. Ik heb dat tot nog toe niet vermeld omdat ik op **APA** zelf niet te diep wou ingaan. Stel nu dat, voor een **APA**-stelling van die vorm, A afleidbaar zou zijn. Dan zou uiteraard ook $\forall x \forall y(x = y \wedge x \neq y)$

[27] Jean Paul heeft weliswaar een keihard bewijs dat alle getallen interessant zijn, maar dat betekent niet dat Diderik ze interessant vindt. Zo vindt Diderik 11 beslist een oninteressant getal. Het is een priemgetal. Ja, maar zo zijn er oneindig veel. Het is ook een decimaal palindroom, net zoals 33 of 2882. Maar wat is daar interessant aan? Die eigenschap deelt 11 met onnoemelijk veel andere getallen en zelfs met 0, 1, 2 en ga maar door. Een beetje inductie en je zou gaan geloven dat alle getallen de eigenschap hebben. Ik stop, maar heb nog argumenten zat.

volgen, zodat **APA** lokaal triviaal zou worden. Met andere woorden, Die getaltheorie valt in omdat alle getallen gelijk aan elkaar blijken te zijn en, terloops gezegd, ook nog eens verschillend van elkaar en van zichzelf. Stel je voor dat je met zo'n rekenkunde een handel moet open houden. Aangezien alle papiergeld en alle munten dezelfde waarde hebben, kan je dan maar beter de munten opsparen. Tenzij je de kachel wil aanmaken uiteraard.

En de rest van ons kennissysteem? Zolang er geen getallen bij te pas komen, blijft de rest van ons kennissysteem overeind. Als je kennissysteem zegt dat die donkere wolken regen aankondigen en dat die rode bessen eetbaar en lekker zijn, dan zal je, als de donkere wolken weer daar zijn, net als vroeger best tijdig gaat schuilen, liefst met een portie van die rode bessen voor als de regen zou aanhouden. Dat je rekenkunde is ingevallen, zal je dan worst wezen, zij het helaas alleen figuurlijke.

De situatie is analoog wanneer je een AFS-theorie hanteert die triviaal zou blijken te worden. Alle verzamelingen zijn dan zowel gelijk aan elkaar als verschillend van elkaar en van zichzelf; ze zijn bovendien allemaal zowel lid als niet-lid van elkaar en van zichzelf. Met zo'n theorie valt er niets zinnig meer te zeggen over verzamelingen. Maar op de rest van onze kennis, op al wat niets over verzamelingen zegt, heeft de triviale theorie geen enkel effect.[28]

Begint er al iets te dagen? Hier lijkt een oplossing te liggen voor het probleem dat werd gesteld aan het einde van afdeling 3. De theorie **APA** en de vele AFS-theorieën kunnen triviaal worden en volstrekt onbruikbaar blijken, zonder dat de rest van onze kennis daardoor wordt aangetast. Het komt er nu alleen nog op aan een methode te vinden om uit deze voorbeelden een methode te brouwen om theorieën van deze eigenschap te voorzien.

Terloops nog iets over de klassieke negatie. Wanneer de onderliggende logica meebrengt dat de materiële implicatie voorkomt in de taal of erin definieerbaar is, dan kan je een formule die trivialiteit uitdrukt gebruiken om de klassieke negatie te definiëren. Er staat een voorbeeld in afdeling 4. Drukt de formule lokale trivialiteit uit, dan zeggen we dat de klassieke negatie *lokaal definieerbaar* is. Eigenlijk is het correcter te zeggen dat we een *lokale klassieke negatie* definiëren. Het gedefinieerde teken heeft namelijk alle eigenschappen van de klassieke negatie in de oorspronkelijke taal, bijvoorbeeld de taal van de rekenkunde, maar niet in alle uitbreidingen van deze taal.

---

[28] In de laatste drie alinea's heb ik het over veronderstellingen. Dat komt omdat **APA** en de AFS-theorieën niet lokaal triviaal kunnen worden, zoals bewezen wordt in de artikels over die theorieën. We kunnen ons echter wel afvragen wat er zou gebeuren indien die theorieën wel lokaal triviaal konden worden. Nadenken over mogelijkheden, zelfs over fictieve mogelijkheden, is belangrijk om tot nieuwe inzichten te komen.

