



TIME-ASYMMETRY IN COUNTERFACTUAL CAUSATION THEORIES

EEN ZOEKTOCHT NAAR DE
DERDE PIJL

Naam: Floris de Vries
Studentnummer: 3536572
Opleiding: Wijsbegeerte
Bachelorscriptie (7,5 ECTS)
Begeleider: Janneke van Lith
Tweede beoordelaar: Thomas Müller
datum: 01/02/2013

Time-Asymmetry in Counterfactual Causation Theories

Een Zoektocht naar de Derde Pijl

Dankbetuiging

Deze scriptie is tot stand gekomen in samenwerking met een aantal mensen die ik kort wil noemen en bedanken. Allereerst ben ik dank verschuldigd aan Janneke van Lith voor het begeleiden van mijn versnelde scriptietraject en voor haar behulpzame feedback en suggesties. Ten tweede wil ik graag Thomas Müller bedanken, omdat hij op zeer korte termijn nog mijn tweede beoordelaar wilde worden. Vervolgens wil ik ook graag Leo van der Ven en Han van der Ven bedanken voor hun inspiratieve gesprekken met mij over 'tjdasymmetrie'. Ten slotte wil ik nog Jemaine Jacobs bedanken voor haar steun tijdens het schrijven van de scriptie, het doorlezen van eerdere versies en het verzorgen van de omslag van dit werk.

Inleiding

Één van de meest opvallende eigenschappen van causaliteit is dat oorzaken kenmerkend voorafgaan aan hun gevolgen – de causale pijl lijkt sterk gelijkgesteld te zijn aan de temporele pijl. Waarom moet dit zo zijn? De asymmetrie van verleden naar toekomst is in eerste instantie vreemd, omdat onze fundamentele natuurkunde volgens standaardopvattingen juist tijdsymmetrisch is.¹ Het is onduidelijk wat de connectie is tussen de tijdasymmetrie van causaliteit en de tijdsymmetrie van natuurkunde. Ik zal dit probleem ‘de puzzel van de tijdasymmetrie van causaliteit’ noemen.²

Het centrale idee van *counterfactual causation* theorieën is dat de betekenis van causale beweringen kan worden verklaard in termen van *counterfactual* voorwaardelijkheden in de vorm van ‘Als A niet was voorgevallen, zou C zich ook niet hebben voorgedaan’.³ Met andere woorden, het is typisch het geval voor causaal gerelateerde gebeurtenissen dat als de oorzaak zich niet had voorgedaan, het gevolg ook niet had plaatsgevonden. Uiteraard is het ook kenmerkend het geval dat als het gevolg niet had plaatsgevonden, dit zou zijn omdat de oorzaak zich niet had voorgedaan. Wat deze analyse nodig heeft om oorzaken en gevolgen van elkaar te kunnen onderscheiden is een analyse van *counterfactual* afhankelijkheid, volgens welke gevolgen *counterfactually* afhankelijk zijn van hun oorzaken, maar niet vice versa. Een dergelijke analyse kan mogelijk een antwoord bieden op de puzzel van de connectie tussen causaliteit en de tijdsymmetrie van natuurkunde. In de traditie van Price zal ik een dergelijke oplossing op de puzzel een ‘derde pijl’ noemen – de eerste twee zijn de temporele en causale pijl.⁴ De eisen, die ik van Price en Weslake zal overnemen, waaraan een derde pijl in dit geval moet voldoen, zijn als volgt: de analyse moet ten eerste duidelijk maken hoe het komt dat gevolgen afhankelijk zijn van hun oorzaken maar niet vice versa, ten tweede hoe deze asymmetrie gebaseerd is op een asymmetrische wetmatigheid of een asymmetrisch feit over onze wereld, consistent met de tijdsymmetrie van fundamentele natuurkunde, en ten derde moet het de relevantie verklaren van het feit dat we kunnen handelen voor toekomstige doelen, maar niet voor doelen in het verleden.⁵

Counterfactual causation theorieën bieden een dergelijke oplossing voor de puzzel van de tijdasymmetrie van causaliteit. Met name Lewis (1979) heeft een belangrijke poging gedaan met zijn

¹ Price, Huw and Brad Weslake. 2009. “The Time-Asymmetry of Causation”.

² Menzies, Peter, “Counterfactual Theories of Causation”, *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>>.

³ Lewis, David. 1973. “Causation”, in *The Journal of Philosophy*.

⁴ Price, Huw. 1996. *Time’s Arrow and Archimedes’ Point: New Directions for the Physics of Time*, Ch. 6.

⁵ Price, Huw and Brad Weslake. 2009. “The Time-Asymmetry of Causation”.

mogelijke werelden analyse van *counterfactuals*. Echter, een aantal filosofen, waaronder Adam Elga, zijn het niet eens met deze oplossing.⁶

Er is dus onenigheid over de houdbaarheid van de *counterfactual* causation oplossing voor het probleem. In dit paper zullen we het bijbehorende debat zo goed mogelijk beslechten aan de hand van de volgende vraagstelling: Biedt de counterfactual analyse van causaliteit een adequate oplossing voor de puzzel van de tijdasymmetrie van causaliteit? Ik zal beargumenteren dat dit niet het geval is. Allereerst zal ik uitleggen hoe dit debat is ontstaan door te kijken naar David Hume, de grondlegger van het discours over causaliteit in de moderne tijd. Daarna zal ik uitleggen hoe de counterfactual analyse van causaliteit van David Lewis (1973b, 1979), een oplossing geeft op de puzzel. Vervolgens zal ik een aantal problemen op zijn antwoord aanmerken. De eerste twee zal ik gewoon kort benoemen, het derde en belangrijkste tegenargument is afkomstig van Elga. In reactie op Elga zal ik één van de meest populaire oplossingen voor Lewis' grootste problemen aandragen: ik zal een *Past Hypothesis* in beschouwing nemen. Ik zal dit doen aan de hand van de gedeelde visie van drie auteurs: Albert, Kutach en Loewer. In verband met het doel van deze tekst zal ik de verscheidene verschillen tussen deze drie auteurs negeren en naar het voorstel van het collectief verwijzen als de 'AKL' visie, die compatibel is met Lewis' voorstel.⁷ Als deze benadering zou kloppen, zouden Elga's argumenten tegen Lewis niet meer van toepassing zijn. Ten slotte zal ik een aantal problemen betreffende de AKL benadering van de *Past Hypothesis* bespreken. Dit zal de weerlegging op Elga tenietdoen, waardoor Elga's argument tegen Lewis weer geldt.

