

## Grondslagen van de Psychologie

college 12  
philosophy of mind & cognition  
13-03-2009

### Functionalisme Deel I

## recapitulatie

- Blok II: Philosophy of Mind;
- We hebben drie typen mentale toestanden;
- We willen weten hoe de relatie tussen geest en de (rest van de) fysische wereld is;
- De filosofische problemen zijn uit de weggeruimd;
- We zijn de verschillende belangrijke posities aan het bekijken.

2

- ~~1. **Substantie dualisme:** de geest bestaat onafhankelijk van het lichaam en vice versa;~~
- ~~2. **Epifenomenalisme:** de geest doet niets;~~
- ~~3. **Idealisme:** de fysische wereld is afhankelijk van de geestelijke wereld;~~
- ~~4. **Behaviorisme:** de geest is eigenlijk gedrag;~~
- ~~5. **Eliminativisme:** de geest bestaat niet;~~
6. **Reductionisme / identiteitstheorie:** mentale toestanden zijn hersentoestanden;
7. **Functionalisme:** mentale toestanden worden gerealiseerd door hersentoestanden.

3

## MBIT

- Argumenten pro MBIT zijn niet weerlegd;
- Argumenten contra MBIT vooralsnog wel;
- Eliminativisme: serieus te nemen voor de termen in de *folk psychology* die niet refereren (en onderzoek moet dat uitwijzen).

4

## Programma college 12 & 13: functionalisme

1. Toch nog een argument tegen MBIT (?);
2. Functionalisme;
3. Functionalisme vs. behaviorisme;
4. Functionalisme vs. identity theory;
5. Functionalisme: welk aspect van de geest verklaart het mogelijk? (Wel: intentionaliteit, niet: qualia?)
6. De eerste stappen in de AI.
7. Naar een moderne AI (college 13).

5

## 1. Meervoudige realisatie als argument tegen MBIT

### Meervoudige realisatie

- Er is een argument wat het aannemelijk maakt dat de MBIT een te sterke claim maakt (maar dat een verwante claim wel juist is);
- Om dit argument te begrijpen hebben we naast de theorie van Kripke over identiteit de notie 'meervoudig realiseerbaar' (MR) nodig;
- Wat is meervoudige realiseerbaarheid?

TB: 40, 48-49, 159-162, B. 2.4; B. 7.1

### Uitleg a.d.h.v. 2 voorbeelden

1. Een klok moet de tijd aan geven;
  2. Vloeibaarheid is een eigenschap van een stof z.d.d. sommige zaken zinken en andere blijven drijven;
- Wat je doet is dus analyseren wat een klok of wat vloeibaarheid is in termen van wat het doet; de **functie** of **causale rol** die het speelt;
  - *Hoe is dit van toepassing op een argument tegen MBIT?*

8

- **Claim MBIT:** pijn = vuren c-vezels;
- **Gegeven Kripke's theorie:** pijn = vuren c-vezels en niets anders;
- **Vergelijk:** water = H<sub>2</sub>O en niets anders, als je zegt water = XYZ dan heb je de taal veranderd;
- **Ergo:** Als de MBIT correct is, dan kunnen wezens met een andere fysische structuur dus geen pijn hebben.

9

### Maar dat is onwaarschijnlijk:

- (1) Vele diersoorten op aarde hebben een andere fysische structuur dan mensen hebben (vissen, reptielen, insecten), deze zouden dan geen pijn kunnen hebben;
- (2) Het is voorstelbaar (wat op zich geen sterk argument is) dat er aliens zijn die wel pijn hebben, maar waarvan hun hersenen van niet van koolwaterstofverbindingen gemaakt zijn.

10

- M.a.w. mentale toestanden lijken eerder bij de categorie vloeibaarheid, en klok te horen, dan bij water.
- Dit argument wordt vaak als fataal voor de MBIT gezien, vooral vanwege de combinatie van de volgende twee zaken: (1) de plausibiliteit dat andere dieren op aarde ook pijn ervaren, en (2) het alternatief voor een MB theorie dat uit deze MR kritiek volgt: het functionalisme.

11

### 2. Functionalisme

TB: 48-50, 147, 156-180, B. 2.4, B. 7.1

### mentaal holisme

- Mentale toestanden zijn toestanden die een causale rol vervullen tussen (1) input; (2) output; en (3) andere mentale toestanden;
- Hierdoor heeft de functionalist de mogelijkheid om mentale toestanden (deels) te karakteriseren in verhouding tot andere mentale toestanden.

