

ROBOTS EN NOUMENA

Bespreking van: Fred A. Keijzer, *The generation of behavior. On the function of representation in organism-environment dynamics*. Proefschrift Universiteit Leiden, 1998. Alle paginaverwijzingen zijn naar dit boek, tenzij anders vermeld.

Er rijdt een robot over een tafel. Hij stopt keurig bij de rand en kiest dan een andere koers. Als hij opnieuw een tafelrand bereikt, stopt hij en verandert weer van richting. Dit fenomeen herhaalt zich eindeloos. De robot heeft geen afstandsbediening, geen ogen of andere zintuigen. Hoe doet hij dat? Welnu, de beweging van de wielen van de robot wordt via tandraden overgebracht op een stel kleinere wielen binnen in de machine. Tussen deze wieltes rolt een stuk schuurpapier dat een exact schaalmodel van de tafel is. Om het schuurpapier zit een rand. Als de wieltes tegen de rand lopen (d.w.z. als de robot van de tafel dreigt te rijden), slaat een pal om en wordt de rijrichting veranderd.

Stel je nu een andere robot voor. Deze heeft twee lichtgevoelige sensoren als ogen aan de voorkant, één voorwiel en twee achterwielen. Hij zoekt elke lichtbron op die binnen zijn bereik valt. En als je hem een lampje voorhoudt, volgt hij daarvan de beweging. Deze robot heeft geen ingebouwd model van de omgeving. Er zit überhaupt weinig in: alleen twee kruisende kabels en twee motortjes. Licht op het rechteroog laat de motor van het ene achterwiel draaien, licht op het linkeroog die van het andere.

Op het eerste gezicht lijken de twee robots niets anders te illustreren dan twee concurrerende ontwerpstrategieën in de cybernetica. Laten wij zeggen, één met en één zonder geheugen. Filosofisch gezien kun je ze echter ook beschouwen als de belichaming van twee visies op de geest: de eerste robot is een rationalist, de tweede een empirist. De R-bot dankt zijn intelligentie aan een aangeboren idee van de omgeving, terwijl de E-bot zijn intelligentie dankt aan zijn zintuigen. Ook wetenschapsfilosofisch kun je de analogie gebruiken, als illustratie van twee onderzoeksstrategieën in de psychologie. De R-bot representeert het cognitivisme, dat zich toelegt op het ontwerpen van geschikte representatiesystemen en bijbehorende algoritmen. De E-bot representeert het interactionisme, dat modellen ontwerpt van de interactie tussen organisme en omgeving. Ofschoon het cognitivisme nog altijd de boventoon voert, winnen de interactionisten tegenwoordig steeds meer terrein.

De controverse tussen cognitivisten en interactionisten staat centraal in Fred Keijzers proefschrift. Of eigenlijk iets ruimer: de controverse tussen de 'actortheorie', waarvan het cognitivisme de meest populaire variant is, en de 'gedragssysteemtheorie', waarvan het interactionisme de bekendste interpretatie is. Keijzer bepleit het interactionisme. Hij heeft daarvoor drie motieven. Het eerste is dat de notie van representatie in de hedendaagse psychologie volgens hem veel te klakkeloos wordt gebruikt. 'Representatie' is in feite een explanans, het explanandum is 'gedrag'. Maar in de cognitivistische *hype* is het explanans het explanandum gaan overschaduwten. Cognitief modelleren is een doel op zich geworden. Keijzer vindt dit niet juist. "Representations should not merely be assumed, they should earn their explanatory keep. They are part of the explanation of something else" (p. 15). De verklaring van 'gedrag' moet dus weer centraal komen te staan.

Bovendien vindt Keijzer dat de psychologie een bepaald soort gedrag systematisch verwaarloost, namelijk het zogeheten 'basale gedrag', vaak afgedaan als 'mechanisch' en 'zui-

ver instinctief'. Keijzers tweede motief is dat dit gedrag niet moet worden afgeschoven naar de fysiologie, maar juist het uitgangspunt moet vormen van de psychologie. Pas als je basaal gedrag goed begrijpt, kun je van daaruit complexere gedragsvormen bestuderen. Niet toevallig is basaal gedrag ook het parapedaardje van de interactionistische psychologie.