⁶ Callender, Craig, "Thermodynamic Asymmetry in Time", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2011/entries/time-thermo/>>.

⁷ Price, Huw and Brad Weslake. 2009. "The Time-Asymmetry of Causation".

Hume's Semantische Conventionalisme

Laten we beginnen met kijken naar de grondlegger van de filosofie van causaliteit in de moderne tijd, David Hume. Vroeg in zijn *Treatise*, biedt Hume de volgende definitie van 'de relatie tussen oorzaak en gevolg' aan:

*We may define a CAUSE to be 'An object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedency and contiguity to those objects that resemble the latter.'*⁸

Dit voorstel maakt het een kwestie van definitie dat oorzaken voorgaan aan hun gevolgen. Hume neemt aan dat de kern van de causale relatie de symmetrische noties van nabijheid en regelmaat zijn, door 'oorzaak' en 'gevolg' te labelen als de eerdere en latere van een paar van de juiste gerelateerde gebeurtenissen. Als Hume gelijk heeft, dan is de relatie tussen de causale pijl en de temporele pijl slechts een kwestie van semantische conventie.⁹

Hume's voorstel heeft een aantal evidente aantrekkelijkheden. Het impliceert dat er geen apart probleem is over de causale asymmetrie, omdat het gewoon een indirecte manier is om naar de temporele asymmetrie te verwijzen. Maar desondanks de economische voordelen, zijn er twee belangrijke voornaamste bezwaren tegen Hume's visie. De eerste is dat de verbinding tussen causale asymmetrie en temporele asymmetrie te strak is. Veel filosofen hadden het idee dat het een interessante kwestie is of er gevallen van simultane causaliteit zijn, of zouden kunnen zijn, waarbij de oorzaak op hetzelfde moment plaatsvindt als het gevolg; of zelfs achterwaartse causaliteit, waarbij de oorzaak zich later voordoet dan het gevolg.¹⁰ Hume's visie verandert deze kwesties in conceptuele verwarringen. David Lewis zegt het als volgt:

*Careful readers have thought they could make sense of stories of time travel . . . ; hard-headed psychical researchers have believed in precognition; speculative physicists have given serious consideration to tachyons, advanced potentials, and cosmological models models with closed timelike curves. . . It will not do to declare [these phenomena] impossible a priori.*¹¹

⁸ Hume, David. *Treatise*, Bk. 1, Part III, sec. XIV

⁹ Morris, William Edward, "David Hume", *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), forthcoming URL = <<http://plato.stanford.edu/archives/spr2013/entries/hume/>>.

¹⁰ Faye, Jan, "Backward Causation", *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2010/entries/causation-backwards/>>.

¹¹ Lewis, David. 1986a. *Philosophical Papers*: p. 40–41

Als we de intuïtie delen dat achterwaartse en simultane causaliteit niet klaarblijkelijk absurd zijn, moeten we – volgens de *a priori* opvatting van conceptuele analyse – Hume’s visie verwerpen, op zijn minst in zijn simpele vorm.¹²

De tweede moeilijkheid bij Hume’s visie is dat de verbinding te zwak is, in de volgende zin: Causaliteit lijkt verbonden te zijn aan beraadslaging. De temporele asymmetrie van causaliteit lijkt iets te maken te hebben met het feit dat het onzin is om iets te beraden met doelen in het verleden in zicht. Hume’s voorstel verklaart dit feit niet. Om dit te zien, moeten we ons voorstellen dat we een ticket hebben in een loterij die gisteren is getrokken. De resultaten zijn nog niet bekendgemaakt, en we hopen dat we hebben gewonnen. Waarom lijkt het zo absurd om nu iets te proberen te doen om te verzekeren, of om het waarschijnlijker te maken, dat ons ticket een paar uur geleden uit de ton was getrokken? Als Hume gelijk heeft, is het geen goed antwoord om te zeggen dat de uitkomst van de loterij een gevolg van een huidige actie, omdat de trekking heeft plaatsgevonden in het verleden. Want volgens Hume’s visie komt dit gewoon neer op het herhalen van de bewering die we probeerden te verklaren, namelijk, dat we handelen voor latere doelen (gevolgen), maar niet voor eerdere doelen (oorzaken). Als er een huidige actie zou zijn die ons succes in de trekking van gisteren zou garanderen, waarom zouden we er dan om geven of we er naar behoren van kunnen zeggen dat die handeling het veroorzaakt heeft?¹³

De beperkingen van Hume’s visie brengen dus twee algemene verlangens aan het licht voor een adequate theorie over de tijdasymmetrie van causaliteit. Het zou het feit moeten verklaren dat de causale pijl kenmerkend – maar wellicht niet noodzakelijk – gelijkgesteld is aan de temporele pijl. En het zou ons moeten helpen om een aangelegenheid van groot praktisch belang in onze levens zinvol te maken, het feit dat we kunnen handelen voor toekomstige doelen, maar niet voor doelen in het verleden (tenminste onder normale omstandigheden).

¹² Frisch, Mathias. “Time and Causation”.

¹³ Frisch, Mathias. “Time and Causation”.