13

### Voorbeeld

	<i>input</i>	<i>output</i>	<i>relatie andere mentale toestanden</i>
<i>pijn</i>	weefselschade	- rillen & kreunen; - je kan minder goed typen;	- je raakt gestresst; - je bent minder alert;

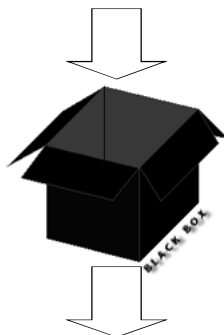
14

- Om verder te bekijken wat functionalisme is, kunnen we het afzetten tegen:  
[1] Behaviorisme; &  
[2] MBIT.

15

### 3. Functionalisme vs. Behaviorisme

- Vaak ziet men het functionalisme als een wat beter uitgewerkte versie van het behaviorisme – immers beide posities karakteriseren het mentale in termen van stimulus (input) en response (output);
- Er is echter een belangrijk verschil: het functionalisme zegt iets over wat er in *the black box* van het behaviorisme zit.



17

### functionalisme accepteert mentaal realisme

- Het functionalisme meent dat mentale toestanden interne toestanden zijn die causale krachten bezitten;
- De behaviorist wil precies af van gepraat over interne toestanden;
- M.a.w. de functionalist is realist omtrent mentale toestanden, en de behaviorist is een anti-realist; een instrumentalist; een nominalist.

18

#### 4. Functionalisme vs. MBIT

#### Het model van MBIT

- **Reductie:**
  - [1] Wat is de causale rol analyse van X?
  - [2] Datgene wat die rol vervult is X.
- Vandaar dat MBIT = reductief materialisme.

20

#### Het model van het functionalisme

- **Reductieve verklaring:**
  - [1] Wat is de causale rol analyse van X?
  - [2] Datgene wat die rol vervult realiseert X.
- Vandaar dat functionalisme = non-reductief materialisme.

21

#### Voorbeeld: pijn

- Wat is de functie van pijn?
- Binnen een evolutionair perspectief is dat “doel” ervoor zorgen dat het wezen dat pijn heeft adequaat reageert op [1] de oorzaak van de weefselschade & [2] de weefselschade zelf;
- Hoe je pijnervaringen maakt (= realiseert) - van hersenen of computers - is irrelevant.



- Als je meent dat zo'n functionalistische kijk op (o.a.) pijn correct is, dan dringt een analogie / metafoor zich op:
- Het lijkt erop dat de geest een informatieverwerkend systeem is (het verwerkt de informatie dat er ergens weefselschade is, zodat je er correct op kan reageren);
- Computers zijn ook informatieverwerkende systemen.

23

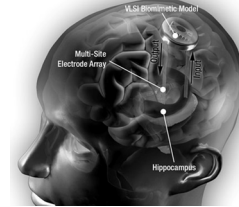
#### 5. Functionalisme m.b.t. wat?

### Twee soorten bewustzijn

- Volgens veel filosofen is cognitie wel maar fenomenaal bewustzijn niet functioneel te analyseren;
- Zie o.a. Chalmers' zombie argument;
- Ook Jaegwon Kim & Ned Block menen dat dat op zijn minst problematisch is;
- **HD:** Laten we eerst kijken naar de AI - dat gaat dus over kunstmatig cognitief bewustzijn (college vandaag & college 13).

25

### 6. De eerste stappen in de AI



### Hoe pakken we onderzoek naar AI aan?

- **Paul Churchland:** We moeten dat doen via *the piecemeal approach* (stukje-bij-beetje aanpak);
- Bewuste intelligentie is complex, en dus moeten we kijken of we niet eerst onderdelen (aparte cognitieve capaciteiten, zoals kunnen rekenen) kunnen namaken;
- Daarna moeten we kijken of we die verschillende cognitieve capaciteiten kunnen integreren in één K.I. wezen.

