



De mens. Een machine?

Het argument van J.R. Lucas
tegen het mechanisme

G.J.E. Rutten



De verhouding tussen mens en machine

- **Mechanisme** (*materialisme, sciëntisme*)
 - De mens is niets meer dan een complexe machine
 - Wij zijn in essentie gedetermineerde automaten
 - Een dualistische visie op de mens is onhoudbaar
- **Dualisme**
 - De mens is meer dan een gedetermineerde automaat
 - Dit ‘meer’ betreft het domein van het mentale
 - Het mentale omvat bewustzijn, gevoelens en vrije wil

Twee klassieke argumenten voor dualisme

1. Inherent aan de menselijke conditie zijn fenomenen als zelfbewustzijn, gevoelens en vrije wil. Deze intrinsiek bij ons als mens behorende fenomenen kunnen niet veroorzaakt worden door de werking van automaten, hoe complex ook. Daarom kan de mens geen machine zijn.
2. Stel dat het dualisme incorrect is. In dat geval ben ik mijn lichaam. We kunnen dus stellen dat ‘Ik = Mijn lichaam’. De eigenschappen die ‘ik’ bezit moeten dus exact overeenkomen met de eigenschappen die ‘Mijn lichaam’ bezit. Nu heb ‘Ik’ de eigenschap dat ik niet consistent kan denken dat ik er niet ben. ‘Mijn lichaam’ heeft deze eigenschap niet. Ik kan immers consistent denken dat mijn lichaam er niet is. We vinden dus een tegenspraak. Ik moet dus meer zijn dan mijn lichaam. Hieruit volgt dat het dualisme correct is.

Zijn deze argumenten overtuigend? Waarom (niet)?

Bestaan er nog andere klassieke argumenten voor dualisme?

Een hedendaags argument voor dualisme

- De filosoof **J.R. Lucas** heeft het in zijn artikel 'Minds, Machines, and Gödel' uit 1961 over een heel andere boeg gegooid.
- Lucas bedacht een scherpzinnig filosofisch argument om aan te tonen dat wij meer zijn dan gedetermineerde automaten.
- Dit argument doet geen beroep op noties als (zelf)bewustzijn, gevoelens en vrije wil.
- Het argument is gebaseerd op **de stelling van Gödel**
- De stelling van Gödel leunt zwaar op de notie van **formeel systeem**

Het argument van Lucas tegen het materialisme

1. Formele Systemen

- 1.1. Wat zijn formele systemen?
- 1.2. Verschillende soorten formele systemen
- 1.3. Kenmerken van formele systemen

2. De stelling van Gödel

- 2.1 Hoe luidt de stelling?
- 2.2 Gevolgen en impact
- 2.3 Een schets van het bewijs

