

AI & Psychologie

Simpel maar effectief: de rol van Computatieve Modellen binnen de functieleer

dr. Stefan Van der Stigchel, Psychologische Functieer Universiteit Utrecht
s.vanderstigchel@uu.nl

Toen eind jaren tachtig de ontwikkeling van computationele modellen sterk toenam, was de belofte groot. Vooral connectionistische modellen zouden een brug vormen tussen breinprocessen en gedrag. Doordat ze gebaseerd zijn op de werking van de hersenen, is het gedrag wat een model vertoont een mogelijke instantie van de werkelijkheid. Deze

modellen kunnen verbale theorieën aanvullen doordat ze *laten zien* dat een bepaalde set hypothesen ook daadwerkelijk werkt. Deze functie is uitermate geschikt voor de functieleer, omdat hier verbale theorieën de meest gebruikte manier zijn om bepaalde fenomenen te verklaren.

Verbaal uitleggen dat het Stroop-effect¹ wordt veroorzaakt doordat semantische informatie sneller wordt verwerkt dan kleurinformatie is informatief, maar de computationele implementatie is een nuttige toevoeging. Geïnspireerd op ideeën van James McClelland en David Rumelhart is er aantal redenen te noemen waarom modellen een nuttige toevoeging zijn. Als belangrijk voordeel kan gezien worden dat een model bevestigt dat een bepaalde set hypothesen binnen een verbale theorie op computationele gronden het gedrag oplevert dat het wil verklaren. Is er bijvoorbeeld een set formules en waarden te vinden dat gedrag oplevert overeenkomstig met het Stroop-effect?

Een model dwingt tot het expliciet maken van een theorie, omdat verbale concepten vertaald moeten worden in wiskundige functies. Wat betekent het precies dat informatie ‘sneller’ verwerkt wordt? Er zijn verschillende manieren om dit proces te implementeren.

In tekstboeken over functieleer worden computationele modellen vaak kort aan het einde van een hoofdstuk behandeld. Natuurlijk bestaan er tekstboeken waarin specifiek de rol van computationele modellen binnen de functieleer wordt besproken, maar deze lijken toch meer bedoeld voor specialisten. Je kunt je afvragen of de stand van zaken binnen het onderwijs een weerspiegeling is van de wetenschappelijke realiteit. Is de koppeling tussen computationele modellen en de functieleer inderdaad problematisch?

Door het opstellen van een model wordt deze ambiguïteit opgelost. Dit levert tevens als voordeel op dat de noodzakelijkheid van de gebruikte concepten kan worden geëvalueerd. Je kunt wel een concept als ‘aandacht’ betrekken in een verbale theorie over het Stroop-effect, maar wanneer blijkt dat een computationeel model dit