27

- Als je op deze manier tegen het AI project aankijkt, dan is het volgende duidelijk:
- “For this reason, AI researchers have tended to single out some one aspect of intelligence, and then concentration on simulating that aspect. As a matter of strategy, problems of integration can temporarily be put aside.” (Churchland 1988: 106)

28

### Program bespreking klassieke AI

#### [I] Case studies:

- [a] Het intelligent oplossen van problemen door doelgericht handelen (drie manieren);
- [b] Leren;
- [c] Het manipuleren van natuurlijke taal;

#### [II] Computer functionalisme

#### [III] Problemen met computer functionalisme

TB: 167-180  
29

### [I] Case study [a]: het intelligent oplossen van problemen

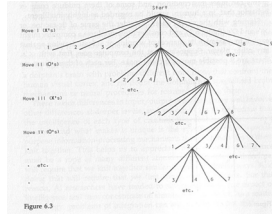
PC bespreekt drie manieren waarop een KI systeem problemen kan oplossen:

- [1] Brute kracht;
- [2] Heuristisch;
- [3] Terugrekenen.

30

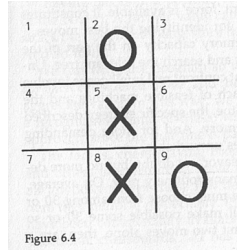
## [1] Brute kracht

- Neem boter-kaas-en-eieren;
- Het doelgerichte gedrag is winnen;
- Je kan dat uiteenzetten in een diagram.



31

- Een computer kan berekenen wat alle mogelijke vertakkingen zijn;
- Een computer die dat kan, zal net als wij inzien dat je geen kruisje moet zetten in vakje 7, omdat je dan verliest.



32

- De computer is in staat om alle mogelijke zetten te berekenen (m.a.w. je zal er nooit van winnen);
- Dit is *exhausted-look-ahead*;
- Probleem: Deze brute kracht methode wordt al snel onbruikbaar;
- Neem schaken, ook hier is winnen het doelgericht gedrag wat je wil.

33

## Schaken

- Per beurt kan je zo'n 30 zetten doen;
- In 1 spel worden zo'n 80 zetten gedaan;
- Wil je de zaak dus in een boomdiagram uiteenzetten, dan heb je dus  $30^{80}$  vertakkingen nodig dat is ongeveer  $10^{118}$ .

34

- "A million computers each examining a million branches per second would still take  $10^{100}$  years to examine the entire tree." (1988: 111)
- Het moet dus anders.

35

## [2] de heuristische methode

Een oplossing is om heuristische middelen te gebruiken, *rules-of-thumb*, zoals:

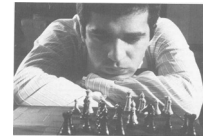
- Een koningin veroveren door een loper te offeren, levert grotere winkansen op;
- Rokade levert kleinere verlieskans op;
- Als je hier dan aan toevoegt dat de computer een aantal (maar dus niet alle) zetten vooruit kan "denken", dan heb je een computer die goed en snel is.

36

“The fact is, chess-playing programs using heuristics as these, and other more cunning, have been written that will beat the pants off anybody except those very few devotes at the master levels, and even here they perform respectably.” (1988: 111)

37

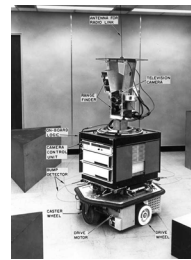
- **HD:** Dat was in 1988 wellicht nog zo,
- In 1997 heeft de IBM computer Deep Blue van Garry Kasparov gewonnen.



38

### [3] terugrekenen

- *Brute force* kost teveel tijd (maar is wel onfeilbaar);
- De *heuristische methode* is sneller, maar is feilbaar;
- Wat je ook zou kunnen doen is slechts één win-situatie nemen, en dan terugrekenen wat je moet doen om die situatie te bereiken;
- De vraag is echter of er dan nog niet teveel mogelijkheden over blijven, zodat je uiteindelijk zolang met rekenen bezig bent dat de zaak toch niet in de praktijk uit kan rekenen.



- Shakey (één van de eerste mobiele robots) werkte volgens dit principe;
- Shakey moest een hele kaart van de omgeving maken alvorens hij van A naar B kon rijden: the filing cabinet method.

40

### Shakey maakte een gedetailleerde kaart

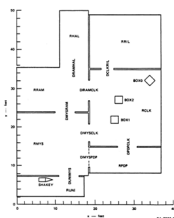


Figure 16. MAP OF SHAKY'S EXPERIMENTAL ENVIRONMENT.  
From Nilsson (1984).

“In the late sixties and early seventies the blocks world became a popular domain for AI research. [...] The key to success was to represent the state of the world completely and explicitly.” (Brooks 1991: 142)

41

### Wat is er (o.a.) mis met Shakey?

- Okay, het werkt, maar is niet echt biologisch realistisch:
  - Shakey moet een uur rekenen om een hele korte afstand af te kunnen leggen;
  - Shakey moet dus een gedetailleerde representatie van de omgeving maken voordat hij van A naar B kan.