6 Een toepassing

We willen de uitdrukking van lokale trivialiteit gebruiken om een kennisonderdeel in te passen in ons kennissysteem. Dat inpassen moet zo gebeuren dat, wanneer het onderdeel onaanvaardbare eigenschappen blijkt te hebben, wel het onderdeel wordt 'geneutraliseerd', maar niet het gehele kennissysteem. Nu zijn er veel manieren om de uitdrukking van lokale trivialiteit zo te gebruiken, maar ik zal hier slechts een van die manieren bespreken. Pogen de toegepaste algemene methode te beschrijven zou tot techniciteiten leiden die zelfs op dit punt nog lezers kunnen doen afhaken. Daarom zal ik slechts een toepassing van de methode van die ene manier bespreken. De lezer zal snel merken dat dit op zich meer dan voldoende boeiende stof tot nadenken biedt.

Laat me beginnen met een grappige vaststelling. In **APA** en in de AFS-theorieën bleken formules lokale trivialiteit uit te drukken en bleek de klassieke negatie lokaal definieerbaar. Die theorieën kunnen echter noch zichzelf noch het gehele kennissysteem 'neutraliseren' om de eenvoudige reden dat ze bewijsbaar niet-triviaal zijn.[29] De Peano rekenkunde **PA** en een aantal gangbare theorieën over verzamelingenleer, zoals **ZF**, zijn niet bewijsbaar consistent. Dit leidt tot een grappige situatie. We vonden een fenomeen in theorieën waar het aanwezig is maar 'geen nut' heeft. Maar we kunnen het fenomeen nuttig aanwenden om ervoor te zorgen dat andere theorieën slechts lokaal triviaal kunnen worden. Ik zal dit illustreren voor **PA**.

Het algemeen opzet is dat we de klassieke negatie in **PA** vervangen door "impliceert de formule die lokale trivialiteit uitdrukt". Op deze manier kunnen we de negatie verwijderen uit de taal van de rekenkunde en ze vervangen door iets wat lokaal op hetzelfde neerkomt. Hierbij duikt er echter een probleempje op. De formule die lokale trivialiteit uitdrukt in de taal van de rekenkunde is $\forall x \forall y(x = y \wedge x \neq y)$, maar in deze formule zelf komt er een negatie voor—$x \neq y$ is namelijk niets anders dan $\neg\, x = y$. Vanzelfsprekend kunnen we de negatie in deze formule niet vervangen door de formule zelf.[30] Wat we echter wel kunnen is $\neg\, x = y$ in deze formule vervangen door $x = y \supset 0 = 0'$ ($x = y$ impliceert $0 = 1$). De formule $0 = 0'$ is in elk geval vals in **PA** want $\neg\, 0 = 0'$ volgt uit het Peano axioma $\forall x \neg\, 0 = x'$. Samen met dit axioma veroorzaakt $0 = 0'$ hoe dan ook trivialiteit. De formule die trivialiteit uitdrukt vervangen we daarom door

$$\forall x \forall y(x = y \wedge (x = y \supset 0 = 0'))\,.$$

Om **PA** te combineren met de rest van onze kennis moeten we de kwantoren uit de **PA**-axioma's beperken tot getallen. De variabelen

[29] Voor de logici: absoluut bewijsbaar, namelijk met finitistische middelen.

[30] Men beweert soms dat dit gebeurt in het bewijs van het Theorema van Gödel, maar dat is louter een slechte metafoor en dus prietpraat.

van de taal van onze kennis variëren immers ook over atomen, boeken, geiten, ideeën en tuinstoelen. De rekenkunde hoort het over getallen te hebben en hoort zijn mond te houden over bronstige konijnen. Getallen hebben immers één enkele opvolger terwijl elk respectabel konijn een ontelbare schare nakomelingen heeft. Dit vergt opnieuw een kleine aanpassing van de taal van de rekenkunde: we voegen een unair predikaat N toe en gebruiken Nx om uit te drukken dat x een getal is. Welke objecten getallen zijn, kan worden vastgelegd door de axioma's $N0$ (0 is een natuurlijk getal) en $\forall x(Nx \supset Nx')$ (de opvolger van een natuurlijk getal is ook een natuurlijk getal). De Peano axioma's worden gekwalificeerd zoals men verwacht. Zo wordt het 'eerste' axioma: $\forall x(Nx \supset \neg x' = 0)$ (0 is niet de opvolger van enig natuurlijk getal). Terloops, ook de formule die trivialiteit uitdrukt wordt mee aangepast:

$$\forall x \forall y((Nx \wedge Ny) \supset (x = y \wedge (x = y \supset 0 = 0')))$$

nog steeds met de bedoeling konijnen buiten te houden.