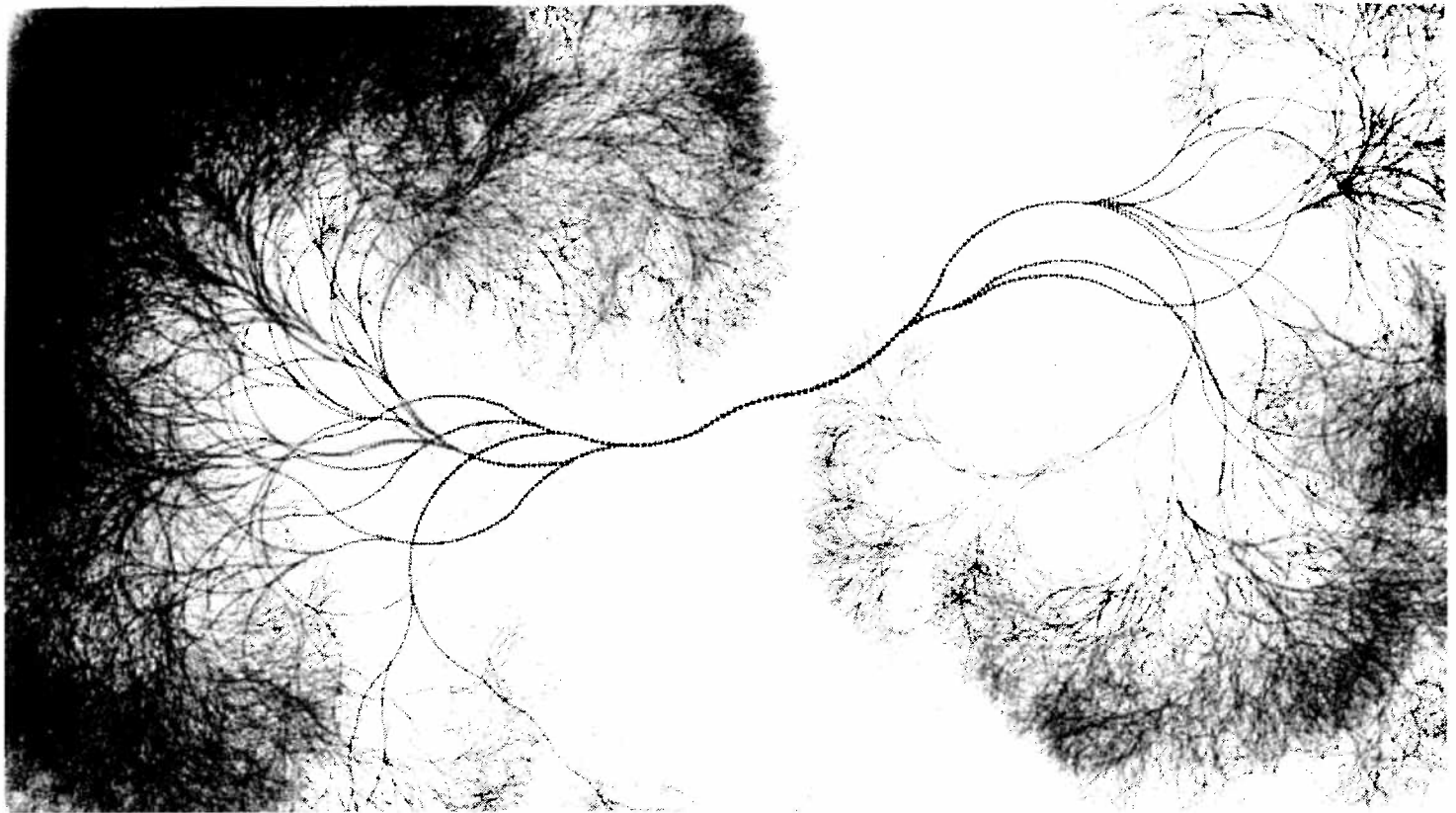
concept niet nodig heeft om het effect te verklaren, is dit concept overbodig. Op deze manier kunnen er nieuwe verklaringen komen voor bestaande bevindingen.

Een ander voordeel is dat een model een voorspelling kan doen over een bepaald type gedrag dat vervolgens experimenteel getest kan worden. Wanneer deze voorspelling inderdaad blijkt te kloppen, levert dit niet alleen extra bewijs op dat een model een correcte weergave van de werkelijkheid is, maar draagt het model ook bij aan de wetenschappelijke zoektocht naar nieuwe hypothesen.

Succesvolle modellen Nu het nut van modelleren toegelicht is, volgt automatisch de vraag wat er voor zorgt dat een model geaccepteerd wordt binnen de functieleer. Hoewel er veel computationele modellen binnen de functieleer bestaan, is er maar een klein aantal dat bijdraagt aan de beeldvorming van onderzoekers binnen de functieleer. Wanneer je kijkt naar de belangrijkste wetenschappelijke tijdschriften binnen de functieleer, zoals *Journal of Experimental Psychology* en *Cognitive Psychology*, merk je dat hier weinig artikelen in staan waarin computationele modellen besproken worden. Deze verschijnen meestal in *proceedings* voor AI-congressen of specialistische tijdschriften als *Neural Networks*.

