

Oriëntatie Kunstmatige Intelligentie

Filosofie en
Architecturen voor Cognitie
Niels Taatgen

Inhoud

Intelligentie Bewustzijn Gedrag
“Ziel” “Geest” “Mind” “Beliefs”

Verklaring

Hersencellen Neurotransmitters
Chemische signalen

- Dualisme
Monisme
- Holisme
Reductionisme

Inhoud

Intelligentie Bewustzijn Gedrag
“Ziel” “Geest” “Mind” “Beliefs”

Syntactisch niveau
Formeel niveau

Hersencellen Neurotransmitters
Chemische signalen

- Functionalisme
- Fysicalisme
- Superveniëntie
- Turing Machine versus de Chinese Kamer

Inhoud

Intelligentie Bewustzijn Gedrag
“Ziel” “Geest” “Mind” “Beliefs”

Symbolen of Neurale netwerken
of logica of dynamische systemen?

Hersencellen Neurotransmitters
Chemische signalen

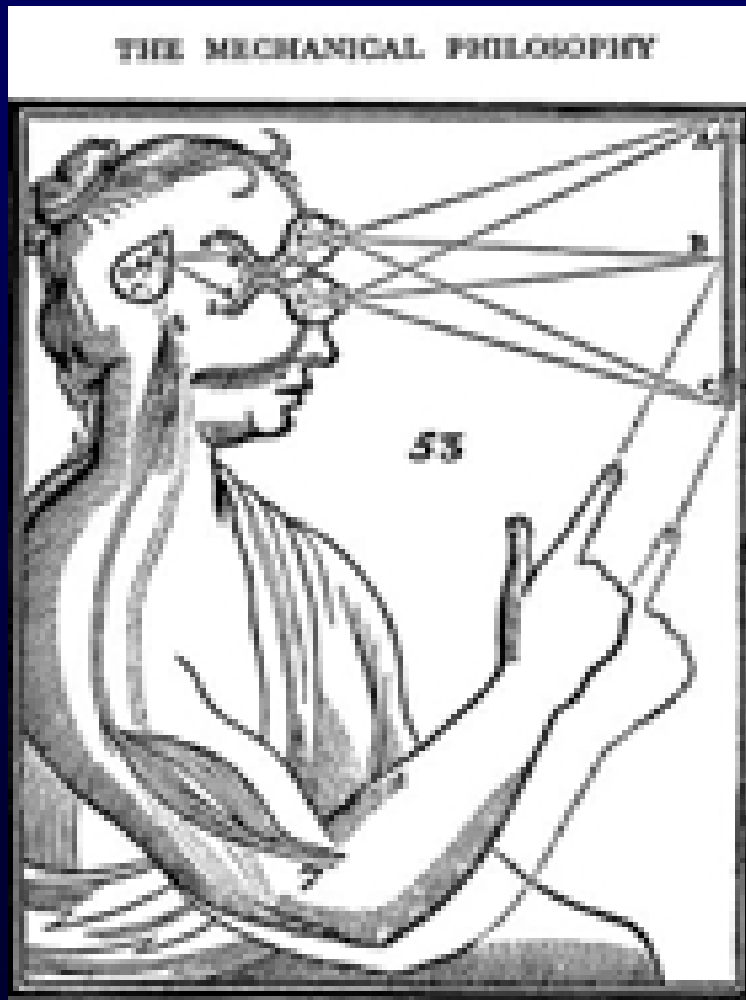
- Symbolisme
- Neurale netwerken
- Hybride Systemen
- Dynamische systemen

Dualisme van Descartes



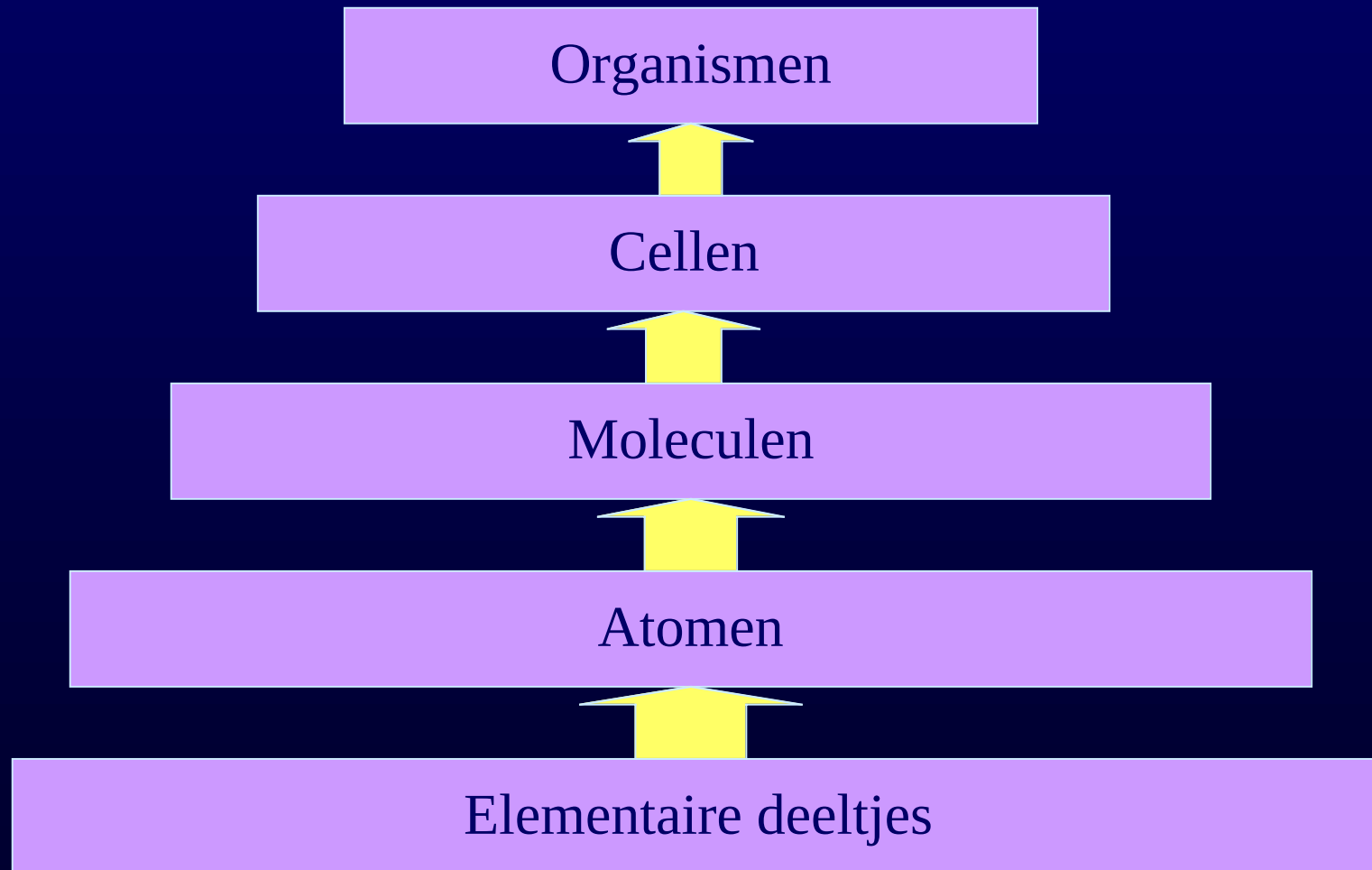
- Lichaam en Geest zijn verschillende dingen
- Lichaam is stoffelijk
- Geest is niet-stoffelijk
- Communiceert met het lichaam via de Pijnappelklier in de hersenen

Problemen met het Dualisme



- Geest is moeilijk wetenschappelijk te bestuderen
- Hoe communiceert iets niet-stoffelijks met iets stoffelijks?

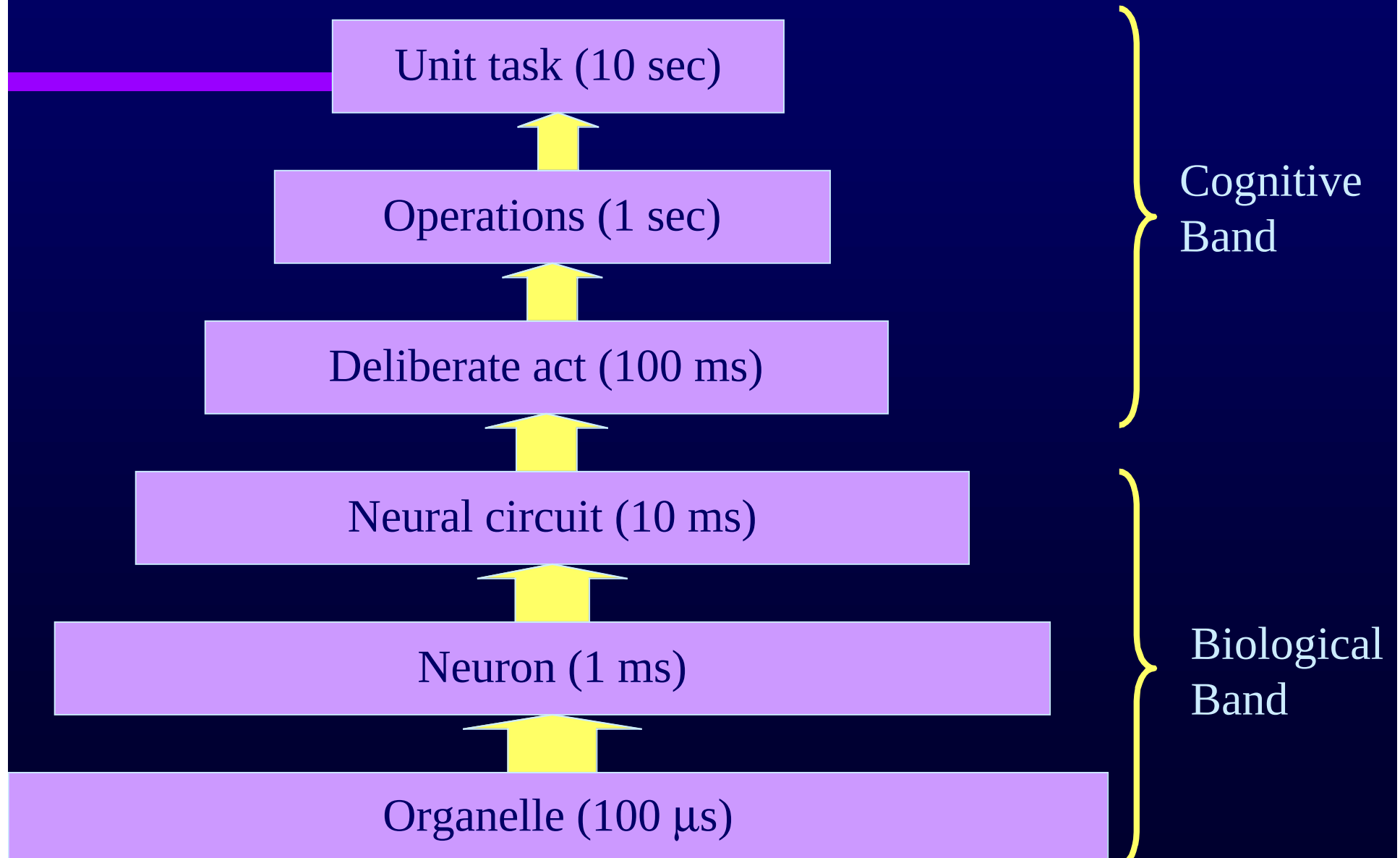
Gelaagde werkelijkheid



Vraag: kan dit ook met Lichaam en Geest?