Deze aanpassing leidt tot een merkwaardig resultaat. Alweer een merkwaardig resultaat? Inderdaad. Uit de aangepaste formule zijn *niet* alle formules van de aangepaste taal van de rekenkunde afleidbaar; $\forall x Nx$ is bijvoorbeeld niet afleidbaar en evenmin $\forall x \forall y\, x \neq y$. Het is ook maar goed dat deze en vele andere formules niet afleidbaar zijn. Indien **PA** immers triviaal zou zijn, dan zouden dergelijke formules eruit volgen. Uit $\forall x Nx$ zou volgen dat Jean Paul een getal is en uit $\forall x \forall y\, x \neq y$ zou volgen dat Jean Paul en Diderik hetzelfde zijn. Met andere woorden, als dergelijke dingen zouden volgen, zou heel ons kennissysteem meteen triviaal worden. Precies dat wilden we vermijden, toch? Wat gebeurt er dan wel wanneer de oorspronkelijke **PA** inconsistent zou blijken? Dan zou onze aangepaste **PA** 'zichzelf neutraliseren' en worden getallen een perfect onzinnige boel. Alle getallen (dingen met eigenschap N) zijn niet alleen verschillend van elkaar maar ook gelijk aan elkaar; 7 maal 4 is gelijk aan eender welk getal en is tegelijk verschillend van eender welk getal; er is geen enkel verschil meer tussen de uitkomst van een vermenigvuldiging en de uitkomst van een som, maar tegelijkertijd zijn de uitkomsten verschillend van elkaar en zelfs van zichzelf; wanneer je $<$ en $>$ definieert, dan blijkt elk getal zowel groter als kleiner dan elk ander getal, maar tegelijkertijd ook noch kleiner noch groter; en zo kan ik nog uren doorgaan. Wie in het onderwijs last had met wiskunde zal niet tot hier hebben gelezen. Dat is heel spijtig want uitgerekend voor dergelijke mensen wordt hier een paradijs geschetst waarbij de aantrekkelijkheid van de echte hemel, die met rijstpap en gouden lepeltjes, verkruimelt.

Wanneer een inconsistentie optreedt in **PA**, dan wordt de getransformeerde **PA** 'geneutraliseerd', maar blijft de rest van onze kennis overeind. Al het absurde wat afleidbaar is betreft getallen en niets anders. Dit betekent uiteraard niet dat er geen probleem is. Vanzelfsprekend hebben we een goed werkende rekenkunde nodig om de wereld te begrijpen en erin te handelen. Maar dat niet heel ons kennissysteem invalt is toch een troost, een magere troost, maar een troost.

Het is niet moeilijk zich voor te stellen hoe het analoge resultaat eruit ziet voor een verzamelingenleer waarvan we nu aannemen dat die consistent is. Ik meen echter dat de toepassing op **PA** voldoende verhelderend is.

7 Waarheen hierna?

In deze laatste afdeling zal ik geen overzicht geven van verder nuttig onderzoek over het hier behandelde onderwerp. Ik zal slechte enkele punten aansnijden die het behandelde verder kunnen verduidelijken en eventueel aansluitende vragen van de lezer beantwoorden.

Wanneer een kennisonderdeel 'geneutraliseerd' werd, bevindt het zich in een toestand die het ongeschikt maakt als vertrekpunt voor het zoeken van een consistente vervanging. Aan het einde van einde van afdeling 2 heb ik er reeds op gewezen dat de start van dat zoekproces als de start van een nieuwe fase kan worden gezien. In die nieuwe fase kan je een heel andere logica hanteren dan in de vorige of kan je een andere drastische herziening van je kennissysteem doorvoeren. De overgang naar de fase waarin een vervanging wordt gezocht, vergt immers hoe dan ook een drastische wijziging. Er moeten zoekmethoden worden toegepast en vanuit logisch oogpunt worden die tot nader order het beste gezien in termen van adaptieve logica's; maar daarover wou ik het dus niet hebben.