Lewis' Derde Pijl

Wat bepaalt de richting van de causale relatie? In antwoord op deze vraag, beargumenteert Lewis dat de richting van causaliteit, de richting is van causale afhankelijkheid; en het is typisch waar dat gebeurtenissen causaal afhankelijk zijn van eerdere gebeurtenissen, maar niet van latere gebeurtenissen. Hij benadrukt de contingentie van het tweede feit, omdat hij achterwaartse causaliteit als een conceptuele mogelijkheid ziet die we niet a priori kunnen uitsluiten. Hiermee overeenkomstig, wijst hij iedere analyse van *counterfactuals* af die de temporele asymmetrie door middel van conceptuele autorisatie zou leveren.¹⁴

Waarom is deze richting kenmerkend gelijkgesteld aan de temporele richting van verleden naar toekomst? Lewis' verklaring van de temporele asymmetrie van *counterfactual* afhankelijkheid is gebaseerd op een *de facto* asymmetrie over de werkelijke wereld. Hij definieert een *determinant* voor een gebeurtenis als een set van voorwaarden die gezamenlijk voldoende zijn, gegeven de natuurwetten, voor het voorkomen van de gebeurtenis. (Determinanten van een gebeurtenis kunnen oorzaken of sporen zijn van de gebeurtenis.) Hij beweert dat het contingent waar is dat gebeurtenissen kenmerkend erg weinig eerdere determinanten hebben, maar veel latere determinanten. Ter illustratie citeert hij Poppers voorbeeld van een sferisch golffront die zich naar buiten uitbreidt vanuit een puntbron.¹⁵ Dit is een proces waar ieder stukje van de golf postdetermineert wat er gebeurt op het punt waarop de golf wordt uitgestoten. Hij zegt dat het achterwaartse proces, waarin een sferische golf samentrekt en ieder stukje golf predetermineert wat er gebeurt op het punt waarop de golf wordt opgenomen, ook de natuurwetten zou gehoorzamen, maar dat dit proces zelden plaatsvindt. Er zijn dus veel latere determinanten en weinig eerdere determinanten.¹⁶

Lewis combineert deze *de facto* asymmetrie van overdeterminatie met zijn analyse van de *comparative similarity relation* tussen verschillende mogelijke werelden.¹⁷ Volgens deze analyse, zijn er verscheidene opzichten van gelijkheid die we in overweging moeten nemen als we niet-backtracking *counterfactuals*¹⁸ evalueren: gelijkheid met betrekking tot natuurwetten en ook gelijkheid met betrekking tot specifieke feiten. Werelden zijn gelijk aan de werkelijke wereld

¹⁴ Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow".

¹⁵ Popper, K. (1956): "The Arrow of Time", *Nature*, 1777: 538.

¹⁶ Menzies, Peter, "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>>.

¹⁷ Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow".

¹⁸ Bij niet-backtracking counterfactuals is het verleden doorgaans vastgesteld tot het tijdstip waarop we het counterfactual antecedent horen te verkrijgen.

naarmate ze minder wonderen of schendingen van de werkelijke natuurwetten bevatten. Ook zijn werelden gelijk aan de werkelijke wereld naarmate het spatio-temporele gebied van perfecte overeenkomst van bepaalde feiten met de werkelijke wereld groter is. Als de werkelijke wereld wordt bepaald door deterministische wetten, zullen deze regels botsen bij het beoordelen van welke *counterfactual* werelden gelijk zijn aan de werkelijke wereld. Een wereld die een *counterfactual* antecedent waar maakt moet verschillen van de werkelijke wereld, ofwel in het toestaan van enige schendingen van de werkelijke wetten, ofwel ze verschilt van de werkelijke wereld wat betreft bepaalde feiten. Lewis' analyse staat een afweging toe tussen deze concurrerende opzichten van gelijkheid in zulke gevallen. Het impliceert dat werelden met een uitgebreid gebied van perfecte overeenkomst van bepaalde feiten kunnen worden beschouwd als zeer gelijk aan de werkelijke wereld, mits de overeenkomst van bepaalde feiten met de werkelijke wereld worden bereikt ten koste van een klein, plaatselijk wonder, maar niet ten koste van een groot wonder. Op zichzelf bevat deze gedachte geen ingebouwde tijdasymmetrie. Dat komt slechts wanneer ze gecombineerd wordt met de asymmetrie van overdeterminatie.¹⁹

Dus, de symmetrische analyse van gelijkheid, gecombineerd met de *de facto* asymmetrie van overdeterminatie impliceert dat werelden die plaats hebben voor *counterfactual* veranderingen, terwijl ze het werkelijke verleden behouden en divergente wonderen toestaan, gelijk zijn aan de werkelijke wereld dan werelden die plaats hebben voor dergelijke veranderingen die, door convergente wonderen toe te staan, de werkelijke toekomst behouden. Dit feit impliceert dat, waar de asymmetrie van overdeterminatie wordt verworven, het heden *counterfactually* afhankelijk is van het verleden, maar niet van de toekomst.

¹⁹ Menzies, Peter, "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>>.

Kritiek op Lewis

Één soort kritiek richt zich op de psychologische ongeloofwaardigheid van Lewis' verklaring.²⁰ De verklaring beroept zich aan de ene kant op een systeem van afgewogen opzichten van gelijkheid tussen mogelijke werelden, dat wordt geleverd door *a priori* conceptuele analyse en, aan de andere kant, op een asymmetrie van overdeterminatie, waarvan wordt beweerd dat het een contingente *a posteriori* waarheid over de werkelijke wereld is. De tweedelige verklaring zou de feiten moeten gebruiken die voldoende bekend zijn om een rol te spelen in de uitleg van ons linguïstieke gebruik van *counterfactuals*. Echter, het is psychologisch implausibel dat het ingewikkelde systeem van afgewogen opzichten van gelijkenis, met vergelijkingen tussen wonderen van verschillende groottes, de intuïtieve gelijkheids-relatie zou kunnen bevangen die gebruikt wordt in *counterfactual* redeneren. Waarom zouden we een dergelijke notie van gelijkenis moeten hebben ontwikkeld? Bovendien, de asymmetrie van overdeterminatie is een esoterische wetenschappelijke hypothese die niet algemeen bekend is bij iedereen die *counterfactuals* gebruikt. Dus het is zeer onwaarschijnlijk dat deze hypothese zou kunnen gelden voor de beheersing van de gewone spreker over de temporele asymmetrie van *counterfactuals*.²¹