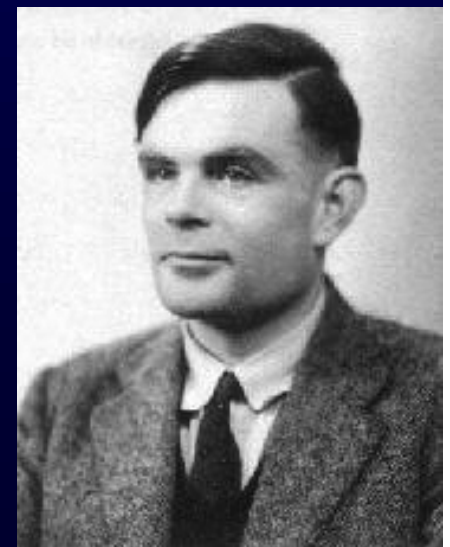
- Reductionisme:
Elk niveau van beschrijving kan volledig verklaard worden vanuit het onderliggende niveau
- Holisme:
Het geheel is meer dan de som der delen

Allen Newell's voorstel

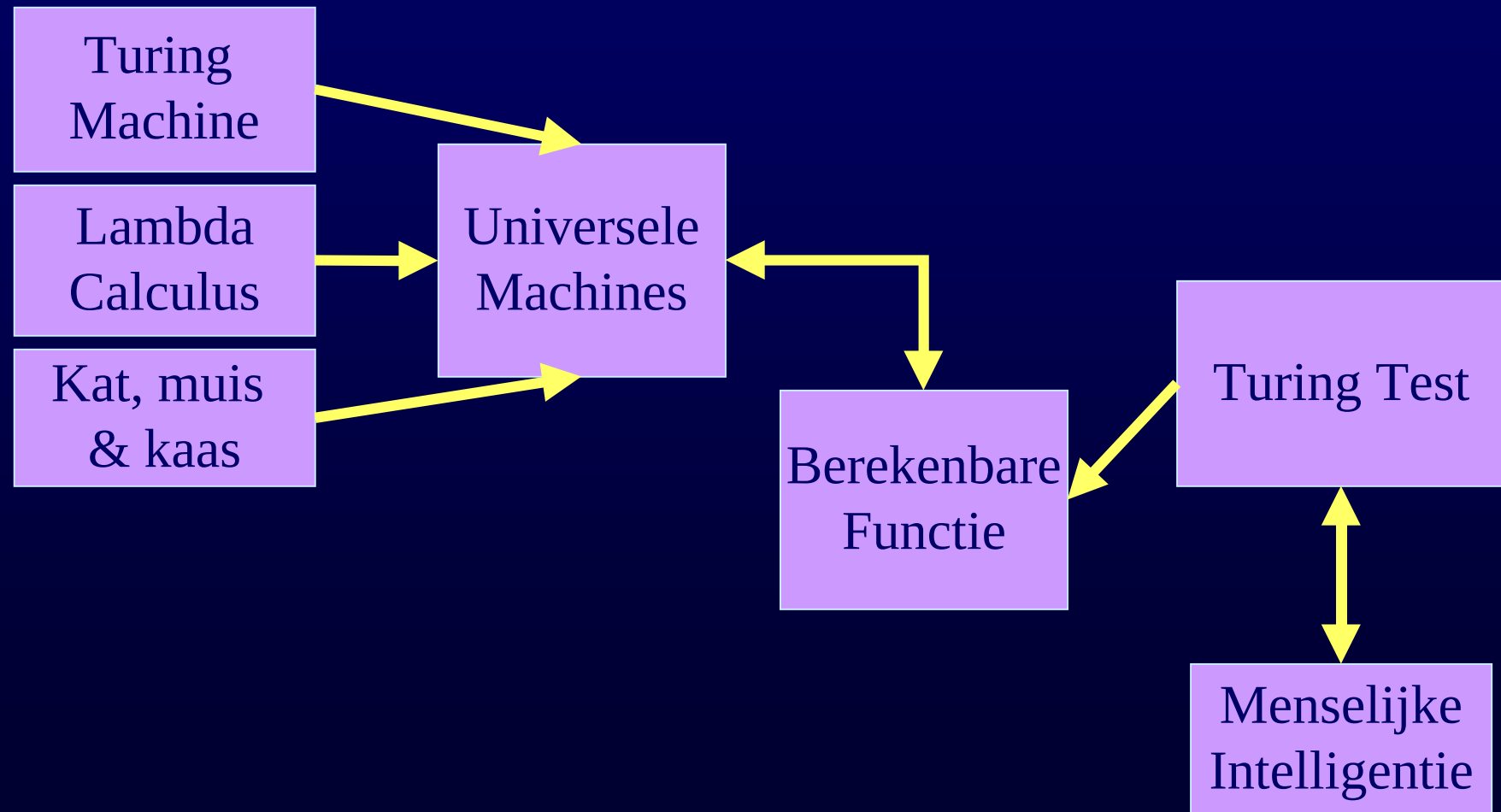


Vraag

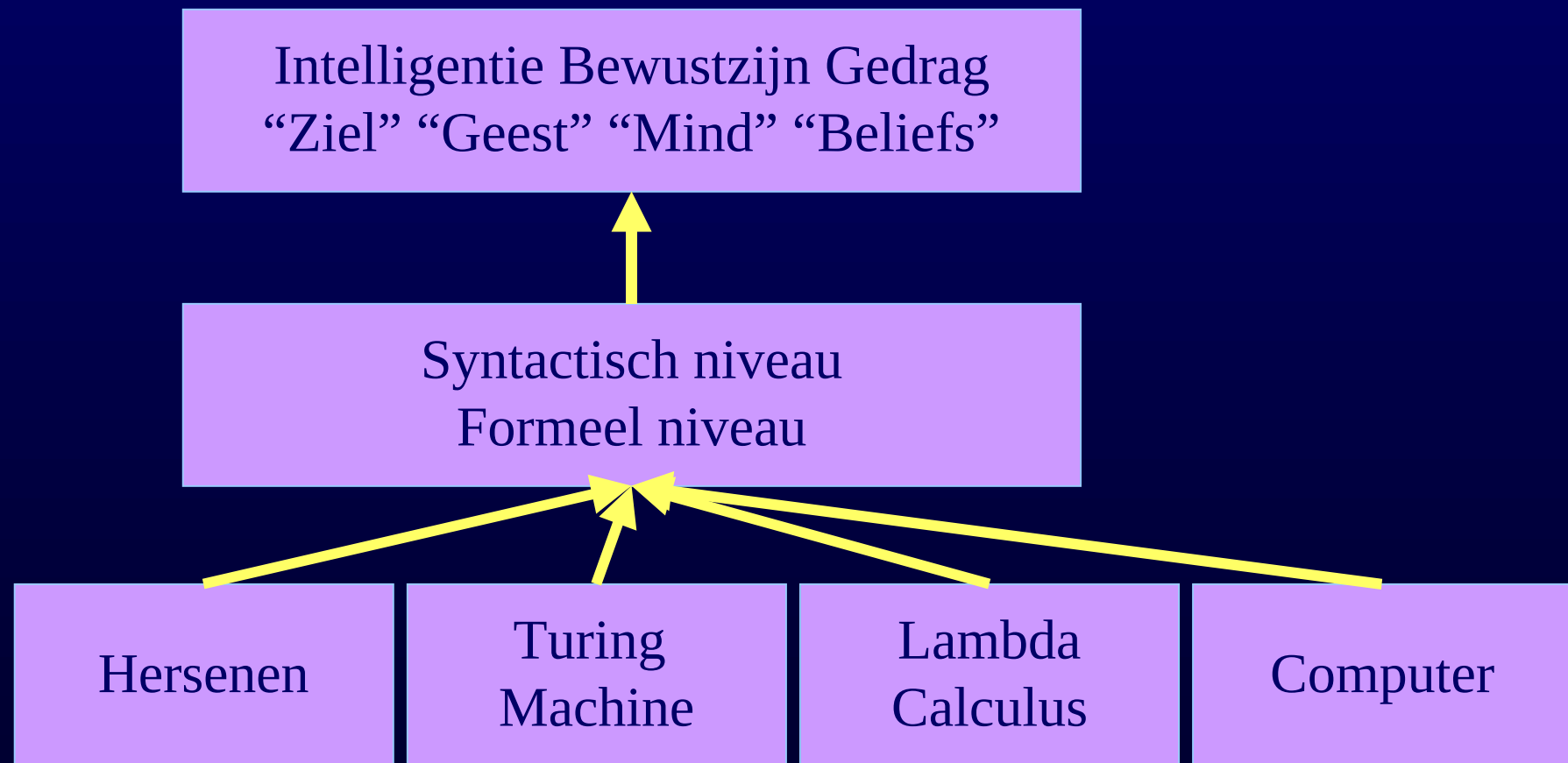
- Is het nodig om alle niveaus tussen neuron en cognitie te doorlopen om cognitie te verklaren?
- Volgens Turing (college 1) is dit niet nodig: principe van Universele Computatie



Turing's redenering



Verklaring voor cognitie



“Multiple Realizability”

- Als het niet uitmaakt of intelligentie wordt voortgebracht door hersenen of een computer, dan lijkt het of het niet langer nodig is om de hersenen te bestuderen.
- Dit idee ligt ten grondslag aan het Functionalisme

Een zonnig landschap



- Is “zonnig” een eigenschap van de verf?
- Het wordt wel veroorzaakt door de verf.