Niet alles wat voorheen werd afgeleid uit ons kennissysteem, moet in een volgende fase worden beschouwd als een integraal deel van het kennissysteem. We kunnen nadenken *over* ons kennissysteem, beslissen dat een onderdeel moet herzien worden, dat een ander onderdeel moet worden verwijderd, en dat andere onderdelen moeten worden behouden en gebruikt bij de herziening van het problematische onderdeel.

Welk nut heeft het dan dat een kennisonderdeel lokaal triviaal wordt? Dat nut ligt in het identificeren van het problematisch kennisonderdeel. Wanneer ons hele kennissysteem triviaal wordt, dan weten we alleen dat er ergens een probleem is. Wanneer lokale trivialiteit optreedt, dan weten we ook welk kennisonderdeel problematisch is.

Er is geen garantie dat elke inconsistentie uit ons kennissysteem kan worden verwijderd [17, §2]. Dit betekent dat het kennisonderdeel dat triviaal werd niet kan worden vervangen door een consistente theorie, maar wel door een theorie die een paraconsistente of zelfs een inconsistentie-adaptieve logica als onderliggende logica heeft. Zo betoog ik in het vermelde artikel over AFS dat het uitwerken van een Fregeaanse verzamelingenleer om meerdere redenen, zowel filosofische als wiskundige, interessant is, ook al is een dergelijke verzamelingenleer noodzakelijkerwijze inconsistent.

Laat me dit essay afsluiten met een puzzel. Veel medemensen beschouwen de (klassieke) negatie als een natuurlijke operator die bovendien een waarheidsfunctie is. De waarheidswaarde van $\neg A$ wordt immers volledig bepaald door de waarheidswaarde van A: van A en $\neg A$ is er een en slechts een waar. Toch laat de materiële implicatie toe een lokale klassieke negatie te bepalen aan de hand van een formule die lokale trivialiteit uitdrukt. Binnen de gegeven taal heeft het gedefinieerde symbool exact alle eigenschappen van de klassieke negatie. Wat bijvoorbeeld de taal van de rekenkunde betreft maakt het geen enkel verschil uit of **PA** wordt geformuleerd met de klassieke negatie als primitief symbool, wat duidelijk een operator is, of met de gedefinieerde lokale klassieke negatie. Beide formuleringen van de theorie hebben precies dezelfde verzameling stellingen en laten toe precies dezelfde rekenkundige regels af te leiden en te gebruiken in een toepassing van de rekenkunde. En toch is er een enorm verschil tussen beide formuleringen wanneer we de taal uitbreiden. Hoe kan dat? Als beloning voor de lezer die tot hier is doorgegaan, bied ik haar of hem deze puzzel aan. Het weze een hulpmiddel om tot zinnig en creatief denken te komen gedurende de verloren momenten waarmee het leven ons soms ongevraagd opscheept, zoals slapeloze nachten, tandartsbezoek en autorijden.

Verwijzingen

[1] Jean Paul Van Bendegem. *Finite, Empirical Mathematics. Outline of a Model*, volume 174 of *Werken Uitgegeven door de Faculteit van de Letteren en Wijsbegeerte.* Rijksuniversiteit Gent, 1987.

[2] Diderik Batens. The consistency of Peano Arithmetic. A defeasible perspective. In Patrick Allo and Bart Van Kerkhove, editors, *Modestly Radical or Radically Modest. Festschrift for Jean Paul van Bendegem on the Occasion of His 60th Birthday*, volume 24 (sometimes 22) of *Tributes*, pages 11–59. College Publications, London, 2014.

[3] Jean Paul Van Bendegem. Strict, yet rich finitism. In Z.W. Wolkowski, editor, *First International Symposium on Gödel's Theorems*, pages 61–79. World Scientific, Singapore, 1993.

[4] Jean Paul Van Bendegem. Strict finitism as a viable alternative in the foundations of mathematics. *Logique et Analyse*, 145:23–40, 1994. Appeared 1996.