Een ander kritiekpunt is dat de asymmetrie van overdeterminatie niet bestaat op de manier die nodig is om Lewis' verklaring van de temporele asymmetrie van *counterfactuals* te ondersteunen. Lewis' idee is dat iedere gebeurtenis *e* vele post-determinanten heeft en weinig pre-determinanten, waarbij een pre-determinant of post-determinant van een gebeurtenis een set van voorwaarden is die gezamenlijk voldoende zijn, gegeven de natuurwetten, voor het voorkomen van de gebeurtenis. Maar als Lewis aanneemt dat de betrokken wetten zijn zoals die van klassieke mechanica, vergist hij zich op dit punt. Want een theorie die tijd symmetrisch is en deterministisch in zowel de voorwaartse als achterwaartse richting, impliceert dat voor iedere plaatselijke gebeurtenis *e* en iedere tijd *t*, er een unieke set van voorwaarden is op *t* die noodzakelijk en voldoende zijn, gegeven de wetten, voor het voorkomen van gebeurtenis *e*. De voorwaarden zijn wellicht geen gelokaliseerde voorwaarden die kenmerkend beschouwd worden als gebeurtenissen, maar ze gelden desalniettemin pre-determinant of post-determinant. Bijvoorbeeld, neem Poppers voorbeeld van de golf die zich uitspreidt vanuit een puntbron. Als er een proces is dat postdetermineert wat er gebeurt op het punt waarop de golf wordt uitgestoten, is er ook een proces, misschien een zeer niet-gelokaliseerd proces,

²⁰ Horwich, P. (1987): *Asymmetries in Time*.

²¹ Menzies, Peter, "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>>.

dat predetermineert wat er gebeurt. Volgens Popper en Lewis, zijn beide processen even waarschijnlijk; en of ze voorkomen hangt af van de randvoorwaarden van het systeem.²²

Een gerelateerd kritiekpunt betreft de asymmetrie van wonderen die centraal staat bij Lewis' gedachte over de temporele asymmetrie van causaliteit. De asymmetrie van wonderen bestaat in het feit dat een wonder dat een *counterfactual* antecedent over bepaalde feiten op tijd *t* realiseert door een mogelijke wereld te hebben die afwijkt van de werkelijke wereld, vlak voordat de tijd *t* kleiner en minder divers is dan een wonder dat hetzelfde *counterfactual* antecedent realiseert en een mogelijke wereld laat samenlopen met de werkelijke wereld na tijd *t*. Adam Elga (2000) heeft beargumenteerd dat de asymmetrie van wonderen niet op gaat in veel gevallen.²³

Elga legt zijn argument uit aan de hand van een voorbeeld: Gretta breekt een ei in een hete pan om 8:00 uur 's morgens en om 8:05 uur is het ei gebakken. Beschouw het proces dat plaatsvindt in de periode van 8:00 uur tot 8:05 uur, spoel het achteruit in de tijd: een gekookt ei zit in de pan; het verenigt zich tot een rauw ei en springt omhoog; en een eierschaal sluit zich om het ei heen. De wetten van de thermodynamica staan toe dat dit proces fysisch mogelijk is, maar het is extreem zeldzaam. Deze wetten vermelden ook dat het proces zeer gevoelig is in zijn initiële voorwaarden: zelfs de kleinste veranderingen in de moleculen die de toestand van het gekookte ei bepalen, zou er in resulteren dat het proces zich op een zodanige manier ontvouwt dat het gekookte ei in de pan blijft zitten, in plaats van dat het samensmelt tot een rauw ei en omhoog springt. Maar dit is, volgens Elga, precies de soort verandering die zou zorgen voor een 'convergent wonder'. Neem de toestand van de werkelijke wereld op 8:05 uur, de vastgelegde toekomst van die toestand na dit punt; breng een aantal kleine veranderingen aan de moleculen toe, die het in deze toestand brengen; en laat dan de wetten van de thermodynamica achterwaarts in de tijd gaan, en we zullen bijna zeker aankomen bij een toestand waarin het ei in de pan kouder wordt. Deze toestand zal er één zijn waarin Gretta het ei niet breekt. De kleine veranderingen in de toestand van de werkelijke wereld op 8:05 uur is een 'convergent wonder' dat een mogelijke wereld oplevert die de *counterfactual* propositie realiseert dat Gretta het ei niet breekt op 8:00 uur terwijl we de vastgelegde werkelijke toekomst behouden na 8:05 uur. Maar dit wonder is niet het grote, diverse wonder wat volgens Lewis een convergent wonder in zou moeten houden.²⁴

²² Ibid.

²³ Elga, Adam. 2000. "Statistical Mechanics and the Asymmetry of Counterfactual Dependence".

²⁴ Ibid.