Het derde motief vat de vorige twee in zekere zin samen: de psychologie moet zich waken voor circulaire verklaringen. Als je jezelf representaties toestaat, kun je in termen daarvan vrij gemakkelijk het psychologische afbakenen van het fysiologische, het geestelijke van het mechanische. Maar de notie van representatie is ontleend aan ons alledaagse zelfbegrip, aan de 'volkpsychologie' van personen en hun handelingen. De *verklaring* van gedrag op het persoonlijke niveau kan, op straffe van circulariteit, niet op datzelfde niveau liggen. Zij moet *subpersoonlijke* mechanismen, processen en eigenschappen blootleggen. Cognitivisten zoeken dan ook naarstig een geschikte notie van 'subpersoonlijke representaties'. Daarin schuilt echter een groot gevaar, dat Keijzer scherp onder woorden brengt. Het is bij dergelijke 'subrepresentaties' vaak erg lastig om te bepalen of zij op eigen benen kunnen staan, dan wel een deel van hun verklarend vermogen ontlenen aan de 'echte' representaties waarop zij gemodelleerd zijn. Keijzer spreekt in dit verband van begrippen die "meeliften", maar die op eigen kracht nog geen meter vooruit zouden kunnen komen (pp. 128vv). Het probleem zal duidelijk zijn: om de 'conceptual mileage' van een begrip te verklaren moet je juist *niet* bij zijn lifters aankloppen.

Keijzer toont zich even goed thuis in de psychologie als in de filosofie, waarvoor hij een groot compliment verdient. Na een inleidende verkenning van het terrein (hfdst. 1), schetst hij in achtereenvolgens de grondtrekken (hfdst. 2) en de problemen (hfdst. 3) van de actortheorie, waarbij met name het 'frame'-probleem en het probleem van 'liftende' representatiebegrippen aandacht krijgen. Saillant is hier het beroep op Von Uexküll, de vader van de *Umweltlehre* (pp. 143vv). Hoofdstuk 4 schetst het interactionistische alternatief, vooral gebaseerd op het omstreden werk van Maturana en Varela. Dat basaal gedrag zich leent voor een non-representationele, interactionistische verklaring wordt inmiddels door weinigen bestreden. In het laatste hoofdstuk probeert Keijzer daarom te laten zien dat ook meer complexe gedragsprocessen op die manier kunnen worden verklaard, waaronder fenomenen als anticipatie en hogere-orde leerprocessen.

Keijzers boek is goed geschreven, onderhoudend en consequent beargumenteerd. Het baseert zich bovendien op een boeiende, deels traditionele en deels persoonlijke keus aan literatuur. Slechts op enkele punten, vaak van ondergeschikt belang, vond ik het boek niet bevredigend. Ik noem er hier een paar.

In de eerste plaats is de actortheorie, zoals Keijzers zelf stelt, "a reconstructed position, it is not a position that is explicitly held, or defended by cognitive scientists" (p. 48). Wanneer je zo'n theorie aanvalt, loop je weldra het risico van 'beating a straw man'. Een mogelijke (zij het niet noodzakelijk afdoende) remedie tegen deze drogredenering is gebruik te maken van een wetenschapsfilosofische inbedding, bijvoorbeeld in termen van Lakatos' rationele reconstructies van wetenschappelijke onderzoeksprogramma's. Keijzer doet iets dergelijks alleen impliciet; een expliciete redenering had zijn betoog kunnen versterken.