Functionalisme

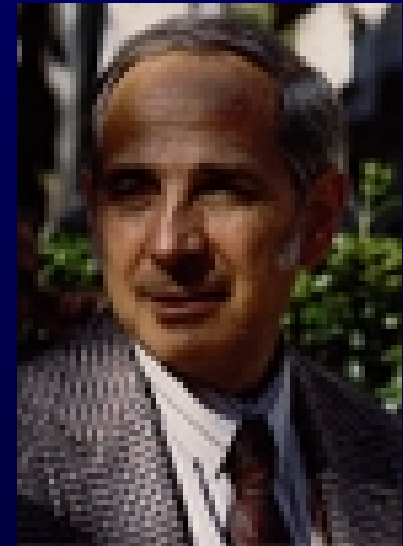
- Mentale eigenschappen (bewustzijn, intelligentie, etc.) worden veroorzaakt door Herseneigenschappen
- Maar mentale eigenschappen zijn niet direct te verklaren vanuit herseneigenschappen
- Boek: de geest superveniëert op het brein

Turing Machine Functionalisme

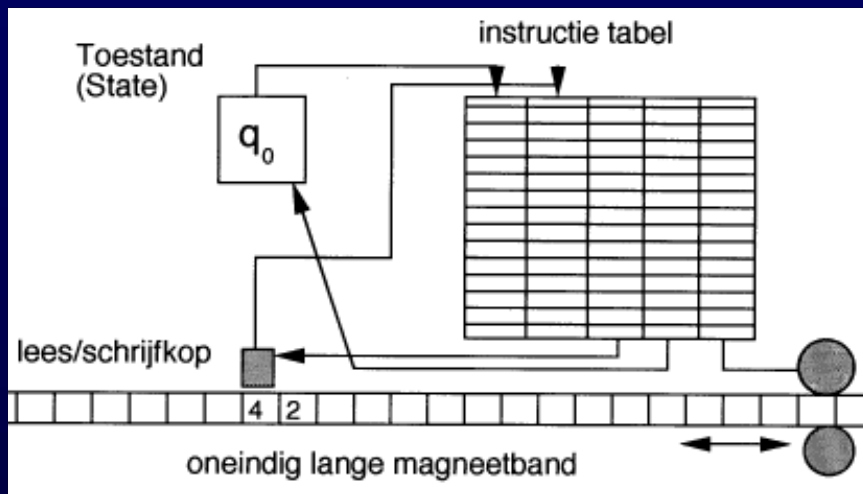
- De menselijke geest wordt veroorzaakt door een computationeel proces, dat net zo goed door een computer uitgevoerd zou kunnen worden

Bezwaar: John Searle

- Searle is het niet eens met deze stelling
- Hij ontwikkelde het z.g. Chinese Kamer argument om aan te tonen dat een intelligente computer onmogelijk is

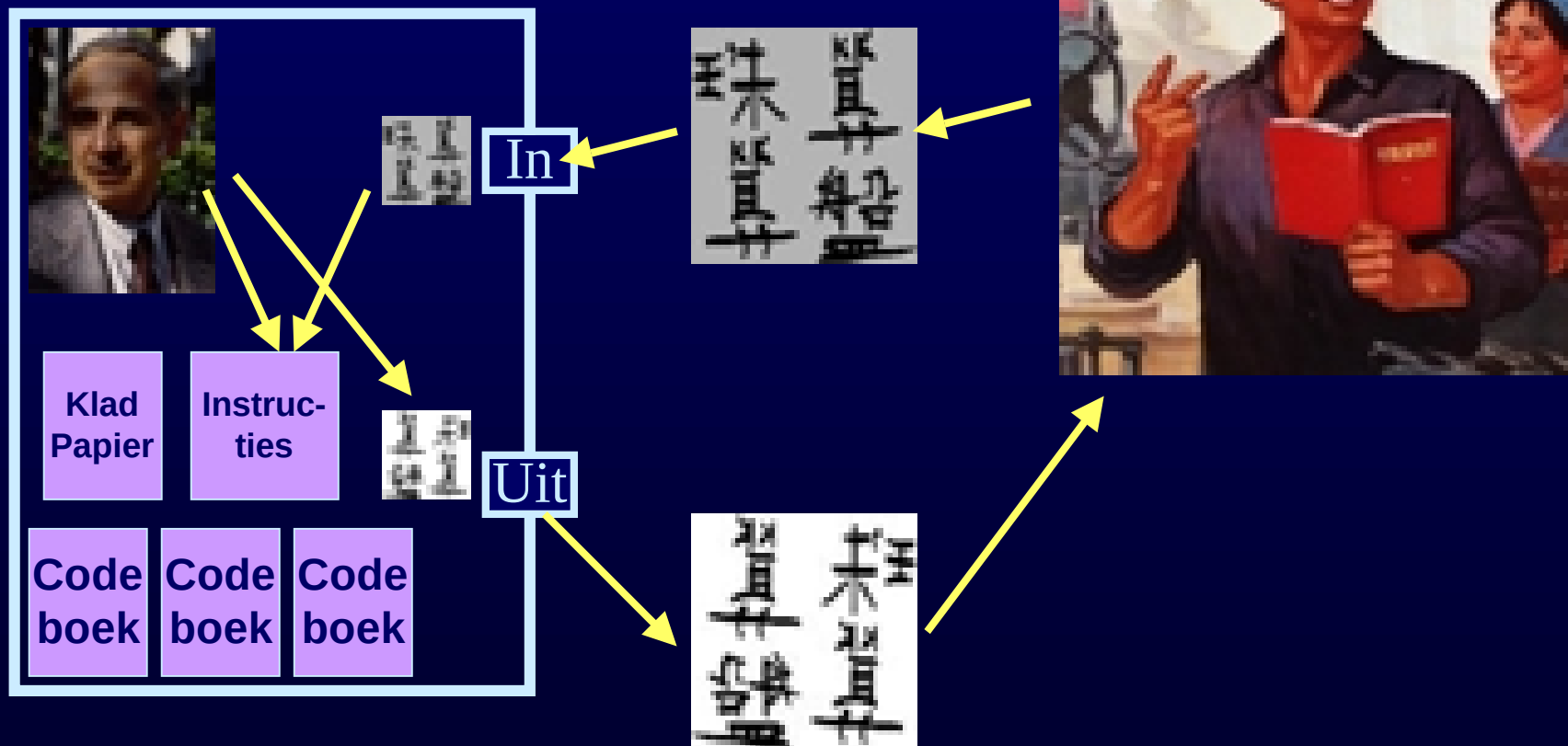


Chinese Kamer argument



- De Turing Machine zou volgens het TM functionalisme intelligent moeten kunnen zijn, gegeven het juiste programma
- Echter: je kunt de TM machine ook best met de hand simuleren

Chinese Kamer argument

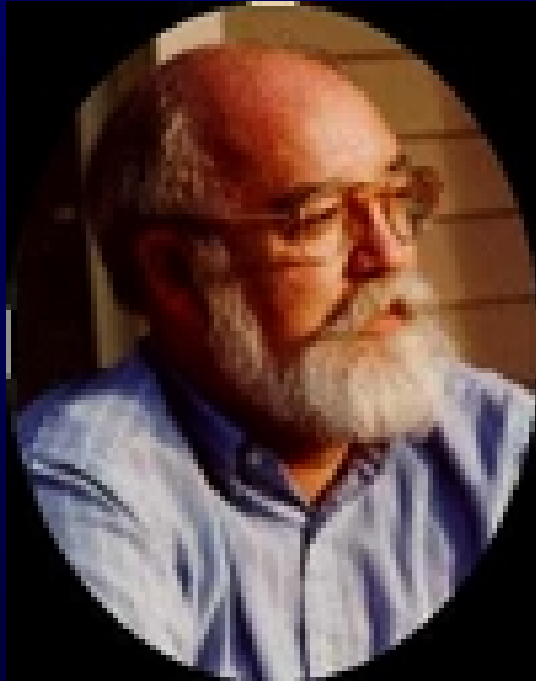


<i>Input</i>	<i>you ± rules</i>	<i>Output</i>
(六) 未. 庚. 酉. (一)	If 亥. then 酉. 壬. 	丙. 辛. (二) 午. 辰.

Searle's redenering

- Uit een formeel systeem (Searle refereert hiernaar als “syntax”) kan nooit betekenis voortkomen (“semantics”)
- Maar Searle doet wel een beroep op ons voorstellingsvermogen, en dit is niet altijd even betrouwbaar (algemeen probleem van Gedachtenexperimenten)
- Voorbeeld van Churchland: schudden van een magneet

Daniel Dennett



- De “eenheid” van menselijk bewustzijn is een illusie
- Vergelijk het met een mierencolonie

Wat is het juiste niveau?

Intelligentie Bewustzijn Gedrag
“Ziel” “Geest” “Mind” “Beliefs”

Symbolen of Neurale netwerken
of logica of dynamische systemen?