Kritiek op Elga

Elga's argument buit een zeer fundamentele eigenschap van een algemeen geaccepteerde statistische verklaring van de tweede wet van de thermodynamica uit. Deze verklaring is afkomstig van Ludwig Boltzmann. Boltzmann's uitleg combineert twee belangrijke ingrediënten. De eerste is een statistische overweging: voor iedere macrotoestand van een fysisch systeem dat zich niet al in een thermodynamisch equilibrium bevindt, zijn er veel meer microtoestanden verenigbaar met die macrotoestand waarvan de evolutie naar een equilibrium zou leiden, dan microtoestanden die weg van een equilibrium zouden leiden. Dit lijkt wellicht voldoende om het feit te verklaren dat in onze ervaring geïsoleerde systemen naar een equilibrium toegaan.²⁵

Echter, de fout in deze redenering werd door Boltzmann's leraar en collega, Josef Loschmidt, naar voren gebracht.²⁶ De statistische overwegingen zijn tijdsymmetrisch. Als ze impliceren dat entropie – een maat van wanorde in een systeem, of liever de waarschijnlijkheid – toeneemt in de richting van de toekomst, dan impliceren ze ook dat entropie toeneemt naar het verleden: tijd symmetrische statistieken kunnen niet de symmetrie doorbreken om de monotone toename van entropie die we daadwerkelijk waarnemen te verklaren. Om te verkaren wat we observeren, moeten we Boltzmann's statistieken aanvullen met een tweede aanname, een tijd symmetrische randvoorwaarde. We moeten aannemen dat het geobserveerde universum is begonnen in een extreem lage toestand van entropie op een punt in het nabije verleden. Albert (2001) noemt deze assumptie de *Past Hypothesis*.²⁷

Loschmidt's punt impliceert dat de werkelijke microtoestand van ons vertrouwde universum altijd 'speciaal' is in de volgende zin: het overgrote merendeel van microtoestanden die verenigbaar zijn met de werkelijke macrotoestand, worden geassocieerd met geschiedenissen die zeer verschillen van die van de werkelijke wereld – geschiedenissen waar entropie toeneemt naar het verleden, in plaats van dat het afneemt naar het verleden. Zoals Elga duidelijk maakt, betekent dit dat er eigenlijk een enorme overvloed is aan microscopische wonderen, die precies de gevallen verstrekt die Lewis' asymmetrie van overdeterminatie zou moeten uitsluiten: werelden die in een punt samenkomen vanuit zeer verschillende geschiedenissen, verschillen van de werkelijke wereld door een klein lokaal

²⁵ Callender, Craig, "Thermodynamic Asymmetry in Time", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2011/entries/time-thermo/>>.

²⁶ Voor degenen die zich afvragen hoe Elga en Loschmidt zich tot elkaar verhouden: Elga's argument berust op Loschmidt's theorie over anti-thermodynamische werelden.

²⁷ Ibid.

wonder. Zonder de restrictie die de *Past Hypothesis* oplegt, zou de asymmetrie van overdeterminatie met andere woorden op grote schaal falen.²⁸

Elga's argument suggereert dat voor zover er een objectieve fysische asymmetrie is van de soort die Lewis nam om de asymmetrie van *counterfactual* afhankelijkheid op te funderen, deze gaat over macroscopische, statistische fenomenen, van dezelfde soort als gewone verschijnselen van de thermodynamische asymmetrie; in het bijzonder afhankelijk van dezelfde initiële voorwaarden. Het is inderdaad verleidelijk om deze fenomenen in het algemeen te typeren als voorbeelden van de verspreiding van precies de soort macroscopische concentraties van energie die worden geproduceerd door de *Past Hypothesis*. Voor zover Lewis' intuïties ons leiden in de richting van een echte fysische asymmetrie – een mogelijke kandidaat voor een derde pijl – lijkt dit de beste richting te zijn.

De meest uitdrukkelijke poging om de asymmetrie van causaliteit en *counterfactual* afhankelijkheid te verbinden met die van de thermodynamica komt van werken van Albert (2001), Kutach (2001, 2002, 2007) en Loewer (2007). Het AKL voorstel probeert de *Past Hypothesis* te gebruiken om de asymmetrie van *counterfactual* afhankelijkheid te verklaren. Het basale idee is om te beargumenteren dat krachtens de *Past Hypothesis*, kleine lokale veranderingen – de soort dingen die we zouden kunnen gebruiken als 'causal handles', zoals Albert beweert – veel grotere en meer diverse veranderingen in de toekomst produceren, dan in het verleden.²⁹ Voor ons gevoel zou *Past Hypothesis* moeten verzekeren dat als we in het heden een *causal handle* (counterfactueel) wiebelen, we overeenkomstige bewegingen in de toekomst produceren, maar niet in het verleden – of op zijn minst, niet in het macroscopische, zichtbare verleden. Loewer verklaart dit idee aan de hand van het figuur van een boom die zich vertakt naar de toekomst, maar beperkt is aan één stam in het verleden. De eerste geloofwaardigheid van dit idee kunnen we gemakkelijk zien door Elga's bezwaar tegen Lewis' asymmetrie van overdeterminatie te herinneren. Elga's bewijs dat convergentie naar de huidige wereld eigenlijk zeer makkelijk is, berust op Loschmidt's anti-thermodynamische werelden – met andere woorden, werelden zonder *Past Hypothesis*.

²⁸ Ibid.

²⁹ Albert, David Z. 2001. *Time and Chance*, p. 128.

Een Web en geen Boom?

Het AKL voorstel is sterk bekritiseerd door Frisch.³⁰ Hij daagt in het bijzonder de bewering uit dat *Past Hypothesis* de benodigde boomstructuur levert. In veel gevallen, beargumenteert hij, lijkt de werkelijke structuur meer op een web dan op een boom. Met andere woorden, de structuur bevat divergentie naar zowel het verleden als de toekomst – wat volgens AKL zou impliceren dat kleine, lokale veranderingen, macroscopische veranderingen in het verleden én de toekomst zou kunnen produceren. Frisch overweegt bijvoorbeeld een gas in een container met twee kamers, die eerst in één van de twee toestanden van lage entropie was: al het gas was in de linker kamer, of al het gas was in de rechter kamer. Nadat het gas is verspreid over de twee kamers, kan het zomaar het geval zijn dat slechts minuscule, lokale veranderingen de microtoestanden scheiden die uit de twee verschillende eerste voorwaarden zijn geëvolueerd. In dit geval, lijkt de AKL benadering te impliceren dat een minuscule huidige verandering zou kunnen veroorzaken dat het gas in de ene kamer had gezeten, in plaats van in de andere. (Zoals Frisch benadrukt, thermodynamica impliceert dat een dergelijk geval waarschijnlijk zeer gewoon is, omdat het simpelweg een gevolg is van gelijktrekking.) Frisch merkt ook op dat, zelfs als we dergelijke grove tegenvoorbeelden opzij zetten, de AKL benadering onbevredigend lijkt. De intuïtieve asymmetrie van causaliteit en *counterfactual* redeneren lijkt scherper, algemener, en veel minder gevoelig voor het micro-macro onderscheid, dan het AKL voorstel mogelijk voor zijn rekening kan nemen.³¹

³⁰ Frisch, 2005a, 2007

³¹ Price, Huw and Brad Weslake. 2009. "The Time-Asymmetry of Causation".