Eén van Keijzers hoofdbezwaren tegen de actortheorie is de drogredenering van de homunculus: aan subpersoonlijke entiteiten eigenschappen toeschrijven die eigenlijk alleen

aan personen toekomen, bijvoorbeeld representaties. Het is op zijn minst merkwaardig dat de schrijver hier de *locus classicus* van dit argument niet noemt, Gilbert Ryles *The concept of mind* (1948), noch verwante Wittgensteiniaanse lijnen van psychologiekritiek.

Daniel Dennett en andere filosofen hebben bepleit dat je de homunculi in de psychologie kunt elimineren. Vervang gewoon elke homunculus door een samenwerkend stel van nét iets dommere homunculi en herhaal dit procédé, tot je een leger van homunculi overhoudt dat zó dom is dat zij kunnen worden vervangen door een mechanische schakeling. Dennett noemt dit ‘intelligentie-op-de-pof’: beschouw de intelligentie voor je homunculi als een lening die je uiteindelijk moet aflossen met mechanismen. Weliswaar noemt Keijzer deze constructie (de ‘lening’ is wat Keijzer ‘liften’ noemt), maar hij spreekt zich niet uit over Dennetts argument als zodanig. Evenmin gaat hij in op het bekende tegenargument dat zelfs de domste homunculus nog altijd een homunculus is. Hoe Keijzer zou reageren op deze argumenten laat zich wel raden, maar de redenering ontbreekt.

Een groot probleem voor de actortheorie is dat van de ‘meeliftende begrippen’. Keijzer wijst erop dat het niet alleen gaat om het begrip ‘representatie’, maar ook om dat van de ‘omgeving’. Een niet-circulaire actortheorie heeft een objectieve, canonieke beschrijving van de omgeving nodig, onafhankelijk van elke specifieke wijze van representeren ervan. Keijzer wijst erop dat dit onhaalbaar is. De waarnemer gebruikt altijd zijn eigen kijk op de omgeving van het organisme als ‘objectieve’ beschrijving. Maar die beschrijving is een functie van de representaties van de waarnemer, dus lift het omgevingsbegrip van het organisme mee met dat van de waarnemer. Mede als prelude op het interactionistische alternatief dat hij voorstaat, beschrijft Keijzer hier de theorie van Von Uexküll, dat elk (type) organisme zijn eigen *Umwelt* heeft als correlaat van zijn fysiologische repertoire, zijn “specific neural system and embodiment” (p. 143). Ter verklaring van het gedrag van het organisme heb je daarom weinig aan een ‘neutrale’ omgeving. Het gaat juist om het interactieve ‘reliëf’ van de omgeving, d.w.z. om de ecologische eigenschappen en processen waarmee het organisme interageert. En dat verschilt van de ene soort tot de andere.

Welnu, Keijzers opmerkingen op dit punt lijken mij beslist juist, maar hun status in de argumentatie omtrent actortheorie en interactionisme is dubbelzinnig. Zijn de opmerkingen metafysisch bedoeld of methodologisch? Metafysisch gesproken is het onmogelijk is om de omgeving-als-zodanig te beschrijven, de *Welt an sich*, los van elk representatie- en referentiekader. Dat is een onkenbaar noumenon. Soms neigt Keijzer inderdaad naar deze uitleg, vooral in zijn bespreking van Maturana en de idealistische traditie (pp. 161vv). Maar soms neigt hij ook duidelijk naar een zwakkere, methodologische uitleg: zonder ingebouwde *Umwelt*-gevoeligheid heeft een gedragstheorie geen kans van slagen. Uiteindelijk staat Keijzer voor een dilemma. Als de opmerkingen metafysisch bedoeld zijn, dan slagen zij weliswaar als argument tegen de actortheorie, maar halen zij ook de gedragssysteemtheorie onderuit. Zijn zij daarentegen methodologisch bedoeld, dan pleiten zij weliswaar voor zoiets als interactionisme, maar vormen zij geen echte bedreiging voor de actortheorie. Die kan immers zonder veel problemen worden verrijkt met üxküllöide elementen.

Leiden

Jan Sleutels

Word count: 1550