¹ Het Stroop-effect verwijst naar het feit dat het makkelijk is om ‘rood-groen-blauw’ te zeggen als die woorden een correspondende kleur hebben dan wanneer de kleur verschilt van het woord (bijvoorbeeld het woord “blauw” in rode letters).

² Het executief systeem kan vertaald worden met cognitieve controle. Hieronder vallen allerlei breinprocessen, zoals plannen, abstract redeneren, mentale flexibiliteit, het inhiberen van acties, leren van regels, selecteren van sensorische informatie en het probleemoplossend vermogen.



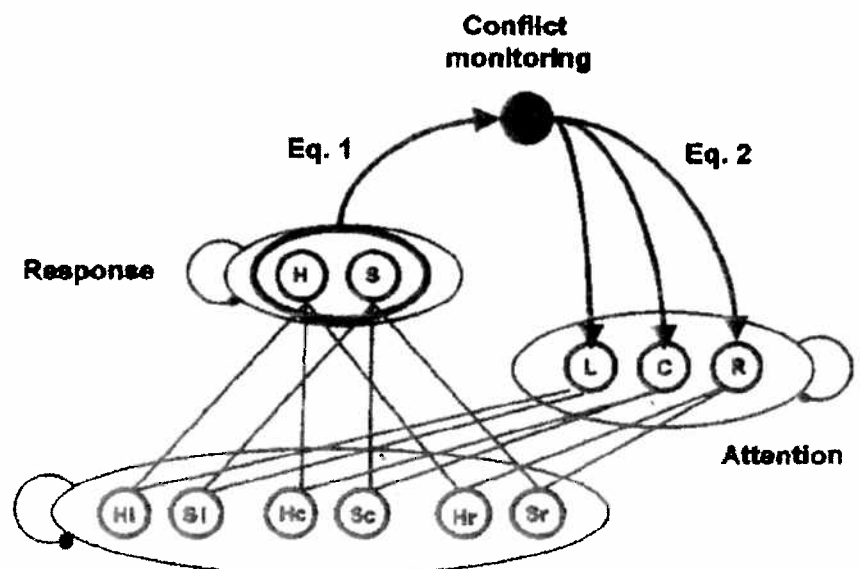
Een belangrijke uitzondering is het invloedrijke *Psychological Review*, waarin regelmatig computationele modellen gepubliceerd worden. Aan de hand van een tweetal succesvolle modellen zal besproken worden waaraan een model moet voldoen om een belangrijke rol te spelen in de functieleer.

Een voorbeeld van een zeer succesvol model is 'Conflict monitoring and cognitive control' (Botvinick, Braver, Barch, Carter, & Cohen, 2001).

Het succes van dit model is te meten aan het aantal citaties van het paper. In zeven jaar tijd is dit paper ruim zevenhonderd keer geciteerd. De kracht van dit model zit hem niet in de biologische plausibiliteit, maar in de introductie van een nieuw concept, waarmee het verschillende fenomenen binnen de functieleer kan verklaren. Dit concept, 'conflict', speelt nu een belangrijke rol in theorieën over executief functioneren².

In het model van Botvinick wordt conflict gedefinieerd als de hoeveelheid energie in de outputlaag van een model. Bij weinig conflict is er weinig energie, en zal het model geen problemen hebben om snel en correct tot de juiste output te komen. Is er veel conflict (en dus

veel energie), dan zal het model trager en ook minder accuraat tot de juiste output komen. Door het toevoegen van een module dat de hoeveelheid conflict in de gaten houdt (*monitoring*) kunnen de waardes van het model zodanig worden aangepast dat het conflict in de volgende trial verminderd wordt. Het enige wat deze module doet, is het monitoren van de hoeveelheid conflict en het aanpassen van de waardes in de modules waar de taakinstructie wordt gerepresenteerd.