42

### Case study [b]: leren

- Kunnen *leren* is onderdeel van conscious intelligence;
- De eenvoudigste manier van leren is de zaak opslaan in een toegankelijk geheugen;
- Een andere manier is om je heuristische regeltjes langzaam aan te passen;
- **Voorbeeld:** Stel dat je teveel gewicht hebt gehangen aan het regeltje een koningin ruilen tegen een minder schaakstuk verhoogt je winkansen. (Door te hoge gewicht verlies je dus.)

43

- Als de computer de win-verlies ratio bijhoudt en dan kan evalueren hoe goed/slecht het gegaan is, dan kan de computer op basis van die gegevens de *rule-of-thumb* bijstellen door er minder gewicht aan te hangen;
- De computer reageert vervolgens anders en heeft dus van zijn fouten geleerd.

44

### case study [c]: natural language manipulation

- Churchland zegt dat hierover nog weinig bekend is, en dat de successen die er zijn voornamelijk neerkomen op succes op het niveau van syntaxis, niet op het niveau van semantiek;
- Prof. Luc Steels (VU Brussel) heeft hier interessant onderzoek naar gedaan.



- Door onderlinge communicatie ontwikkelen de *Talking Heads* een gemeenschappelijke taal om naar de gekleurde objecten te verwijzen. (Hoewel er ook synonymie en polysemie ontstaat.)
- **Probleem voor AI:** Mensen zijn zeer goed in het herkennen van gezichten, en computers niet.

46

### Peicemeal approach

- PC stelt (in 1984) dat we dit soort zaken uit moeten zoeken en ze dan samen moeten voegen tot 1 systeem dat dan wellicht op menselijke cognitie lijkt;
- We hebben al gezien: als je functionalisme accepteert, dringt zich de computer metafoor op;
- Kan je niet verschillende softwareprogramma's maken en die dan later samenvoegen?

47

### [II] computer functionalisme



Alan Turing



Hilary Putnam

- Door de opkomst van de AI werd het mogelijk om het-brein-is-een-computer metafoor te gebruiken (waarbij de-geest-is-de-software) hoort;
- Doordat Alan Turing een universele computer beschreef (die MR was!!) werd dit dan ook echt uitgewerkt (door Putnam).

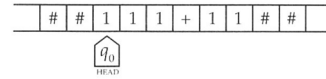
48

TB: 163-164, 174-176, 178-179, B, 7.2, B, 7.3, B, 7.5

## De Turing Machine

Een TM bevat de volgende componenten (vergelijk het met een cassette-speler):

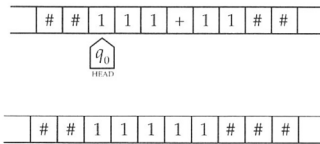
1. Een tape die verdeeld is in vakjes; deze tape kan naar links en naar rechts bewogen worden;
2. Een kop: deze kan lezen en schrijven;
3. Een eindige set interne toestanden (= configuraties); deze geven we aan met  $q_0$  t/m  $q_n$ ;
4. Een alfabet bestaande uit de symbolen  $b_1$  t/m  $b_m$ ; (er kan slechts één symbool per vakje op de tape).



Een Turing Machine kan de volgende dingen doen:

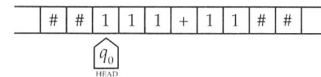
- (1) Op elk tijdstip  $t$  is de TM in een bepaalde interne toestand  $q_i$  en de kop scant een stukje van de tape;
- (2) Wat de TM doet is afhankelijk van (1) de interne toestand & (2) het symbool op de tape; dit kunnen drie dingen zijn:
  - (i) Het symbool vervangen door een ander (evt. zelfde) symbool;
  - (ii) Een vakje naar links of naar rechts gaan (tenzij het proces afgelopen is, dan blijft de kop stilstaan);
  - (iii) De TM komt nu in een nieuwe interne toestand (die ook evt. dezelfde kan zijn).

50



- Hoe komen we van de eerste naar de tweede situatie?
- Dat kunnen we in een machine tabel uiteenzetten.