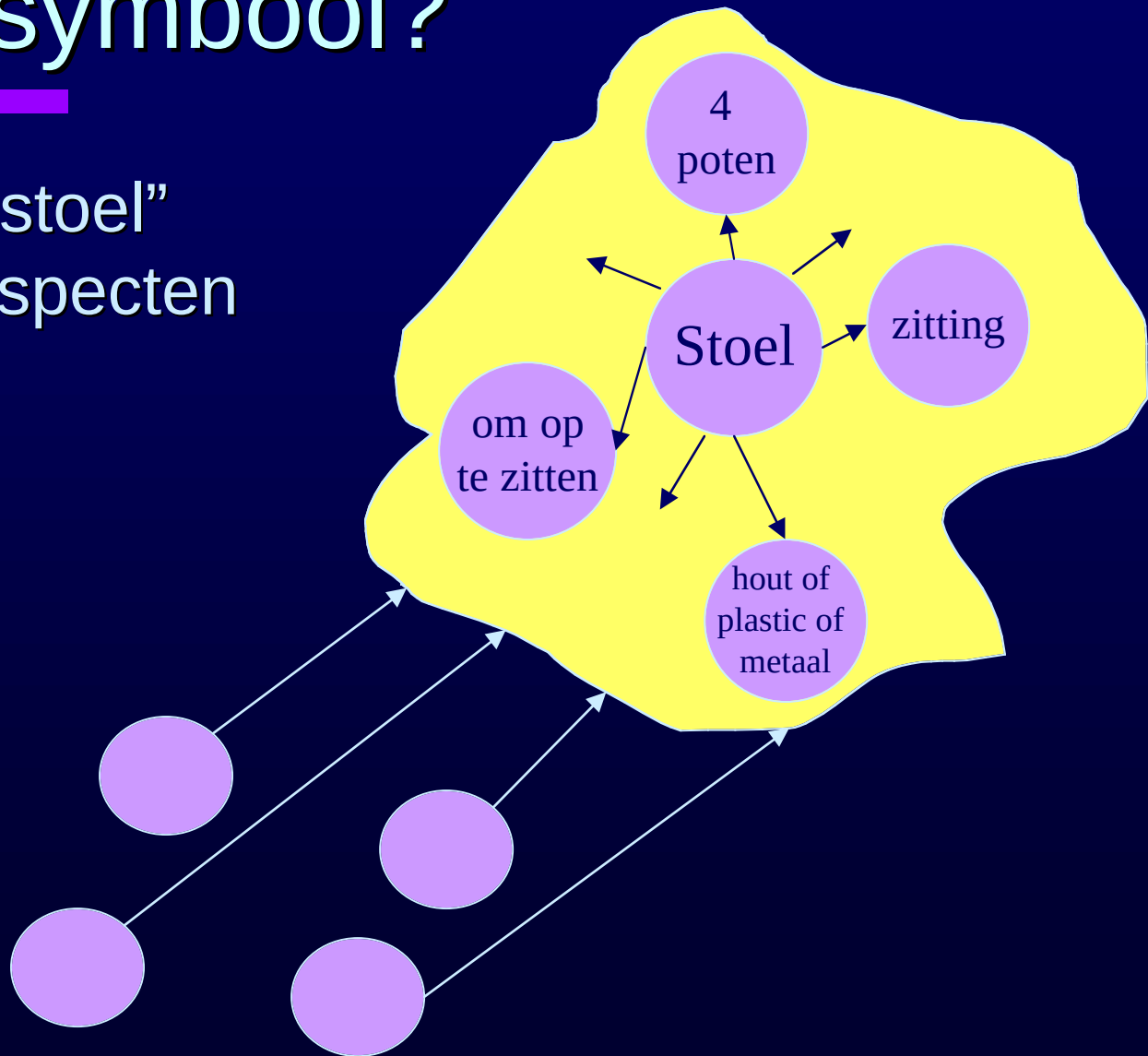
Hersencellen Neurotransmitters
Chemische signalen

- Symbolisme zit heel dicht bij het gedrag
- Connectionisme zit dicht bij de hersenen
- Hybride systemen zoeken een compromis

Symbolisme: wat is een symbool?

- Het begrip “stoel” heeft vele aspecten
- Vorm
- Functie
- etc.

“Vier Stoelen”



Symbolisme

- Voordeel:
Symbolen zijn makkelijk te gebruiken, vier stoelen is geen probleem
- Oftwel: symbolen zijn productief
- Nadeel:
Een stoel is altijd hetzelfde, ook in een andere context

Connectionisme (neurale netwerken)

- Symbolen zijn niet fundamenteel maar zijn opgebouwd uit kleinere stukjes (subsymbolen)

Kop koffie

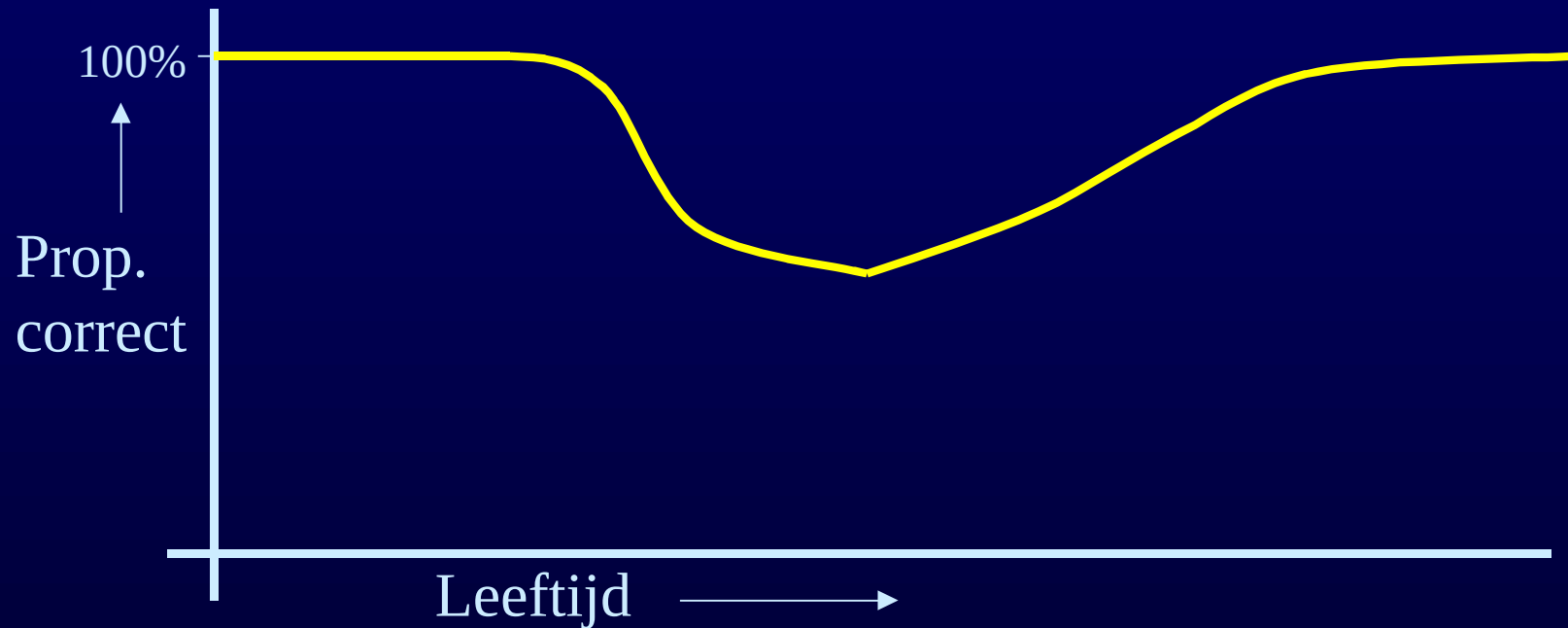
- + Rechtopstaand voorwerp waar iets in kan
- + Hete vloeistof
- Glas in aanraking met hout
- + Aangebrande geur
- Langwerpig zilveren voorwerp

Hybride systemen

- Symbolen zijn fundamenteel, maar hebben subsymbolische aspecten
- Voorbeeld: ACT-R heeft symbolen (chunks in declaratief geheugen), maar deze symbolen hebben een activatie, en hebben verbindingen met elkaar

Voorbeeld: het Verleden Tijd debat

Waar gaat het om?



- Stage 1: break-broke
- Stage 2: break-broke
- Stage 3: break-broke

Symbolische theorie

- Regelmatige werkwoorden worden door een regel gerepresenteerd, en onregelmatige werkwoorden worden als uitzondering hierop onthouden
- De U-vorm in het leren begint op het moment dat het kind de regel ontdekt

Probleem

- Hoe wordt die regel dan ontdekt?

Connectionistische Theorie

- Alle kennis wordt op één manier gerepresenteerd, namelijk in een neurale netwerk
- De U-vorm wordt verklaard doordat het vocabulaire groeit

Problemen

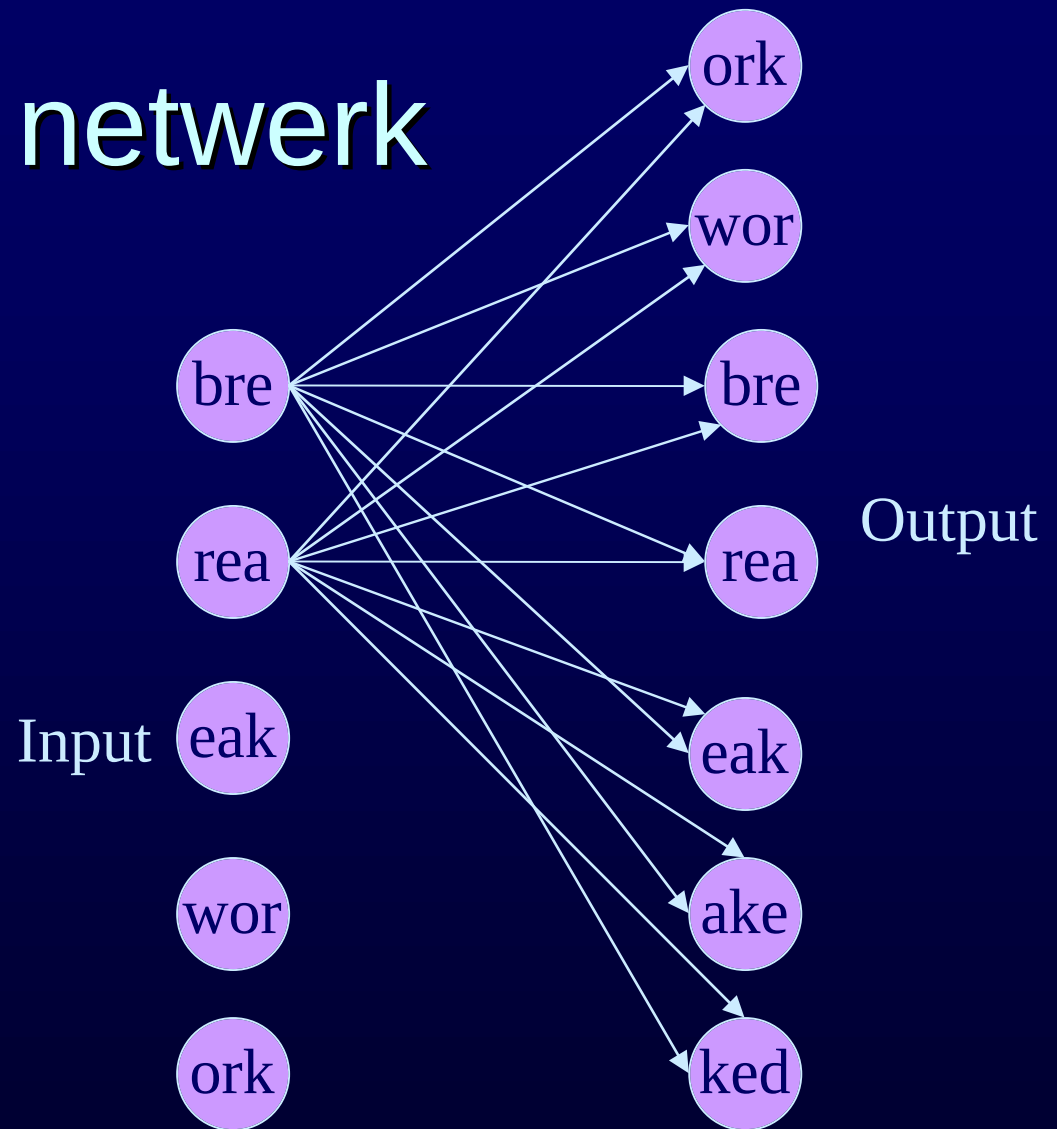
- Regelmatige werkwoorden maken slechts 30% uit van het gebruik
- Plotselinge groei in vocabulaire valt niet samen met begin van overgeneralisatie