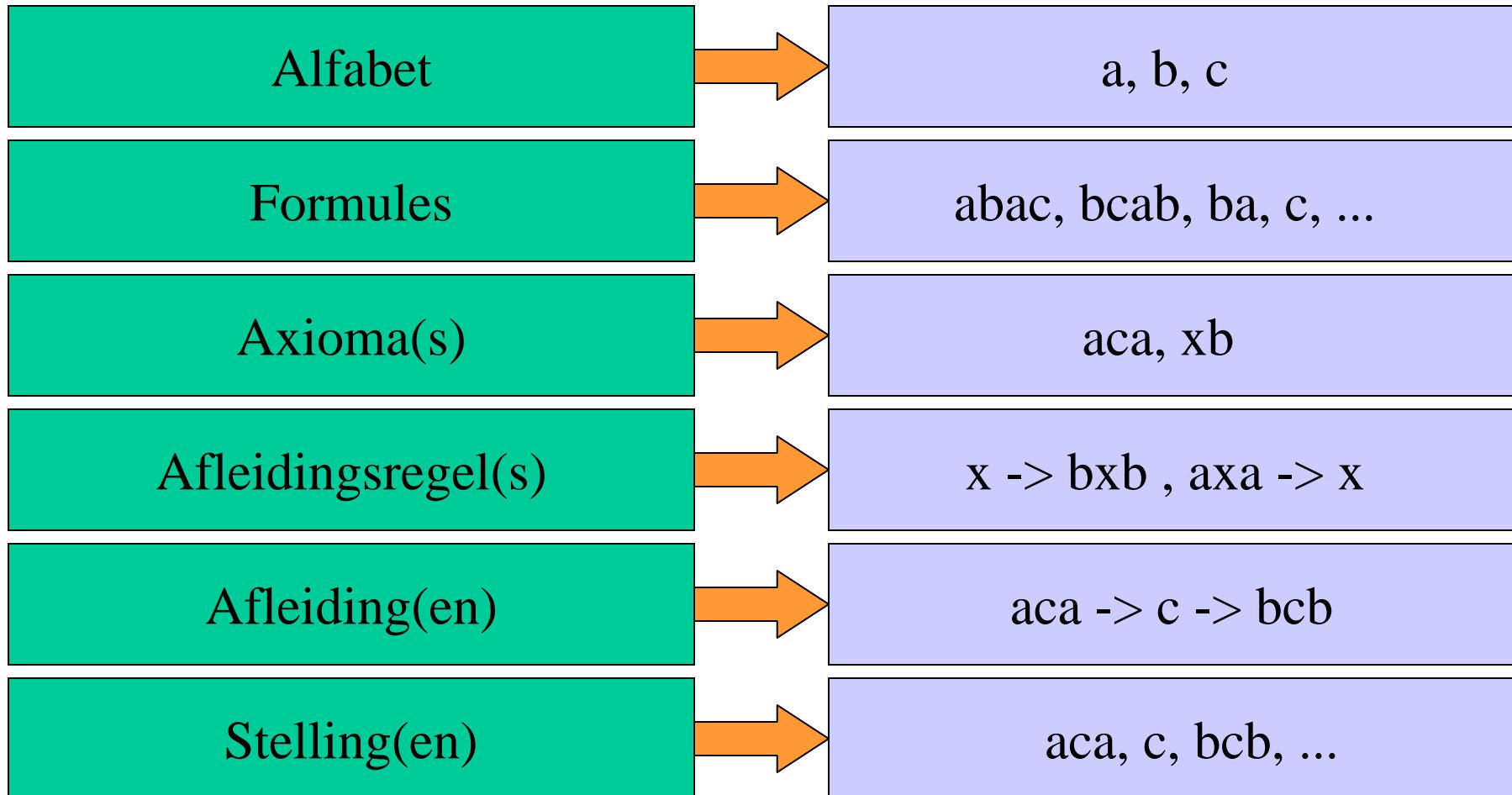
3. Lucas' argument tegen het mechanisme

4. Discussie: is het Lucas argument overtuigend genoeg?

Formeel systeem

Waaruit bestaat een
formeel systeem?

Eenvoudig voorbeeld:
Het 'abc' systeem



Abstracte en geïnterpreteerde formele systemen

Abstracte formele systemen:

De *formules* hebben *géén* enkele *betekenis*.

Zij verwijzen naar niets buiten het formele systeem.

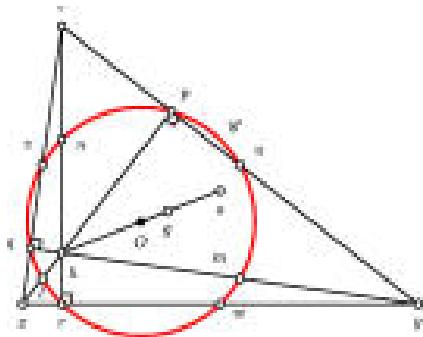
(bijvoorbeeld het ‘abc’ systeem)

Geïnterpreteerde formele systemen:

De *formules* hebben *wél* een bepaalde *betekenis*.

Zij verwijzen naar iets buiten het formele systeem.