51



	Toestand $q_0$	Toestand $q_1$
Symbool op tape: 1	Opdracht: $1Rq_0$	Opdracht: #halt
Symbool op tape: +	Opdracht: $1Rq_0$	
Symbool op tape: #	Opdracht: $\#Lq_1$	

52

- Elke berekening die je wiskundig kan bedenken kan een Turing machine uitvoeren;
- Dus: als iets computationeel is, kan een TM dat uitvoeren;
- Bovendien: het maakt niet uit waar je de TM van maakt – **MR van de TM dus!!**

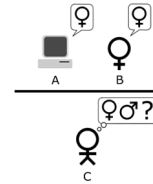
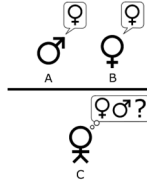
53

## De geest als een Turing machine

- *What it is for an organism, or system, **to have a psychology**, that is, what it is for that organism to have mentality – **is for it to realize an appropriate Turing Machine.** [...]*  
*In short, our brain is our mind because it is a computer, not because it is the kind of organic, biological structure it is. (Kim 1996: 91)*
- Dit komt neer op het **slagen voor de Turing Test = denken.**

### The Turing test

- AT noemde dit “the imitation game”;
- C (de ondervrager) communiceert via email met persoon A & persoon B;
- Is C in staat om te achterhalen wie de echte vrouw is?



What will happen when a machine takes the part of A in this game? Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‘Can machines think?’

Alan Turing 1950: 434

### Als de computer slaagt voor de Turing Test

- Dan is de computer dus net zo goed als een mens;
- Dan moet je dus eigenlijk zeggen dat de computer kan denken (zoals mensen denken): de computer is een denkend ding, een bewust ding;
- En dan is denken dus meervoudig realiseerbaar.

57

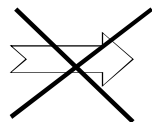
### [III] Problemen met de geest als een Turing Machine

- Werkt de geest wel op deze manier [A] symbool als input, [B] regel om dit te verwerken (in relatie tot andere symbolen) en [C] symbool als output?
- Velen menen van niet, en John Searle heeft die intuïtie omgezet in een “argument”.

58

### Dus (aldus Searle):

Syntax



Semantiek

TB: 178-179, B. 7.4

### The Filing Cabinet Method

- *The filing cabinet method* is de methode die hier gebruikt wordt: je stopt alle informatie in het systeem (de machine tabel) en laat het dan pas los in de wereld;
- Dat is biologisch onrealistisch als model voor cognitie en bewustzijn;
- Bovendien faalt het systeem als je de wereld ook maar een heel klein beetje verandert.

60

### **Dit is een serieel proces**

- In dit model gaat alles stapje voor stapje: cognitie zou een serieel proces zijn;
- Hersenen werken niet serieel maar parallel (Cf. komend college);
- Dat is wellicht zo voor een leerproces, maar niet als je iets al kan (denk aan schaken: je ziet direct wat je kan doen, cognitie is holistisch).

61

### **Geen schadetolerantie**

- Een computer heeft geen schadetolerantie: als er één stukje van de computer stuk is, dan werkt het hele systeem niet meer;
- Als er een neuron wegvalt dan werken de hersenen nog steeds.

62

### **Het software level is niet autonoom**

- Als je de geest als de software ziet die draait op het brein (de hardware), dan is het idee dat die software ook op andere hardware zou kunnen draaien – bijvoorbeeld op de constitutie van een alien met een silicium brein;
- Maar lichaam, brein & geest zijn tegelijkertijd geëvolueerd...

63

- En evolutie werkt met wat ze al heeft;
- Je hebt een vis met een blaas, en dat is waarvan je “start” als je een landdier wil “maken;”
- Je kan die blaas gebruiken om er longen uit te evolveren;
- Hetzelfde mechanisme heb je nodig voor de evolutie van brein & geest, waardoor het onwaarschijnlijk is dat de geest autonoom is.

64

### **Conclusie / vooruitblik**

- M.a.w. traditioneel / computer functionalisme is biologisch erg onrealistisch;
- Volgende college zullen we een aanpassing zien, waarbij het computemodel voor het brein nog steeds biologisch onrealistisch is, maar waarbij wel een aantal verbeteringen is aangebracht – het heeft niet meer de problemen die we zojuist opgesomd hebben.

65

### **Wat voor vragen kan je over deze stof verwachten?**

66

## **Tot volgende keer**

### **Vragen?**

Mail naar [a.j.p.w.dooremalen@uva.nl](mailto:a.j.p.w.dooremalen@uva.nl)

Of stel ze op het discussieforum

### **Inhoudelijke discussie over de stof?**

Bekijk het forum op Blackboard

67