Figuur 0 - Uit de paper 'Conflict Monitoring and Cognitive Control' by Botvinick, Braver, Barch, Carter, and Cohen



Een ander succesverhaal is “A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus” (Trappenberg, Dorris, Munoz, & Klein, 2001). Voor mijn eigen onderzoek is dit van groot belang geweest. Het model beschrijft hoe de beslissing wordt genomen over de doellocatie van een volgende oogbeweging. Een dergelijke beslissing is afhankelijk van twee stromen van informatie. Ten eerste is er ‘top-down’ informatie; dit is de informatie van alle interne processen. Als je gevraagd wordt om naar de deur te kijken, dan kun je deze taak

In de paper van Botvinick worden bestaande modellen gebruikt waaraan een conflictmodule is toegevoegd. Van Strooptaken is bijvoorbeeld bekend dat de hoeveelheid incongruente trials (bijvoorbeeld het woord ‘rood’ in het groen) het Stroopeffect bepaalt. Hoe meer incongruente trials, hoe sterker het stroop-effect. Door het toevoegen van een conflictmodule aan een bestaand Stroopmodel werd dit effect gerepliceerd.

De voorspellende waarde van het Botvinick model bleek hoog. Niet alleen gedragsmatig, maar ook met betrekking tot het brein. De *anterior cingulate cortex* (ACC) lijkt namelijk de rol te spelen van conflictmonitor. Dit gebied is actief op het moment dat er een trial wordt uitgevoerd waarin veel conflict is. De verschillende hypothesen die gegenereerd werden over de activiteit van de conflictmodule bleken te kloppen met activiteit van de ACC.

Dit model voldoet dus aan de eisen van een succesvol model zoals eerder opgesteld. Het geeft een nieuwe verklaring voor bestaande bevindingen en het maakt deze verklaring expliciet. Er wordt aangetoond dat de voorgestelde set hypothesen ook daadwerkelijk werkt, en de theorie genereert een nieuwe set hypothesen. Voorspellingen worden gedaan, en voorspellingen blijken te kloppen.

uitvoeren door een top-down oogbeweging naar de deur te maken. Ten tweede is er ‘bottom-up’ informatie; dit is de informatie van buitenaf. Als je in gesprek bent en er komt ineens iemand de kamer binnen, maak je vaak automatisch een bottom-up oogbeweging naar de deur. Er is dus een continue competitie tussen bottom-up en top-down informatie om te bepalen wat de doellocatie is van de volgende oogbeweging.

Het model van Trappenberg geeft een mogelijke implementatie van deze competitie. Op basis van gegevens uit de neurofysiologie bestaat het model uit een tweedimensionaal neurale veld waar top-down en bottom-up informatie worden geïntegreerd. De neuronen in dit neurale veld corresponderen met locaties in de buitenwereld, zoals dit ook in het brein is gecodeerd. De twee stromen van informatie (bottom-up en top-down) zijn in competitie met elkaar, maar worden gerepresenteerd op één enkele map. De activiteit die het gevolg is van de competitieve integratie van deze twee stromen van activiteit bepaalt vervolgens het gedrag. Daar waar als eerste de activatie-threshold wordt bereikt, daar kijk je naar. Zo zal bijvoorbeeld een nieuw opvallend element gemakkelijk een oogbeweging uitlokken, omdat deze sterke bottom-up activiteit veroorzaakt.