Voorbeelden van geïnterpreteerde formele systemen



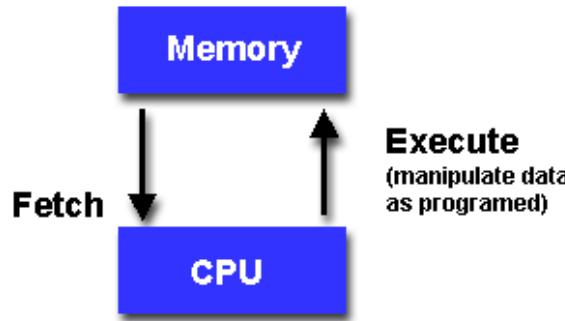
De meetkunde

$$\begin{array}{r} 457 \\ 683 \\ \hline 1140 \end{array}$$

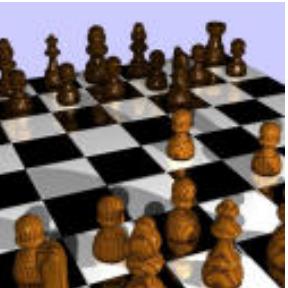
De rekenkunde



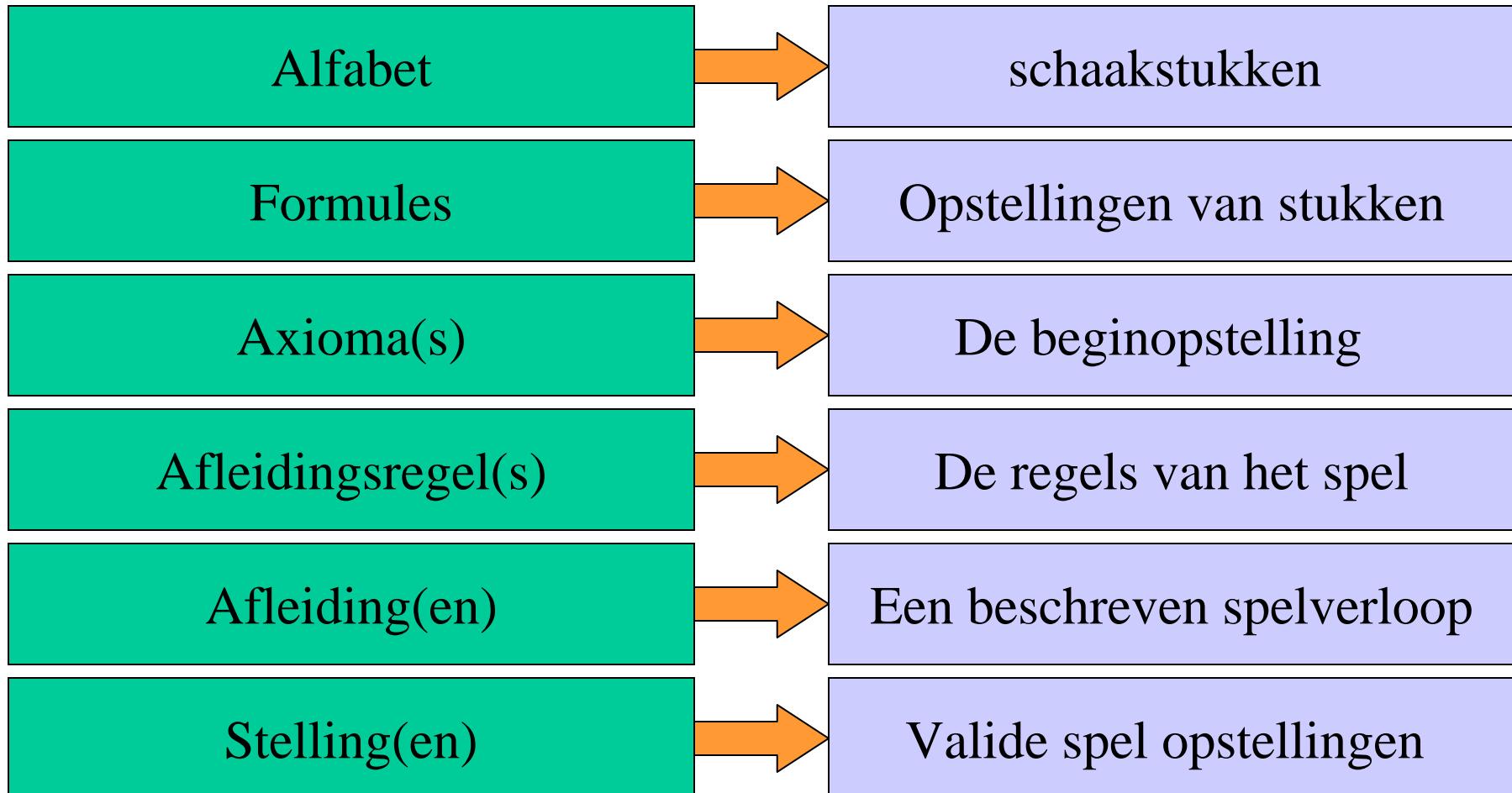
Het schaakspel

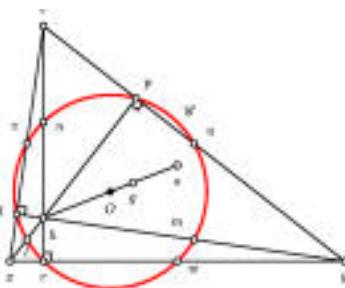


Een machine

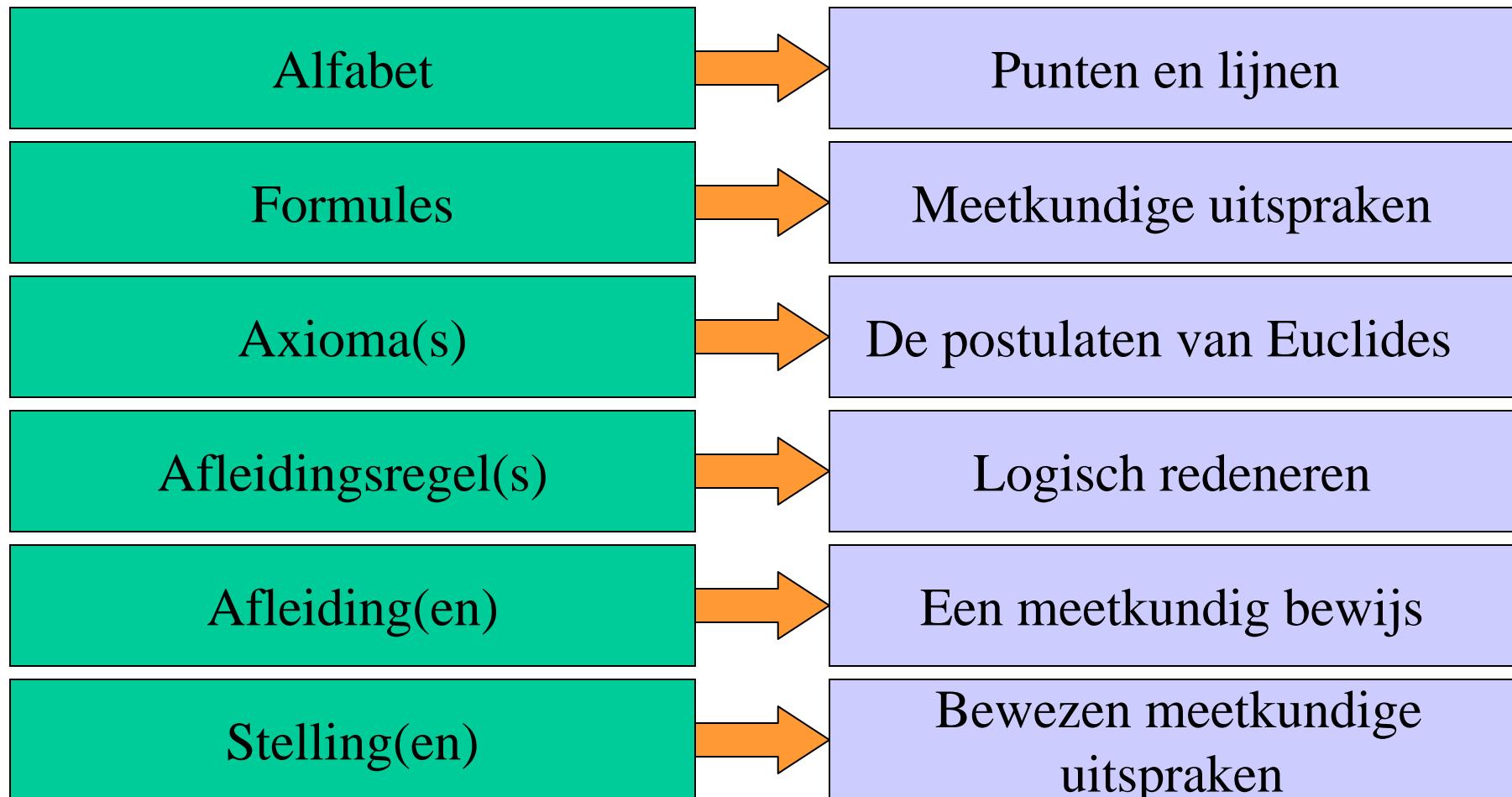