Neuraal netwerk

- Woorden worden gerepresenteerd in “Wickelfeatures”

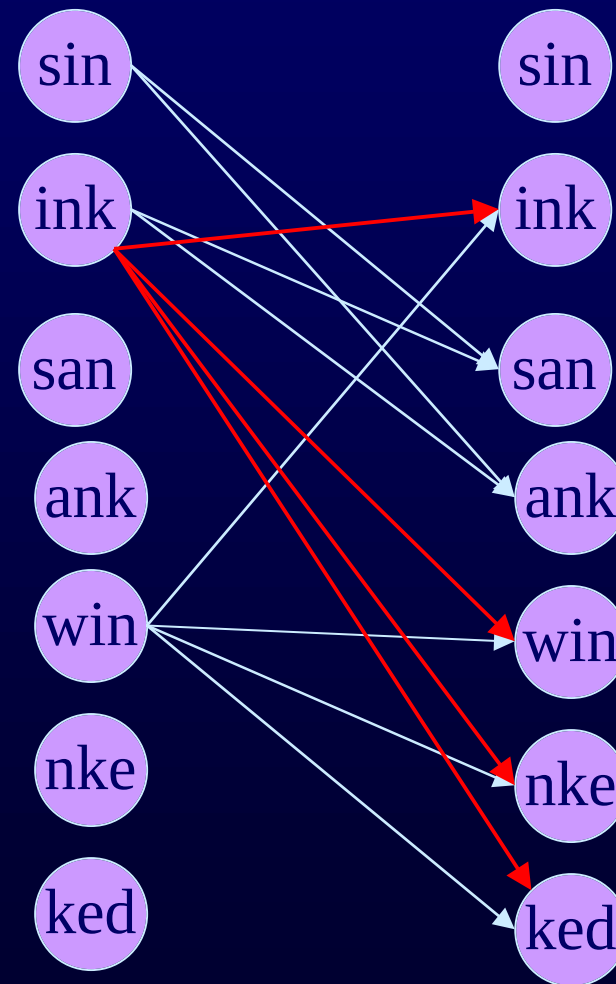
- Break wordt:

- bre
- rea
- eak



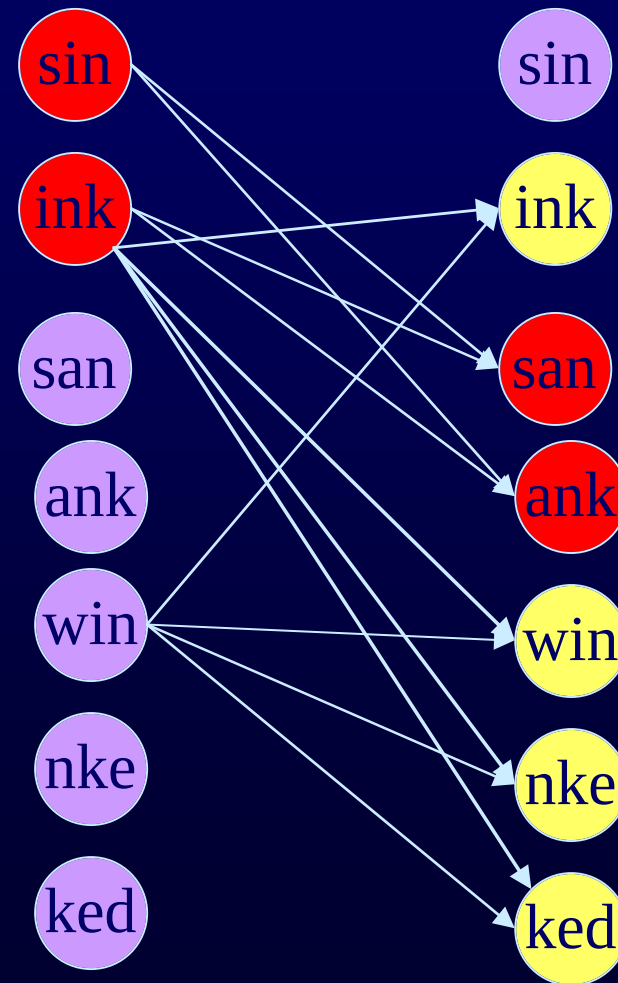
Neuraal netwerk

- Het leren van sink/sank
- en wink winked



Neuraal netwerk

- Vraag het netwerk: wat is de verleden tijd van sink?
- Nu gaat het nog goed, maar als we ook nog blink-blinked leren, en stink-stinked, dan gaat het fout



Probleem voor zowel symbolische als connectionistische modellen

- Kinderen kunnen verleden tijden leren zonder dat ze feedback nodig hebben

Het hybride ACT-R model

- Maakt gebruik van twee representaties: regels en voorbeelden
- Het maken van een verleden tijd gebeurt door:
 - Het in het geheugen opzoeken van het antwoord
 - Het toepassen van de regel (voeg -ed toe aan de stam)

Aanname: onregelmatige werkwoorden hebben voordelen

- Ze zijn vaak korter, dus dat is voordeliger om uit te spreken
- Regelmatige werkwoorden zijn vaak fonetisch onregelmatig
- Onregelmatige werkwoorden kunnen dus sneller zijn, maar je moet ze wel allemaal onthouden

Voordeel regelmatige werkwoorden

- Is erg productief: je hebt maar één regel nodig om alle mogelijke werkwoorden te vervoegen

Wat is dus slimmer: regelmatig of onregelmatig?

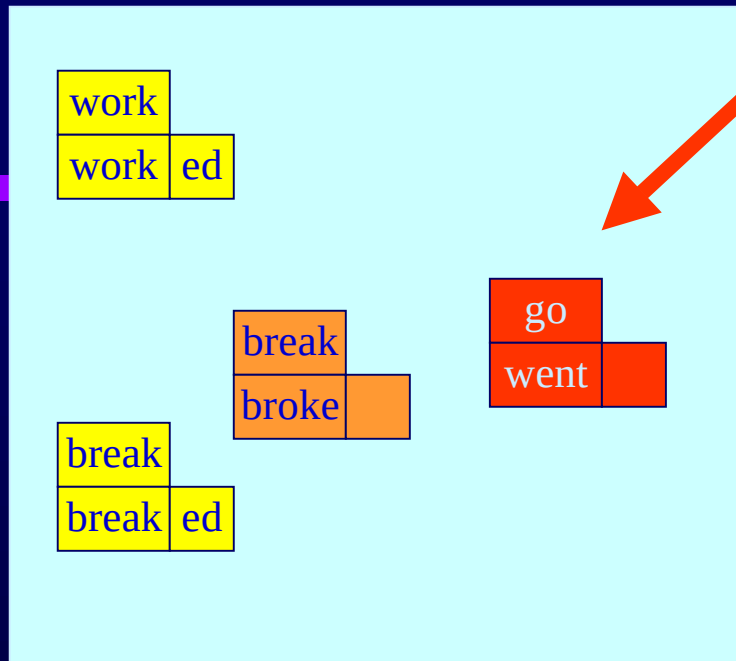
■ Rationele analyse

- Hoogfrequente woorden komen vaak voor, dus die kunnen we wel onthouden
- Laagfrequente woorden komen zo weinig voor dat je beter een regel er op na kunt houden om ze allemaal te produceren

Hoe werkt het model: 4 strategieën

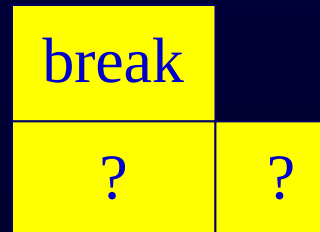
- Probeer een verleden tijd uit het geheugen te halen
- Gebruik analogie om een verleden tijd te maken uit een oud voorbeeld
- Gebruik gewoon de stam als verleden tijd, dus doe gewoon niets
- In het begin kent het model de regel (voeg ed toe) nog niet, maar zal deze later leren

Declarative Memory

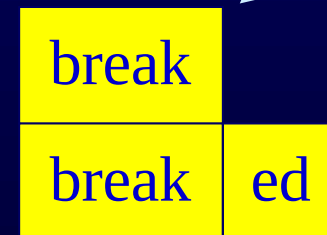


Environment
(i.e., parents)