[5] Graham Priest. Is arithmetic consistent? *Mind*, 103:337–349, 1994.

[6] Diderik Batens. Adaptive Fregean set theory. In preparation.

[7] Jean Paul Van Bendegem. *Inleiding tot de moderne logica en wetenschapsfilosofie: een terreinverkenning.* VUBPRESS, Brussel, 1991.

[8] Diderik Batens. *Logicaboek. Praktijk en theorie van het redeneren.* Garant, Antwerpen/Apeldoorn, 1992. 2: 1993; 3: 1996; 4: 1999; 5: 2002; 6: 2004; 7: 2008; 8: 2017.

[9] Bryson Brown. How to be realistic about inconsistency in science. *Studies in History and Philosophy of Science*, 21:281–294, 1990.

[10] Joke Meheus. Adaptive logic in scientific discovery: the case of Clausius. *Logique et Analyse*, 143–144:359–389, 1993. Appeared 1996.

[11] Joke Meheus. Clausius' discovery of the first two laws of thermodynamics. A paradigm of reasoning from inconsistencies. *Philosophica*, 63:89–117, 1999. Appeared 2001.

[12] Joke Meheus. Inconsistencies in scientific discovery. Clausius's remarkable derivation of Carnot's theorem. In Helge Krach, Geert Vanpaemel, and Pierre Marage, editors, *History of Modern Physics. Acta of the XXth International Congress of History of Science*, pages 143–154. Brepols, Turnhout (Belgium), 2002.

[13] Nancy Nersessian. Inconsistency, generic modeling, and conceptual change in science. In Joke Meheus, editor, *Inconsistency in Science*, pages 197–211. Kluwer, Dordrecht, 2002.

[14] John Norton. The logical inconsistency of the old quantum theory of black body radiation. *Philosophy of Science*, 54:327–350, 1987.

[15] John Norton. A paradox in Newtonian gravitation theory. *PSA 1992*, 2:421–420, 1993.

[16] Joel Smith. Inconsistency and scientific reasoning. *Studies in History and Philosophy of Science*, 19:429–445, 1988.

[17] Diderik Batens. In defence of a programme for handling inconsistencies. In Joke Meheus, editor, *Inconsistency in Science*, pages 129–150. Kluwer, Dordrecht, 2002.

[18] Peter Vickers. *Understanding inconsistent science.* Oxford University Press, Oxford, 2013.

[19] Diderik Batens. It might have been Classical Logic. *Logique et Analyse*, 218:241–279, 2012.

[20] Diderik Batens. Meaning, acceptance, and dialectics. In Joseph C. Pitt, editor, *Change and Progress in Modern Science*, pages 333–360. Reidel, Dordrecht, 1985.

[21] Diderik Batens. Do we need a hierarchical model of science? In John Earman, editor, *Inference, Explanation, and Other Frustrations. Essays in the Philosophy of Science*, pages 199–215. University of California Press, 1992.

[22] Diderik Batens. *Menselijke kennis. Pleidooi voor een bruikbare rationaliteit.* Garant, Antwerpen/Apeldoorn, 1992. 2: 2004; 3: 2008.

[23] Bryson Brown and Graham Priest. Chunk and permeate, a paraconsistent inference strategy. Part I: The infinitesimal calculus. *Journal of Philosophical Logic*, 33:379–388, 2004.

[24] Bryson Brown and Graham Priest. Chunk and permeate II: Bohr's hydrogen atom. *European Journal for Philosophy of Science*, 5(3):297–314, 2015.

[25] Diderik Batens. Pluralism in scientific problem solving. Why inconsistency is no big deal. *Humana.Mente Journal of Philosophical Studies*, 32:149–177, 2017.

[26] Diderik Batens. Looting liars masking models. In Can Başkent and Thomas Macaulay Ferguson, editors, *Graham Priest on Dialetheism and Paraconsistency.* Springer, in print.

[27] George S. Boolos, John P. Burgess, and Richard J. Jeffrey. *Computability and Logic.* Cambridge University Press, 4 edition, 2002.

[28] Diderik Batens and Kristof De Clercq. A rich paraconsistent extension of full positive logic. *Logique et Analyse*, 185–188:227–257, 2004. Appeared 2005.