Het schaakspel



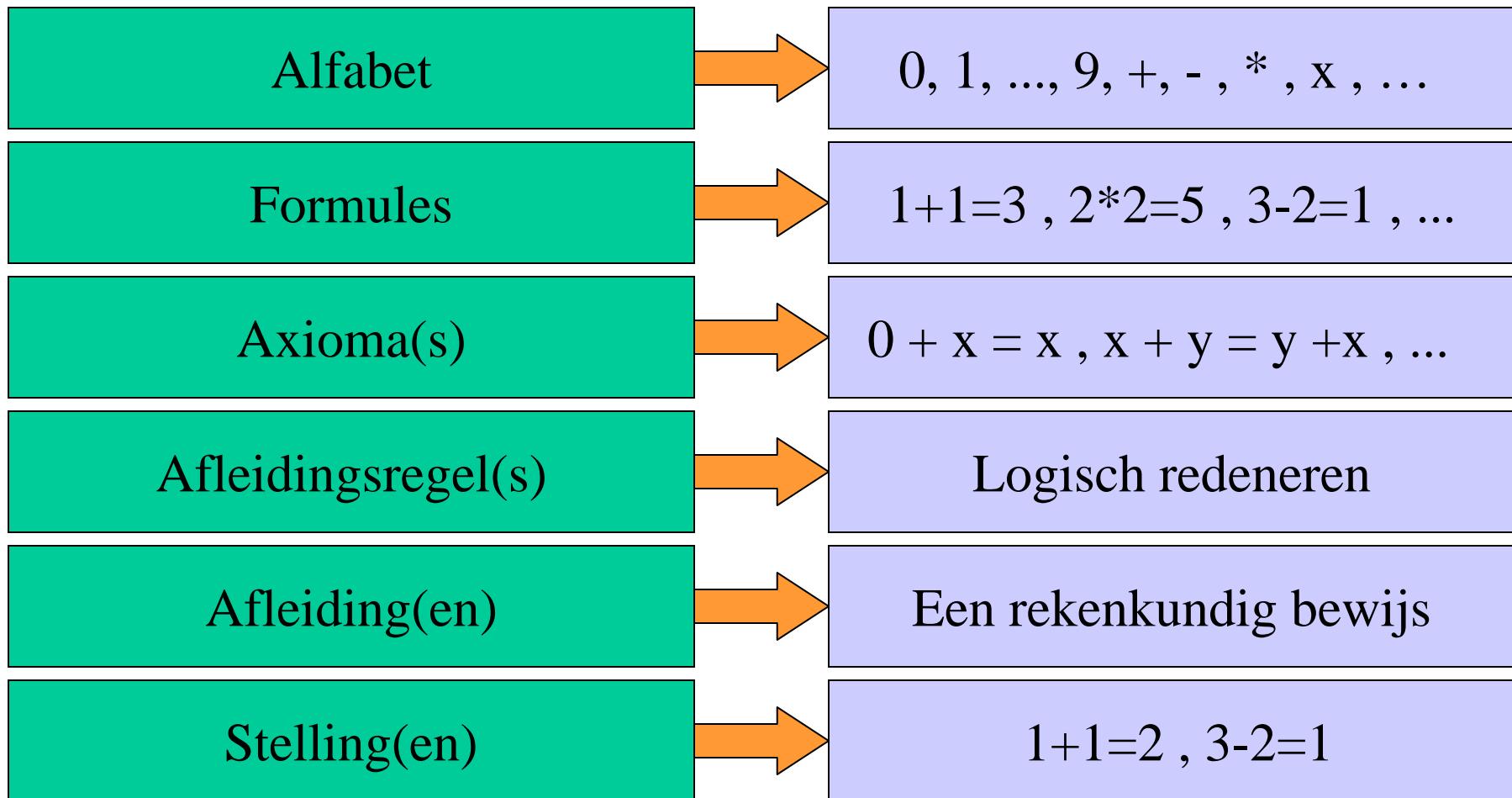


De meetkunde

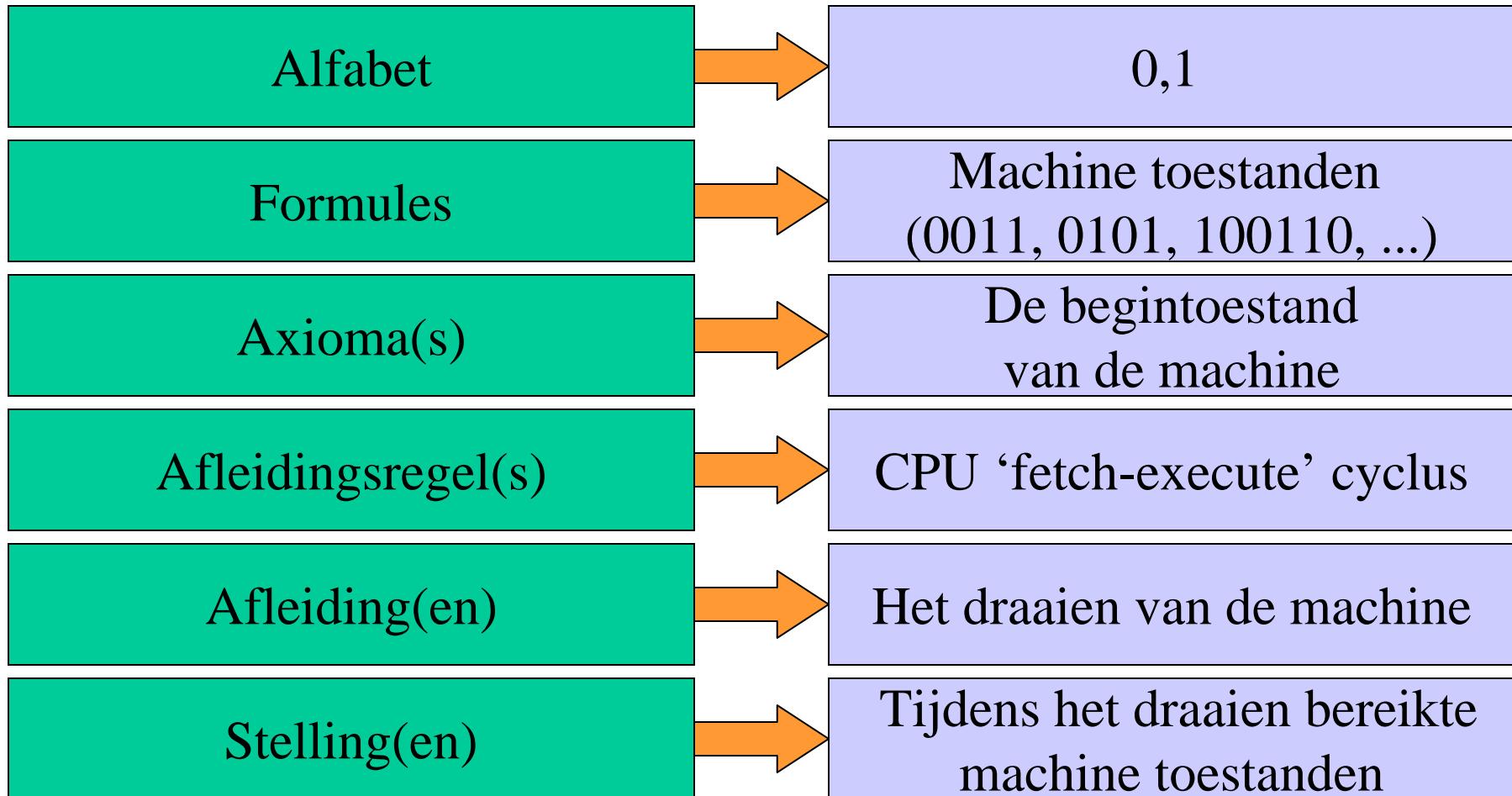
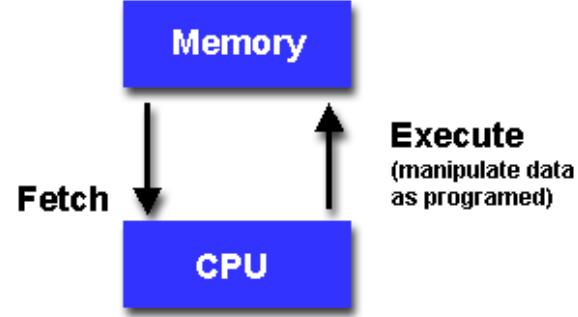