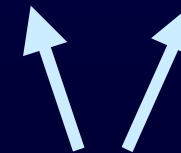
Current goal



stem



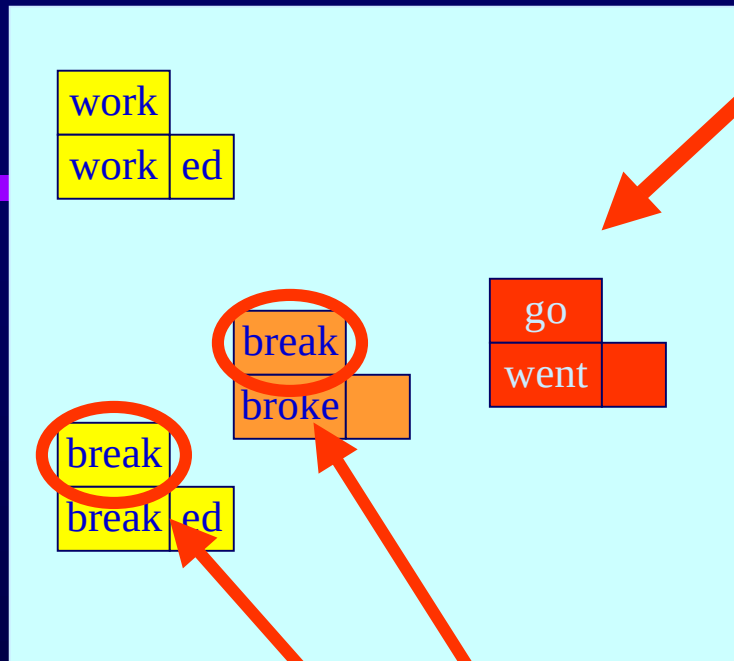
past tense



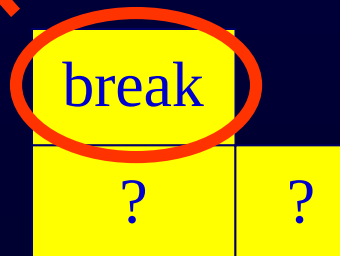
Declarative Memory

Environment
(i.e., parents)

Strategy 1: Retrieval



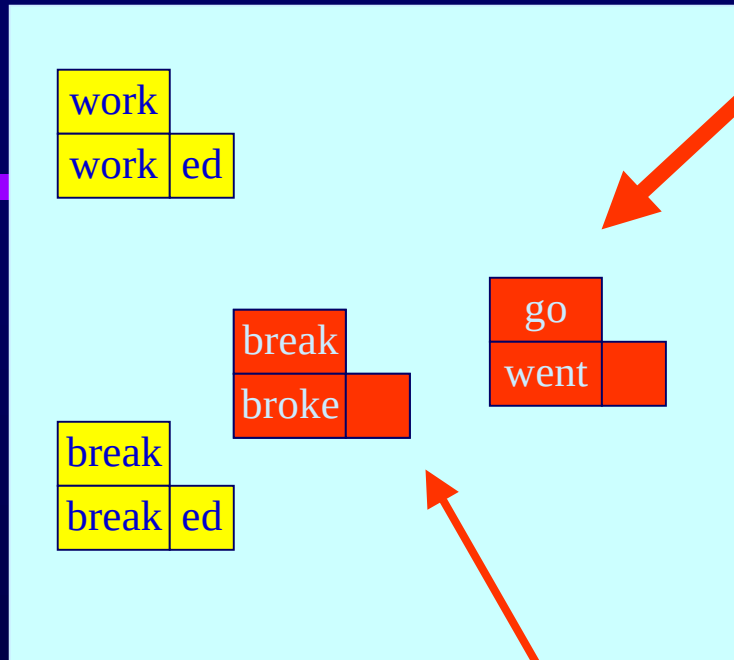
Current goal



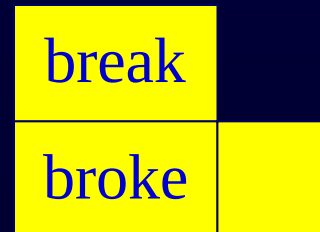
Declarative Memory

Environment
(i.e., parents)

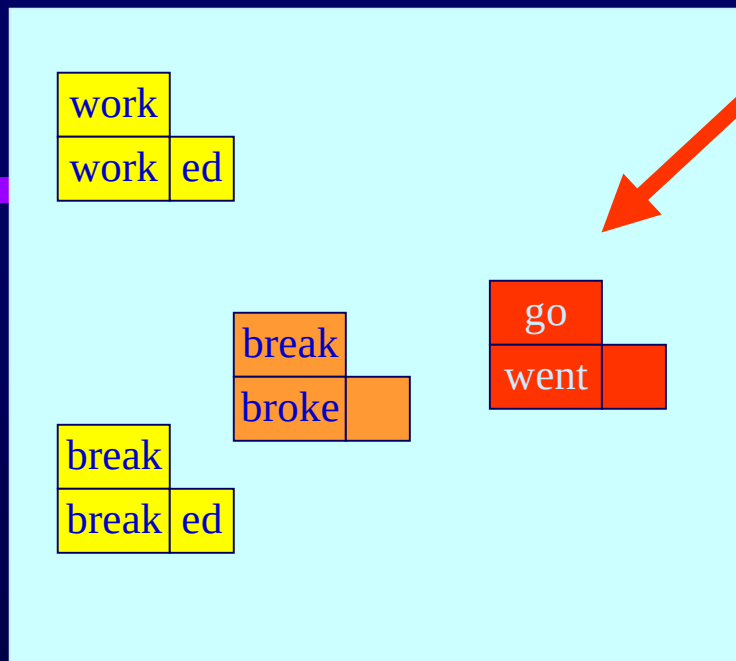
Strategy 1: Retrieval



Popped goal strengthens
Current goal declarative memory



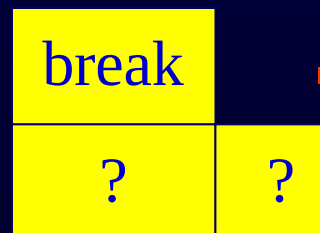
Declarative Memory



Environment
(i.e., parents)

Strategy 1: Retrieval
Strategy 2: Do nothing

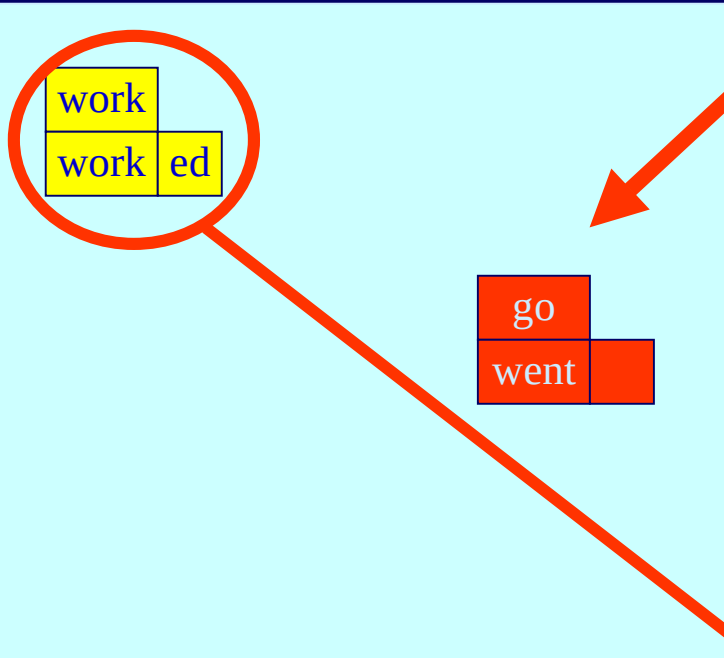
Current goal



“Yesterday break”

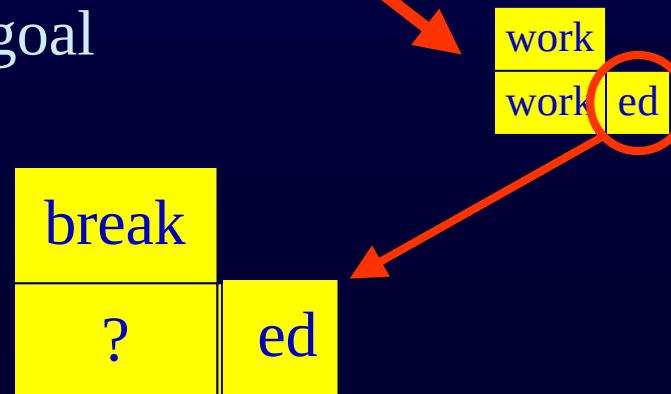
Declarative Memory

Environment
(i.e., parents)



Strategy 1: Retrieval
Strategy 2: Do nothing
Strategy 3: Analogy
(pattern matching)

Current goal



Declarative Memory

work	
work	ed

go	
went	

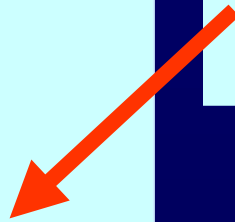
Environment
(i.e., parents)

Strategy 1: Retrieval
Strategy 2: Do nothing
Strategy 3: Analogy
(pattern matching)