Future Hypothesis

Een ander bezwaar tegen de AKL benadering rust op de waarneming dat als het waar zou zijn dat de *Past Hypothesis* (in combinatie met de eerder vastgestelde tijd symmetrische bronnen) voldoende zou zijn om ons onvermogen om het verleden te beïnvloeden te verklaren, dat dan – door symmetrisch te redeneren – een analoge lage entropie randvoorwaarde in de toekomst ons er van zou weerhouden om de toekomst te beïnvloeden. Maar zou een *Future Hypothesis* dit gevolg hebben? Ik denk van niet. In deze sectie zullen we zien dat AKL niet in staat is om de tijdasymmetrie van beraadslaging te verklaren en de tekortkoming van Hume's voorstel deelt.³²

De eerste vraag is of een dergelijke toekomst-beperking zou impliceren dat onze weloverwogen fenomenologie een toekomstgericht analogon zou zijn van wat we proberen te verklaren met betrekking tot het verleden: de schijnbare absurditeit, tenminste in gewoonlijke gevallen, van zodanig handelen dat je het verleden beïnvloedt. Het is moeilijk om te zien waarom dit zo zou moeten zijn. Beperkingen in de verre toekomst – zelfs extreme beperkingen, veel strakker dan de *Past Hypothesis* zelf – lijken vrijwel geen invloed te hebben op onze huidige intuïtie dat we de toekomst kunnen beïnvloeden. Veronderstel dat God ons vertelt dat als wet, de uiteindelijke toestand, over zo een 15 biljoen jaar van nu, beperkt zal worden binnen een minuscule deel van de faseruimte (vergelijkbaar qua formaat met wat we nodig hebben voor *Past Hypothesis*). Beter nog, veronderstel dat hij ons aanbiedt om de werkelijke uiteindelijke microtoestand te vertellen. Hoe dan ook, de AKL boom van mogelijke trajecten lijdt aan dezelfde soort beperking richting de toekomst als die *Past Hypothesis* nodig heeft in de richting van het verleden. Vervallen we in fatalisme, als we bedenken dat het absurd is om te proberen onze nabije toekomst te beïnvloeden? Het is moeilijk om te zien waarom dat het geval zou (moeten) zijn. Vanwege de symmetrie is het moeilijk om te zien waarom een verafgelegen *Past Hypothesis* onvereenigbaar zou moeten zijn met het in staat achten van onszelf om het nabije verleden te kunnen beïnvloeden.³³

Hier kun je wellicht tegen inbrengen dat dit argument te veel draait om het feit dat het slechts een verafgelegen toekomst-beperking betreft. Laten we, zonder dat we de meest voor de hand liggende reactie dat *Past Hypothesis* ook behoorlijk ver weg is zullen gebruiken, ons richten op de gevolgen van toekomst-beperkingen die dichterbij liggen. Zouden deze gevolgen noodzakelijk worden waargenomen als dat ze deliberatie absurd maken? Daarentegen denken we dat ze wellicht voorzien

³² Price, Huw and Brad Weslake. 2009. "The Time-Asymmetry of Causation".

³³ Ibid.

van een nieuwe graad van controle, een invloed over zaken waarvan we eerder dachten dat ze onafhankelijk waren van onze acties.³⁴

Om een oud voorbeeld uit de beslissingstheorie literatuur te gebruiken, veronderstel dat we bestemd zijn om Dood te ontmoeten in de middag van een bepaalde dag. In het gedachte-experiment is het het geval dat deze wet geldt.³⁵ Het is nu 09:05 uur van de noodlottige ochtend en we begeven ons op het vliegveld van Aleppo, met een boarding pass voor de vlucht naar Damascus. We weten dat Dood ons zal treffen in de ene plek of de andere; en daarbij dat hij al onderweg is naar welke plek het zal worden. Is het absurd om te denken dat we nog steeds vrij zijn om te kiezen of we in het vliegtuig stappen of niet? Overduidelijk is het tegendeel waar. Hoewel de randvoorwaarde ons zeker berooft van veel opties – the mogelijkheid om in de middag ergens anders te zijn dan Damascus of Aleppo, bijvoorbeeld, of om überhaupt ergens te zijn, later op de dag – brengt het ook een aantal mogelijkheden: in het bijzonder, de mogelijkheid om Dood's bewegingen te beïnvloeden, zelfs wat eerder op de dag in kwestie.