$$\begin{array}{r} 457 \\ 683 \\ \hline 1140 \end{array}$$

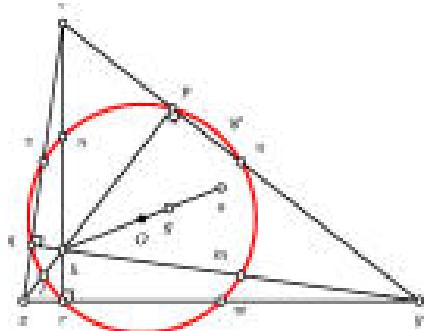
De rekenkunde



Een machine



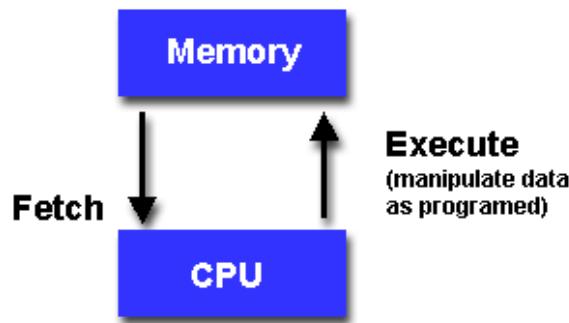
Vier voorbeelden van geïnterpreteerde systemen. Zijn ze in twee groepen te verdelen?



$$\begin{array}{r} 457 \\ 683 \\ \hline 1140 \end{array}$$

*... Formules als
uitspraken over ...*

*... Formules als
toestanden van ...*



Twee groepen van geïnterpreteerde systemen

Uitspraak geïnterpreteerde formele systemen

Het is zinvol om over de waarheid of onwaarheid van stellingen van deze systemen te spreken. Uitspraken (over de meet- of rekenkunde) zijn immers waar of onwaar

Toestand geïnterpreteerde formele systemen

Het is onzinnig om over de waarheid van stellingen van deze systemen te spreken. Toestanden (van een spel of machine) zijn immers niet waar of onwaar

Wat willen we eigenlijk zoal van formele systemen weten?

- Is het systeem **beslisbaar**?
 - Een formeel systeem is beslisbaar indien er een test bestaat die voor iedere formule bepaalt of deze formule al dan niet bewijsbaar is binnen het systeem
 - Beslisbaarheid is als begrip zinvol voor alle typen systemen
- Is het systeem **volledig**?
 - Een formeel systeem is volledig indien alle ware formules bewijsbaar zijn
 - Volledigheid is als begrip alléén zinvol voor uitspraak geïnterpreteerde systemen
- Is het systeem **consistent**?
 - Een formeel systeem is consistent indien alle bewijsbare formules waar zijn
 - Consistentie is als begrip alléén zinvol voor uitspraak geïnterpreteerde systemen



Citaat uit een brief van de Oostenrijkse econoom Morgenstern aan de latere Oostenrijkse kanselier Kreisky (1965):

“Einstein once told me that after Gödel his own work no longer meant much to him, and that he simply came to Princeton to have the privilege of walking home with Gödel”

J. Von Neumann: “Gödel is definitely the greatest logician since Aristotle.





Hoe luidt de Stelling van Gödel (1931)?