Current goal

work	
work	ed

break	
break	ed



De regel wordt geleerd door de analogie strategie te specialiseren

Retrieve example

Fill empty suffix slot
with value from example

Paste tense of
work is work-ed

Add -ed to stem

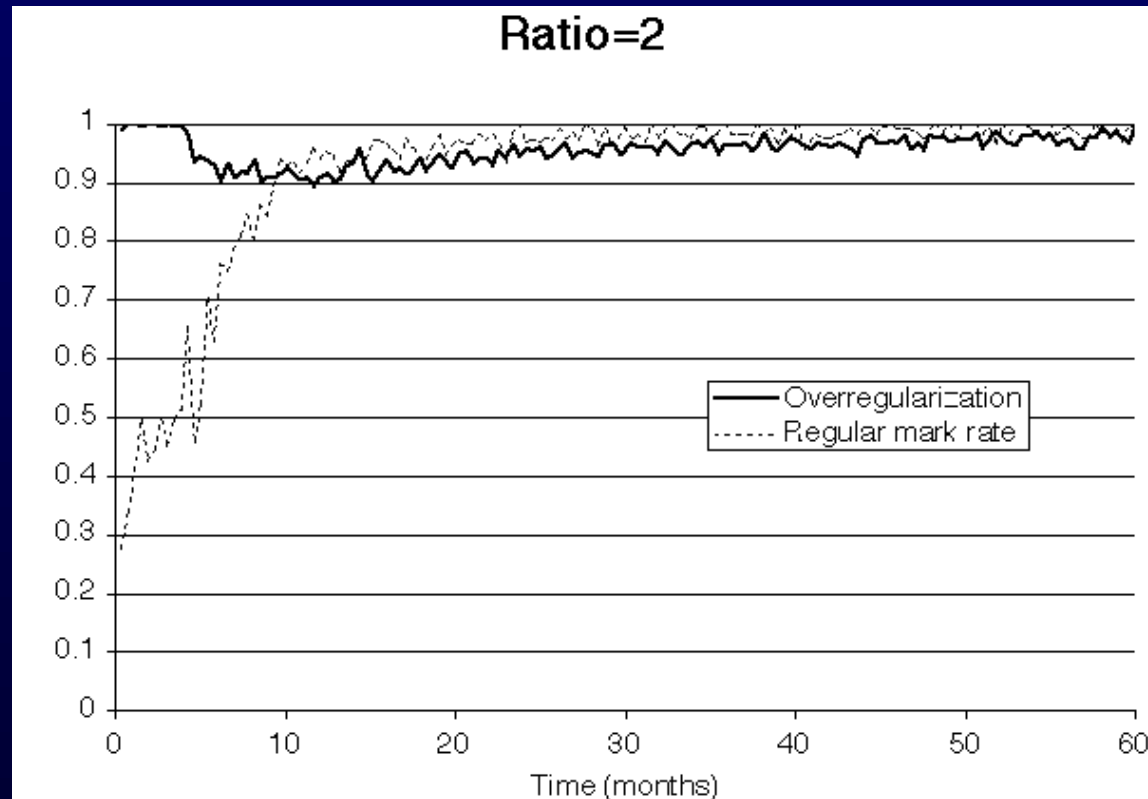
Procedural

Declarative

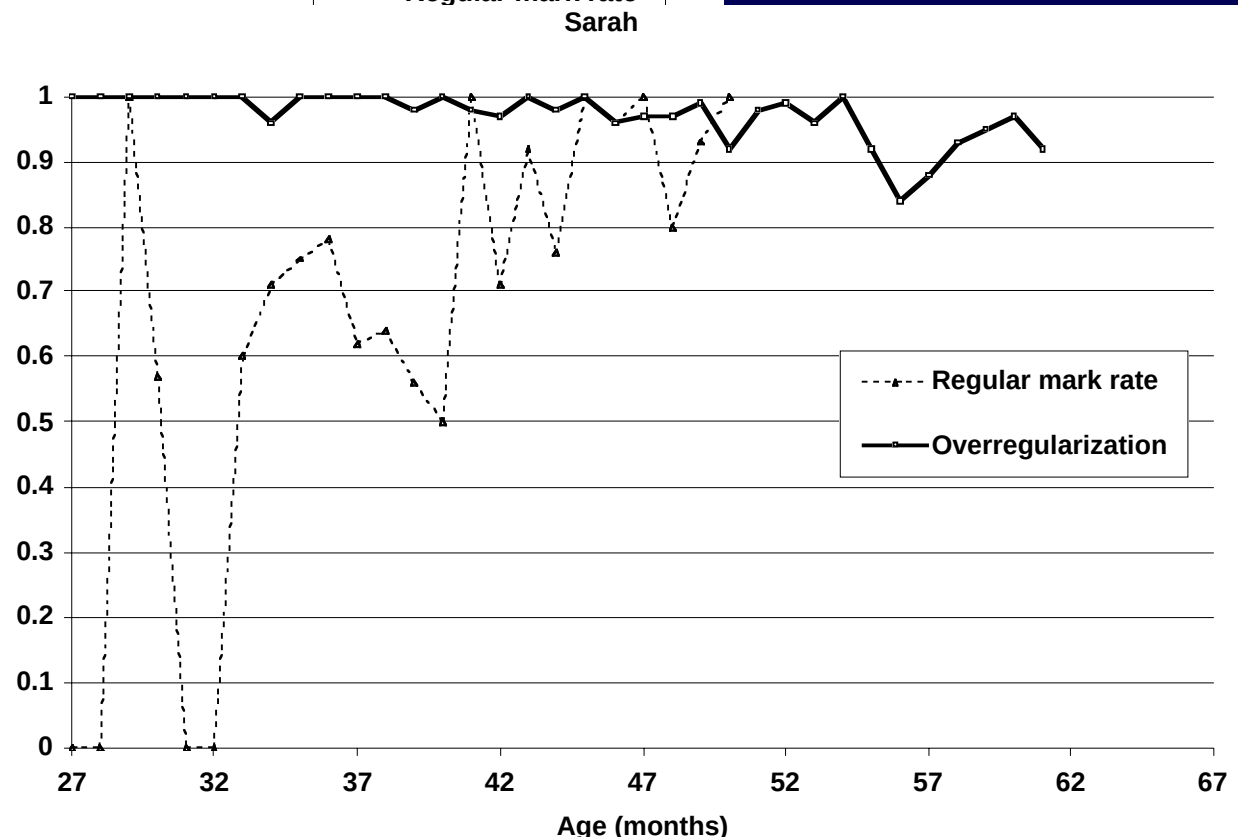
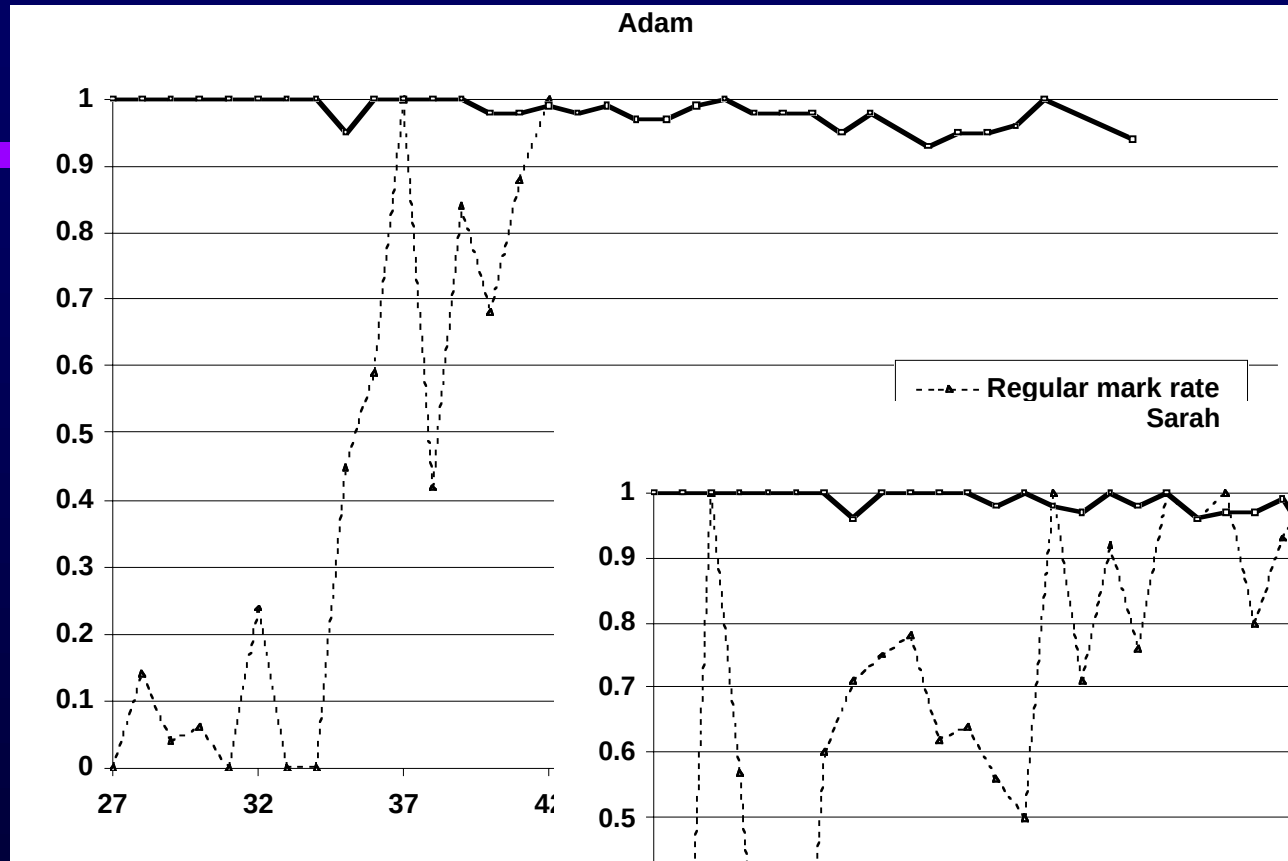
Wat gebeurt er met de strategieën

- Opzoeken in het geheugen werkt in het begin niet zo goed, maar later wel.
- Analogie is niet zo goed, want kostbaar (je moet meerdere denkstappen ondernemen)
- De regelmatige regel, die pas later geleerd wordt, is wel goed, maar niet zo goed als opzoeken in het geheugen uiteindelijk wordt.