Er zijn drie sterke punten aan dit model. Ten eerste geeft het een raamwerk om te kunnen denken over oogbewegingprocessen. Als ik nadenk over verklaringen voor data uit een oogbewegingsexperiment denk ik vaak in termen van dit model. Ik heb hiermee een manier om de processen die in mijn experiment een rol spelen te 'visualiseren' om tot theorievorming te komen. Ten tweede kan dit model verschillende fenomenen verklaren door een vrij simpel gemeenschappelijk mechanisme, namelijk de competitieve integratie van twee stromen van informatie. Ten derde is dit model geïnspireerd op de werking van de superior colliculus, een gebied in middenhersenen. Van dit gebied is het bekend dat het een belangrijke rol speelt in de aansturing van oogbewegingen. De activiteitspatronen van de twee stromen van informatie zijn gebaseerd op opnames van neuronen in dit gebied en de gebieden waar deze informatie vandaan komt.

Natuurlijk zijn deze modellen slechts twee voorbeelden van de tientallen die er zijn binnen alle

domeinen van de functieleer. Opvallend genoeg kun je stellen dat de modellen die wel succesvol zijn ook meteen zeer bepalend zijn voor theorievorming binnen dat vakgebied. Het lijkt een kwestie van alles of niets. Wat maakt een model zo invloedrijk?

Overgecompliceerde modellen worden overgeslagen

Buiten dat het model moet voldoen aan de kenmerken zoals eerder beschreven, lijkt het succes vooral af te hangen van de simpelheid waarmee het bestaande fenomenen kan verklaren. Beide besproken modellen zijn zo sterk omdat iedere onderzoeker binnen de functieleer het principe achter het model begrijpt en kan gebruiken om de eigen ideeën vorm te geven. Deze onderzoekers zullen misschien niet de formules begrijpen, maar de exacte implementatie doet er eigenlijk in dat geval niet toe. Het gaat om de innovatieve gedachten achter de formules. Hierbij is het uiteraard ook van belang dat een model gepubliceerd is in een tijdschrift dat ook gelezen wordt door onderzoekers die zuiver experimenteel werk doen. Op deze manier zal ook dit publiek kennis kunnen nemen van het model.

Natuurlijk is het prachtig als een model biologisch plausibel is, maar het lijkt wel of modeleerders daar meer waarde aan hech-

ten dan niet-modeleerders dat doen. Bij dit laatste publiek gaat het er niet om dat een model neuronen gebruikt die sterk lijken op neuronen in de hersenen, maar om hoe zij het model kunnen gebruiken in hun eigen beeldvorming van de processen die zij onderzoeken. Het toevoegen van biologische kenmerken op zich geeft een model ook niet meer verklarende kracht. Overgecompliceerde modellen die perfect alle data kunnen repliceren maar dermate ingewikkeld zijn dat alleen wetenschappers die hogere wiskunde begrijpen de puzzel kunnen oplossen, zullen worden overgeslagen door onderzoekers zonder modelleerervaring. Ze gebruiken niet zozeer de meest biologisch plausibele modellen, maar vooral de meest begrijpbare en bruikbare. Het blijft altijd een gesimplificeerde weergave van de werkelijkheid, dus je kunt je afvragen hoeveel detail nuttig is.

Nu weet ik uit eigen ervaring dat het bijzonder lastig is om een model gepubliceerd te krijgen in een wetenschappelijk tijdschrift

waarin ook experimentele studies worden gepubliceerd. Je model moet echt baanbrekend zijn, wil het geaccepteerd worden door zo'n blad. Het zou goed zijn als dit soort tijdschriften de drempel van innovativiteit op een zelfde niveau zetten als voor experimentele studies. Een dergelijke verandering zal zeker zorgen voor een betere integratie van computationele modellen binnen de functieleer. Ik hoop met bovenstaande voorbeelden te hebben aangetoond dat de ontwikkeling van dit soort modellen binnen de functieleer van groot belang is. Zolang het model maar simpel is en een raamwerk biedt om data te verklaren, zijn dit soort modellen zeer invloedrijk en cruciaal om een stap verder te komen binnen de functieleer. ◊

Referenties:

- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. *Psychological Review* 2001; 108(3):624-652.
- Trappenberg TP, Dorris MC, Munoz DP, Klein RM. A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience* 2001; 13(2):256-271.