“De elementaire rekenkunde is ofwel inconsistent ofwel onvolledig”

Gevolgen

“De rekenkunde kan niet zowel consistent als volledig zijn”

“Aannemende dat de rekenkunde consistent is, moet geconcludeerd worden dat er ware rekenkundige uitspraken bestaan die niet bewijsbaar zijn binnen de rekenkunde”

“Rekenkundige waarheid is *sterker* dan rekenkundige bewijsbaarheid”

Impact

De droom van Cantor, Frege, Russell, Whitehead, Hilbert en een groot aantal andere wiskundigen stortte compleet in.

Alan Turing en Alonzo Church gaven, gebruikmakend van Gödel’s ideeën, in 1936 de absolute genadeslag.

De kern van het bewijs

- De volgende bewering is één van de vele bekende formuleringen van de beroemde paradox van Epimenides (*liar's paradox*):

P = ‘Deze zin is niet waar’

- Levert de volgende bewering ook een paradox op?

Q = ‘Deze zin is niet bewijsbaar’

- Gödel wil zijn stelling bewijzen door een bewering in de rekenkunde te formuleren (i.e. een bewering over getallen) die over zichzelf zegt dat zij niet bewijsbaar is binnen de rekenkunde. De gezochte bewering moet dus tegelijkertijd een bewering over de rekenkunde zijn.

Zelf-referentie in de rekenkunde?

Buiten het systeem (Ons praten *over* de rekenkunde)

De rekenkundige formule '1+1=2' bevat drie getallen

'2+5' betreft het linkerdeel van de formule '2+5=4+3'

Er zijn oneindig veel verschillende rekenkundige formules

...

...

Binnen het systeem (formules *in* de rekenkunde)

1+1=2

Er bestaat een x zodat $x + 1 = 4$

Voor alle x en y geldt dat $x+y = y+x$

...

...

...

Hoe denkt Gödel een dergelijke rekenkundige formule te kunnen vinden? Rekenkundige formules bestaan toch alléén *in* de rekenkunde? Zij betreffen toch immers uitsluitend rekenkundige relaties tussen getallen? Hoe kan een rekenkundige formule daarom überhaupt iets zeggen *over* de rekenkunde, laat staan over zichzelf?

Zelf-referentie in de rekenkunde?

- Rekenkundige formules gaan inderdaad over (relaties tussen) getallen.
- Stel echter dat we besluiten om met iedere rekenkundige formule precies één uniek getal te laten corresponderen (Gödel getal).
 - Formule **1+1=2** krijgt bijvoorbeeld Gödel getal **200**
 - Formule **Er bestaat een x zodat $x + 1 = 4$** krijgt bijv. Gödel getal **400**
 - Formule **Voor alle x en y geldt dat $x+y = y+x$** krijgt bijv. Gödel getal **600**
 - Etc...
- Dankzij deze correspondentie gaan rekenkundige formules voortaan indirect dus eveneens over (relaties tussen) rekenkundige formules.
- Wellicht bestaat er dan inderdaad een rekenkundige formule die van zichzelf zegt dat zij niet binnen de rekenkunde bewijsbaar is...

Stap 1 van het bewijs: Het kiezen van een formeel systeem

<i>Alfabet van de rekenkunde</i>	<i>Toelichting</i>
\sim	Negatie
$+$	Optelling
$*$	Product
E	Er is een ...
$=$	Gelijkheid
0	Nul
s	Successor
$($	Haakje open
$)$	Haakje sluit
,	Komma
x	Variabele
...	...

$0, s(0)=1, s(s(0))=2, s(s(s(0)))=3, \dots$

$(x) / \text{voor alle } x \text{ geldt} \dots$

$(x)(x=x)$

$\text{Ex} / \text{er is een } x \text{ zodanig dat} \dots$

$\text{Ex}(x=s(0))$

Stap 1 van het bewijs: Het kiezen van een formeel systeem

<i>Alfabet van de rekenkunde</i>	<i>Toelichting</i>
\sim	Negatie
$+$	Optelling
$*$	Product
E	Er is een ...
$=$	Gelijkheid
0	Nul
s	Successor
$($	Haakje open
$)$	Haakje sluit
,	Komma
x	Variabele
...	...

Eenvoudige formules

$$0=0 \quad \text{“0 is 0”}$$

$$\sim(0=s(0)) \quad \text{“0 is niet 1”}$$

$$s(0)+s(0)=s(s(0)) \quad \text{“1+1=2”}$$

Complexere formules

$$(x)(x+x=s(s(0))*x) \quad \text{Voor alle } x \text{ geldt } x+x=2x$$

$$(Ex)(x=s0) \quad \text{Het getal 0 heeft een opvolger}$$

Zeer complexe formules

$$\sim Ex(y) Ez(y+z=x) \quad \text{Er is géén grootste getal}$$

Stap 2 van het bewijs: Uitdelen van Gödel getallen

Alfabet van de rekenkunde	Bijbehorend Gödel nummer
~	1
+	2
*	3
E	4
=	5
0	6
s	7
(8
)	9
,	10
x	11
...	...