Het voorbeeld suggereert dat, hoewel een wetmatige toekomst-beperking de mogelijkheden kan limiteren, het niet absurd is om te denken de we controle kunnen uitoefenen binnen deze limieten. Binnen deze limieten, lijkt het effect ons niet te weerhouden van het bereiken van onze doelen, maar ons er van te verzekeren dat de wereld samenspannt om deze doelen tot stand te brengen. Dit is iets compleet anders dan ons weerhouden om de doelen te bereiken. Het geeft ons een nieuwe soort van controle over andere gebeurtenissen – de gebeurtenissen die op de juiste wijze gerangschikt moeten worden, in het licht van de nieuwe beperking, om onze doelen te geschieden. Dit betekent, in het bijzonder, dat we wellicht het verafgelegen heden kunnen beïnvloeden, en het verleden, via een soort van zigzag. We kiezen de toekomst in sommige opzichten, en de toekomst-beperking verzekert dat het verafgelegen heden en verleden synchroon blijven, om de benodigde uiteindelijke toestand te bereiken. Als dit is hoe dingen zouden gaan in een wereld met wetmatige toekomst randvoorwaarden, zou de *Past Hypothesis* dan niet hetzelfde soort effect moeten hebben?³⁶ Zou het niet slechts ons vermogen om het verleden te beïnvloeden moeten beperken, en dit compenseren door ons nieuwe krachten te geven – krachten om bijvoorbeeld het verafgelegen heden te beïnvloeden, door stukjes van het verleden te beïnvloeden waar de antecedenten van het verafgelegen heden noodzakelijk mee verbonden zijn? Deze mogelijkheid is blijkbaar gemist, omdat AKL heeft gefaald om een ambiguïteit te zien in de vereiste dat we de gevolgen van kleine, lokale veranderingen overwegen – *causal handles*, om Albert's term te gebruiken. De vereiste dat de

³⁴ Ibid.

³⁵ Gibbard, Alan, and William Harper. 1978. "Counterfactuals and Two Kinds of Expected Utility", p. 136

³⁶ De *Past Hypothesis* wordt hier beschouwd als een wetmatige conditie.

handles lokaal zijn, is nodig om een triviale vervalsing van de theorie te voorkomen, omdat in de veronderstelde context van een deterministische theorie, grootschalige verschillen onmiddellijk een verschil zullen maken op zowel eerdere als latere tijdstippen. Maar deze beperking tot kleine, lokale *handles* moeten we niet laten impliceren dat de gevolgen van het wiebelen van de *handles* niet tegelijk kan – anders zouden we gelijktijdige causaliteit door middel van volmacht uitsluiten.³⁷

Deze overwegingen kunnen we op twee manieren uitspelen. Ten eerste suggereren ze dat de gevolgen van wetmatige toekomstbeperkingen anders zouden zijn dan een toekomstgericht analogon van wat we proberen te verklaren ten opzichte van het verleden: de klaarblijkelijke absurditeit, op zijn minst in gewone gevallen, van zo handelen dat je het verleden beïnvloedt. Zoals we hebben gezien, geven verafgelegen beperkingen weinig reden tot fatalisme, en hoewel onmiddellijke beperkingen zeker onze keuzes zouden beperken, zouden ze ons ook nieuwe mogelijkheden bieden.³⁸

Ten tweede, het argument suggereert dat microscopische effecten op het verre verleden – wat AKL toestaat als een gevolg van hun visie – niet verhinderd kunnen worden van vergroting naar minder microscopische effecten op het minder verre verleden, en het verafgelegen heden, door middel van een zigzag.³⁹ De motor van de tweede fase van dit proces – de ‘zag’ waarmee de invloed van een huidige actie terugkomt uit het verre verleden – zal het proces van versterking van kleine verschillen zijn, die centraal staat in het voorstel van de gedachte over macroscopische vertakking. Veronderstel dat het waar is (zoals AKL toestaat) dat, als ik mijn vinger een moment geleden had op gelift, er verschillen waren geweest in de posities van een aantal atomen, biljoenen jaren in het verleden. Wat voor veranderingen zou de beweging van deze oeroude atomen hebben vervaardigd over zo een grote tijdsperiode? Zeker geen veranderingen groot genoeg voor de verwijdering van mij en mijn vinger, want ik ben nu hier, door bepaling, in de geschiedenis in kwestie. Maar er is geen dergelijke bescherming voor de rest van mijn vertrouwde universum, overal binnen de toekomstige lichtkegel van deze oeroude microscopische veranderingen.⁴⁰

We kunnen concluderen dat de AKL benadering onvoldoende verklaring geeft van de asymmetrie van beraadslaging die we eerder bij Hume zagen. Daarnaast suggereert het zojuist beschreven argument een krachtig bezwaar tegen iedere poging om de tijdasymmetrie van causaliteit te funderen op de macroscopische statistische asymmetrieën die we in onze wereld kunnen vinden. Zoals ik reeds opmerkte, lijkt het zeer geloofwaardig dat deze asymmetrieën hun oorsprong vinden bij de *Past*

³⁷ Price, Huw and Brad Weslake. 2009. “The Time-Asymmetry of Causation”.

³⁸ Ibid.

³⁹ Kutach, Douglas. 2002. “The Entropy Theory of Counterfactuals”.

⁴⁰ Price, Huw and Brad Weslake. 2009. “The Time-Asymmetry of Causation”.

Hypothesis. Maar we hebben zojuist beargumenteert dat, aangezien een *Future Hypothesis* het niet absurd zou maken om te beraden over toekomstige doelen, de *Past Hypothesis* niet kan verklaren waarom we niet overleggen voor doelen in het verleden. Dus iedere theorie over de causale pijl die ernaar op zoek is om de tijdasymmetrie van causaliteit te reduceren tot de soort asymmetrie die voortkomt uit de *Past Hypothesis*, lijkt eveneens voorbestemd machteloos te zijn om de tijdasymmetrie van beraadslaging te verklaren – voorbestemd, met andere woorden, om in dit opzicht de tekortkoming van Hume's voorstel te delen.