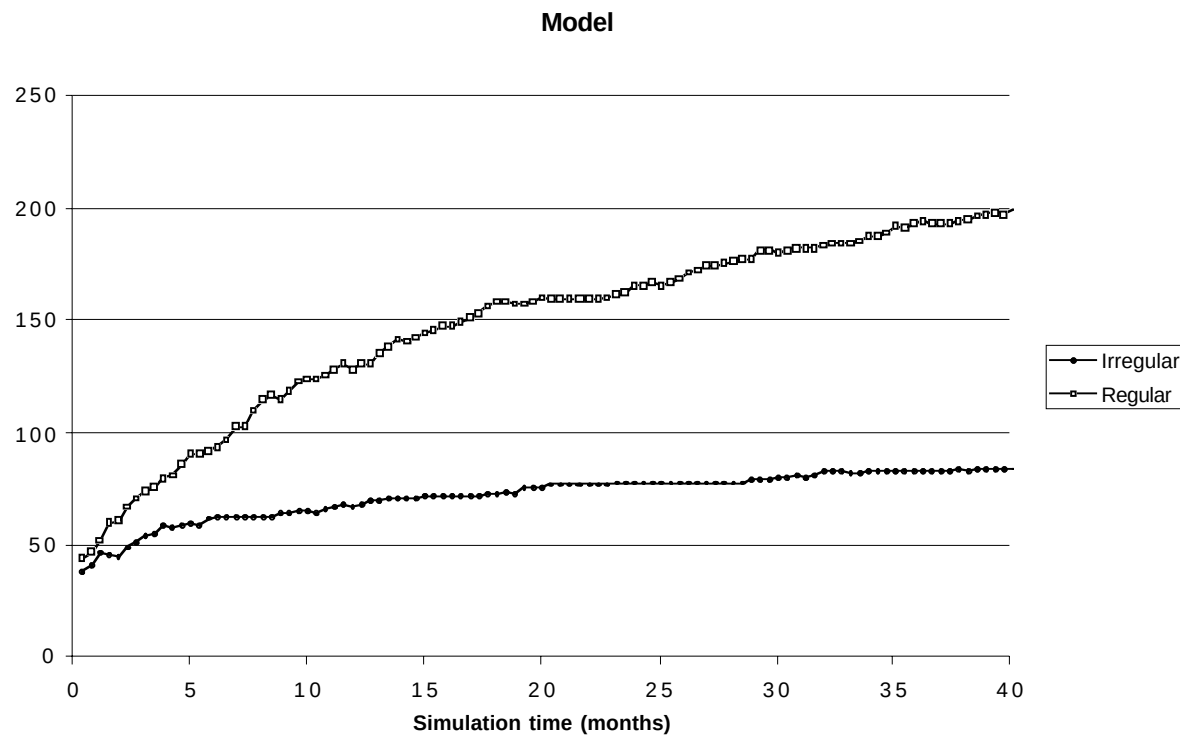
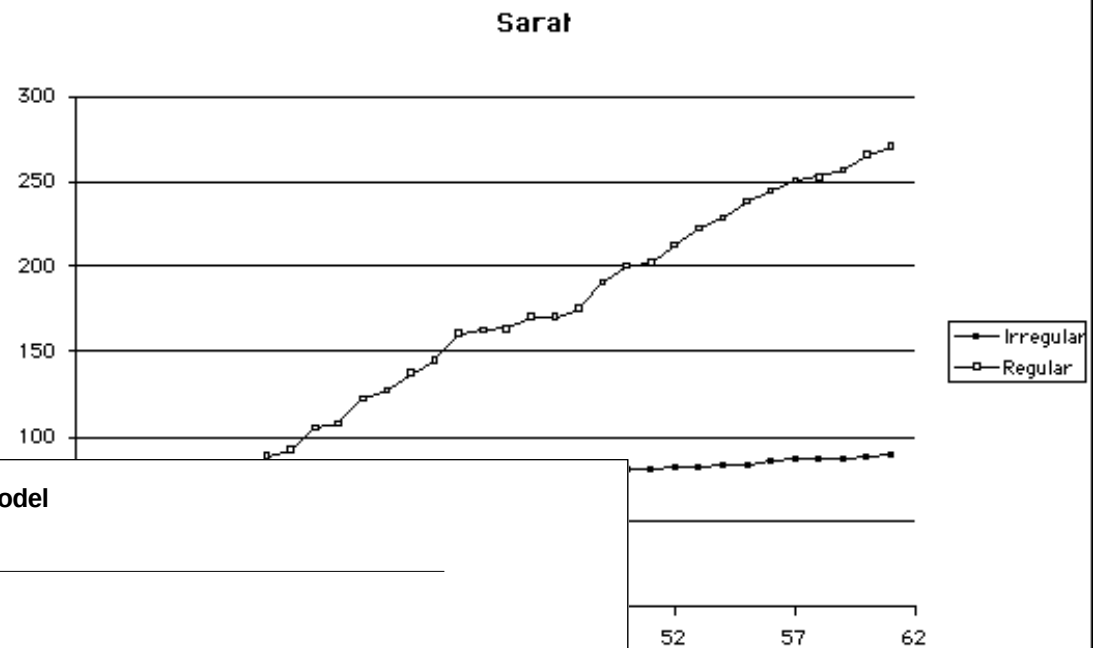
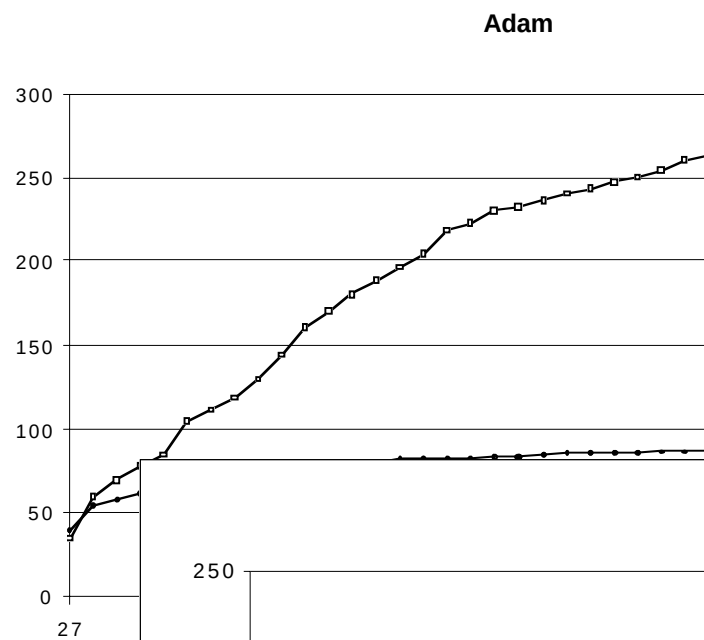
U-shaped learning



Emperical data

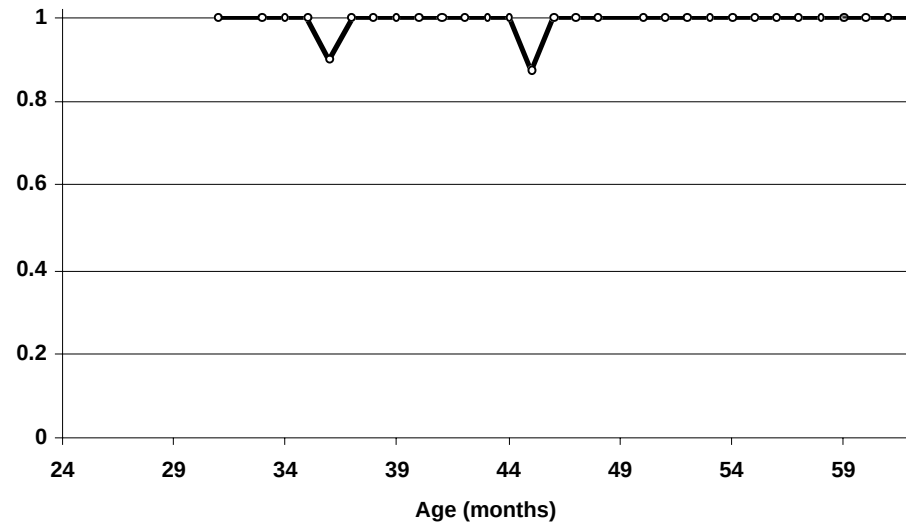


Vocabulary growth

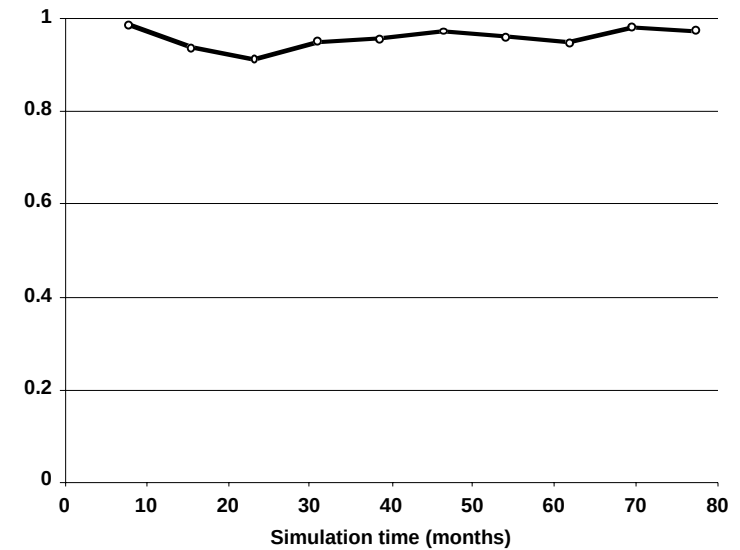


Individual words

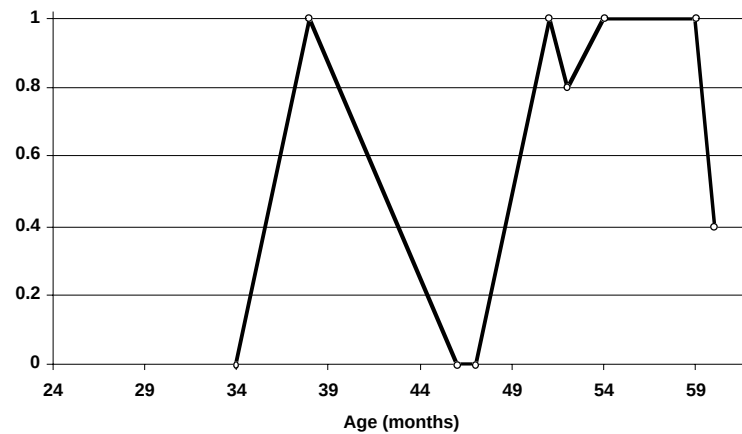
Abe: Say



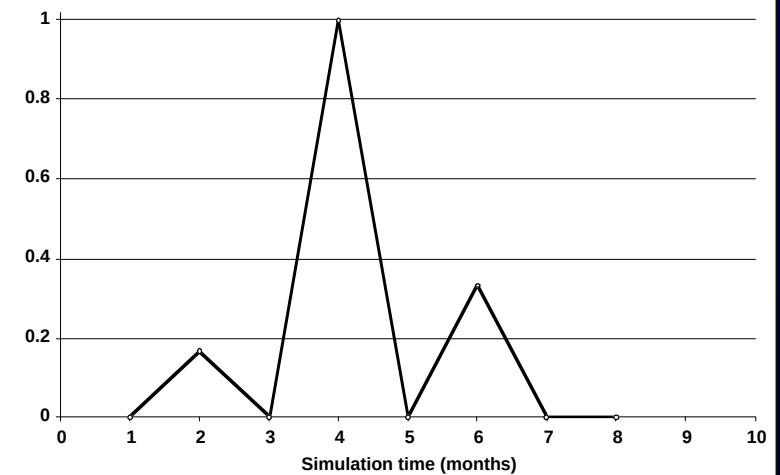
Say



Abe: Win



Tear



Conclusies

- Het hybride model lost een aantal problemen van zowel het symbolische als het connectionistische model
- Het kan leren zonder feedback op de eigen productie
- Het te generaliseren naar andere talen, bijvoorbeeld het Duitse meervoud