0=0

$$2^6 \text{ maal } 3^5 \text{ maal } 5^6 = 243000000$$

(Ex)(x=s0)

$$2^8 \text{ maal } 3^4 \text{ maal } 5^{11} \dots 23^6 \text{ maal } 29^9 = \dots$$

Bij iedere rekenkundige formule hoort precies één Gödel getal en bij ieder Gödel getal hoort precies één rekenkundige formule!

Een afleiding in de rekenkunde is niets meer dan een rijtje rekenkundige formules ('van axioma tot stelling'). Het Gödel getal van een afleiding is het product van de Gödel getallen van de formules in het bijbehorende rijtje

Dankzij stap 2 kunnen uitspraken *over* de rekenkunde vertaald worden in formules *binnen* de rekenkunde

Stel $(x + x) * x$ heeft Gödel getal **10000**

Stel $(x + x)$ heeft Gödel getal **100**

De volgende uitspraak *over* de rekenkunde...

$(x + x)$ is een linkerdeel van $(x + x) * x$

... wordt vertaald in onderstaande formule *binnen* de rekenkunde:

$(Ex)(100*x=10000)$

Stap 3 van het bewijs: Vertalen van ‘bewijsbaarheid’

De volgende uitspraak over de rekenkunde...

Het rijtje rekenkundige formules met Gödel getal x is een bewijs voor de rekenkundige formule met Gödel getal z

... wordt vertaald in onderstaande formule binnen de rekenkunde:

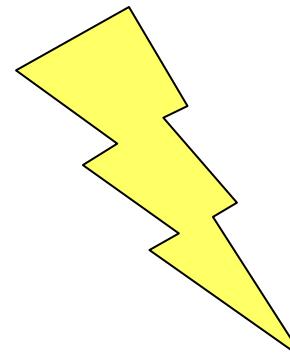
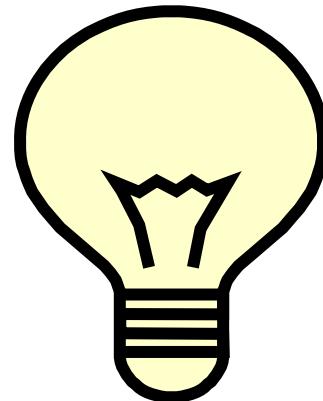
Dem(x,z)



Stap 4 van het bewijs: De beroemde Gödel formule

Gödel vond een Gödel getal (laten we deze **G** noemen) zodanig dat de volgende rekenkundige formule **G** als Gödel getal heeft...

$$(x) \sim \text{Dem}(x, G)$$



Het Lucas Argument

Lucas begint zijn artikel ‘Minds, Machines, and Gödel’ uit 1961 met de volgende woorden:

“Volgens mij toont de stelling van Gödel aan dat het Mechanisme onwaar is, dat wil zeggen dat de menselijke geest niet verklaard kan worden als machine”

De mens zou dus altijd méér zijn dan een machine...

Welke argument gebruikt Lucas?

Het Lucas Argument

Lucas in zijn artikel:

“Hoe gecompliceerd we een machine ook maken, zij zal, als het ook werkelijk een machine is, corresponderen met een formeel systeem, dat weer onderhevig is aan de procedure van Gödel om een ware formule te vinden die onbewijsbaar is in dat systeem. De machine zal niet in staat zijn deze formule als waar te produceren, ook al kan de menselijke geest zien dat zij waar is. En daarom is de machine nog steeds géén adequaat model van de menselijke geest. Het mechanistische model, moet in zekere zin, eindig en begrensd zijn: en dan doet de geest het altijd iets beter. Dankzij de stelling van Gödel heeft de menselijke geest altijd het laatste woord.”



Het Lucas Argument (*bewijs uit het ongerijmde*)

- Stel ik ben een machine
- De machine die ik ben correspondeert met een formeel systeem M
- Het beoefenen van de rekenkunde is één van mijn menselijke vermogens
- De rekenkundige waarheden die ik kan ontdekken vallen samen met de rekenkundige stellingen die binnen M bewijsbaar zijn (*)
- M omvat de rekenkunde en is dus volgens de stelling van Gödel onvolledig
- Door de Gödelprocedure op M toe te passen ontdek ik een rekenkundige waarheid W die binnen M niet bewijsbaar is
- Uit (*) volgt echter dat ik W helemaal niet kan ontdekken. Tegenspraak!
- Ik ben dus géén machine. QED

Overtuigd? Waarom (niet)?