Conclusie

Biedt de counterfactual analyse van causaliteit een adequate oplossing voor de puzzel van de tijdasymmetrie van causaliteit? Ik heb beargumenteerd dat dit niet het geval is. Allereerst heb ik meer verteld over de oorsprong van het debat: De beperkingen van Hume's visie brachten twee algemene verlangens aan het licht voor een geschikte analyse van tijdasymmetrie van causaliteit. Een dergelijke analyse moet kunnen verklaren dat de causale pijl kenmerkend – maar wellicht niet noodzakelijk – gelijkgesteld is aan de temporele pijl. En het zou ons moeten helpen om het feit te verklaren dat we kunnen handelen voor toekomstige doelen, maar niet voor doelen in het verleden. Daarna heb ik uitgelegd hoe de *counterfactual* analyse van causaliteit van David Lewis (1973b, 1979), een oplossing geeft op de puzzel: de symmetrische analyse van gelijkheid, gecombineerd met de *de facto* asymmetrie van overdeterminatie impliceert dat werelden die plaats hebben voor *counterfactual* veranderingen, terwijl ze het werkelijke verleden behouden en divergente wonderen toestaan, gelijk zijn aan de werkelijke wereld dan werelden die plaats hebben voor dergelijke veranderingen die, door convergente wonderen toe te staan, de werkelijke toekomst behouden. Dit feit impliceert dat, waar de asymmetrie van overdeterminatie wordt verworven, het heden *counterfactually* afhankelijk is van het verleden, maar niet van de toekomst. Vervolgens heb ik een aantal problemen van zijn antwoord bespreken: de psychologische ongeloofwaardigheid van Lewis' verklaring, een kritiekpunt over dat de asymmetrie van overdeterminatie niet bestaat op de manier die nodig is om Lewis' verklaring van de temporele asymmetrie van *counterfactuals* te ondersteunen en als belangrijkste tegenargument heb ik Adam Elga's redenering besproken waarin hij Lewis' asymmetrie van wonderen bekritiseert. Om Elga's argument te weerleggen heb ik een Past Hypothesis in beschouwing genomen. Ten slotte heb ik een aantal problemen van de Past Hypothesis besproken. We kunnen concluderen dat de AKL benadering onvoldoende verklaring heeft gegeven van de asymmetrie van beraadslaging die we eerder bij Hume zagen. Kortom, de *counterfactual* analyse van causaliteit biedt geen adequate oplossing op de puzzel van de tijdasymmetrie van causaliteit. Sterker nog, iedere theorie over de causale pijl die ernaar op zoek is om de tijdasymmetrie van causaliteit te reduceren tot de soort asymmetrie die voortkomt uit de *Past Hypothesis*, lijkt in dit opzicht de tekortkoming van Hume's voorstel te delen.

Literatuur

Albert, David Z. 2001. *Time and Chance*, Harvard University Press, Cambridge MA.

Callender, Craig, "Thermodynamic Asymmetry in Time", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/time-thermo/>, bekeken 21-1-2013.

Elga, Adam. 2000. "Statistical Mechanics and the Asymmetry of Counterfactual Dependence", in *Philosophy of Science*, Vol. 68, No. 3, September 2000, pp. S313–S324.

Faye, Jan, "Backward Causation", *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2010/entries/causation-backwards/>.

Frisch, Mathias. 2005a. "Counterfactuals and the Past Hypothesis", in *Philosophy of Science*, Vol. 72, No. 5, December 2005, pp. 739–750.

Frisch, Mathias. 2007. "Causation, Counterfactuals, and the Past-Hypothesis", in Price and Corry (2007).

Frisch, Mathias. "Time and Causation", to appear in *The Blackwell Companion to the Philosophy of Time*, ed. Heather Dyke and Adrian Bardon, URL = https://docs.google.com/viewer?a=v&q=cache:ZnP1F0tHoQoJ:www.philosophy.umd.edu/Faculty/mfrisch/papers/time-and-causation.pdf+&hl=nl&gl=nl&pid=bl&srcid=ADGEESgddGZcKxxzf1PdbFCOn4-afol-ichika_M6CX7O4WNM_JS4Hq-Qh3RPltDiFyw6O5-M_WX8ih8vWpiaxpqHaoPEVwSbn6lj5uS4ZuWHsn8bEgmJVIW0btpGecsCeQ37tV9fibK&sig=AHIEtbRVrBrNMqy_9kkBpuZ-iniSIRqnUw, bekeken 21-1-2013.

Gibbard, Alan, and William Harper. 1978. "Counterfactuals and Two Kinds of Expected Utility", in Clifford Hooker, James Leach, and Edward McLennen (Ed.), *Foundations and Applications of Decision Theory Foundations and Applications of Decision Theory*, D. Reidel, Dordrecht, Holland, pp. 125–162.

Horwich, Paul. 1987. *Asymmetries in Time: Problems in the Philosophy of Science*, MIT Press, Cambridge MA.

Horwich, P. 1987. *Asymmetries in Time*. Cambridge, Mass: MIT Press.

Hume, David. *A Treatise of Human Nature*. URL = <<http://www.gutenberg.org/ebooks/4705>>, bekeken 21-1-2013.

Kutach, Douglas. 2002. "The Entropy Theory of Counterfactuals", in *Philosophy of Science*, Vol. 69, No. 1, March 2002, pp. 82–104.

Lewis, David. 1973b. "Causation", in *The Journal of Philosophy*, Vol. 70, No. 17, October 1973, pp. 556–567. Reprinted in Lewis (1986a: 159–171).

Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow", in *Noûs*, Vol. 13, No. 4, November 1979, pp. 455–476. Reprinted in Lewis (1986a: 32–52).

Lewis, David. 1986a. *Philosophical Papers*, Volume II, Oxford University Press, New York.

Loewer, Barry. 2007. "Counterfactuals and the Second Law", in Price and Corry (2007), pp. 293–326.

Menzies, Peter, "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>>, bekeken 21-1-2013.

Morris, William Edward, "David Hume", *The Stanford Encyclopedia of Philosophy (Spring 2013 Edition)*, Edward N. Zalta (ed.), forthcoming URL = <<http://plato.stanford.edu/archives/spr2013/entries/hume/>>, bekeken 21-1-2013.

Popper, Karl. (1956): "The Arrow of Time", *Nature*, 1777: 538.

Price, Huw. 1996. *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*, Oxford University Press, Oxford.

Price, Huw and Brad Weslake. 2009. "The Time-Asymmetry of Causation". In *The Oxford Handbook of Causation*, ed. Helen Beebe, Peter Menzies, Christopher Hitchcock, 414-443. Oxford University Press.