Les B-08 Kunstmatige intelligentie en logica

8.1 De Turing test

Toen halverwege de 20e eeuw de computer zijn intrede deed, stelde de Brit **Alan Turing** (1912-1953) in een wetenschappelijke publicatie de vraag "Kunnen machines denken?".

Omdat deze vraag wat duidelijker te stellen, schetste hij als voorbeeld de **Turing test**. De Turing test komt erop neer dat als een computer iemand voor de gek kan houden en deze kan laten geloven dat hij een mens is, de computer intelligent moet zijn. In zijn oorspronkelijke vorm is de test een spel dat gespeeld wordt door een man (A), vrouw (B) en een mannelijke of vrouwelijke ondervrager (C). De ondervrager kent de man en vrouw onder de namen X en Y (die hun antwoorden op papier noteren om met hun stemgeluid niet hun sekse te verraden) en moet er door het stellen van vragen achter komen of X de man (A) is of juist Y de man (A) is. De ondervrager stelt bijvoorbeeld aan X de vraag "Kan je me vertellen hoe lang je haar is?". Stel dat X de man (A) is, dan is het juist zijn bedoeling om de ondervrager te misleiden. Hij zal een antwoord geven als "Ik draag het nu redelijk kort, mijn langste haren zijn 20 cm lang". De vrouw (B) zal de ondervrager willen helpen door het geven van correcte antwoorden, maar dat draagt niets bij aan de onzekerheid van de ondervrager als hij van de man (A) dezelfde informatie krijgt.

Turing, die een belangrijke bijdrage leverde aan de ontwikkeling van de computer, stelde vervolgens de vraag of een machine de rol van ondervrager zou kunnen spelen in dit spel. Natuurlijk zou een machine het stellen van vragen en het verwerken van de antwoorden uit kunnen voeren! Zolang het denken beperkt is tot het denken in een redeneersysteem met een vast aantal mogelijkheden moet het mogelijk zijn om machines denktaken te laten verrichten.

De geschiedenis heeft Turing gelijk gegeven. De computer kan intelligente handelingen verrichten als:

- schaken en andere spellen spelen
- vertalen
- tekst, spraak en beeld (oog, vingerafdruk, gezichten) herkennen

Het **kunstmatige** van deze **intelligentie** is dat de computer "slechts" voorgeprogrammeerde handelingen kan verrichten. Computerintelligentie is dus geen vorm van intelligentie die zichzelf kan ontwikkelen. Een goed voorbeeld van kunstmatige intelligentie is het schaakspel. De computer kan een groot aantal zetten "vooruit denken", dat wil zeggen doorberekenen welke mogelijkheden ontstaan bij het verplaatsen van een bepaald schaakstuk. Daarvoor hoeven slechts de regels bekend te zijn welke bewegingsmogelijkheden elk van de stukken op het bord heeft. We noemen het schaakspel ook wel een **gesloten redeneersysteem**. Er kan niet ineens een stuk het speelveld inlopen dat nieuwe mogelijkheden aan het spel toevoegt.

Het menselijk denken vormt een **open redeneersysteem**. Wij zijn vaak in staat om in nieuwe situaties nieuwe, oorspronkelijke oplossingen te vinden voor de problemen die we tegenkomen. Steven Spielberg heeft het in zijn film "Artificial Intelligence" zo ver laten komen dat computers dit menselijke vermogen zijn gaan beschikken. Of het ooit zo ver zal komen?

In ieder geval hebben machines bewezen te kunnen denken voor zover het om gesloten redeneringen gaat.

8.2 Onzuiverheden in de natuurlijke taal en de wiskunde

Sinds het begin van de 20e eeuw, toen de computer nog niet bestond, hebben filosofen en wiskundigen zich bezig gehouden met de vraag of er een zuivere logische taal geformuleerd kon worden. Een taal waarin geen onzuiverheden meer bestonden.

onzuiverheden in de natuurlijke taal

In de natuurlijke taal treden allerlei onzuiverheden op. Een bekend voorbeeld is de **paradox**, een bewering die zichzelf tegenspreekt.

Een bekend voorbeeld van een bewering die zichzelf tegenspreekt is (paradox van Epimenides van Kreta):

Alle Kretenzers liegen.

Als niet alle Kretenzers liegen, is de bewering onjuist maar ook als alle Kretenzers liegen, is de bewering onjuist. Want degene die de bewering doet spreekt dan de waarheid.

onzuiverheden in de wiskunde

Ook in de wiskunde zijn paradoxen mogelijk (paradox van Russell):

Laat S de verzameling zijn van alle verzamelingen die zichzelf niet bevatten.
Is S een element van zichzelf ?

Als S een element is van zichzelf, dan is het geen verzameling die zichzelf niet bevat en dus is S geen element van zichzelf. Als S geen element is van zichzelf, dan is S een element van de verzameling S en dus een element van zichzelf.

Kennelijk laat de manier waarop we taalkundige en wiskundige beweringen opbouwen de ruimte om beweringen op te bouwen die niet logisch zijn.

OPDRACHT

Opdracht 8.1
Zoek op internet nog een taalkundige en een wiskundige paradox op.

8.3 De taal van de logica

Filosofen en wiskundigen zagen het in het begin van de 20e eeuw als opdracht om een taal te ontwikkelen zonder onzuiverheden. De taal die zij ontwikkelden heet wel mathematische (wiskundige) **logica**.

De logica gaat uit van beweringen (**proposities**). Een bewering kan zijn “Het regent”. Het wiskundige van de logica bestaat er uit dat beweringen worden vervangen door letters, net als bij het letterrekenen. Zo kunnen we de bewering “Het regent” bewering A noemen.

In de logica is een aantal symbolen opgenomen die beweringen aan elkaar kunnen koppelen (**connectieven**):

$\neg$	niet
$\wedge$	en
$\vee$	of
$\rightarrow$	als …. dan

Is A de bewering “Het regent” en B de bewering “Ik wordt nat”, dan betekent:

$\neg A$	Het regent niet.
$A \vee \neg A$	Het regent of het regent niet.
$A \wedge B$	Het regent en ik wordt nat.
$\neg B \vee A$	Ik wordt niet nat of het regent.
$A \rightarrow B$	Als het regent, dan wordt ik nat.

OPDRACHT

Opdracht 8.2

Gegeven zijn de beweringen:

X: “Het regent”
Y: “Ik ga voetballen”
Z: “Ik ga naar de stad”

a) Vertaal de volgende logische bewering in natuurlijke taal:
   $Y \wedge \neg Z$
b) Vertaal de volgende logische bewering in natuurlijke taal:
   $Y \rightarrow X$
c) Vertaal de volgende zin in een logische bewering:
   “Als het niet regent dan ga ik voetballen of ga ik naar de stad”

8.4 Logisch redeneren

Met behulp van de koppelsymbolen kunnen we logische beweringen samenstellen. Met deze logische beweringen kunnen logische redeneringen worden opgezet. De logica houdt zich bezig met de vraag welke beweringen logisch af te leiden zijn (uit enkele basisbeweringen) en welke niet af te leiden zijn.

Er zijn verschillende manieren om te laten zien of een bewering af te leiden is (waar is) . Eén van die manieren is door gebruik te maken van waarheidstabellen. Hieronder staat de **waarheidstabel** die hoort bij het $\vee$-symbool.

A	**B**	**A $\vee$ B**
1	1	1
1	0	1
0	1	1
0	0	0

In de tabel lees je af dat A $\vee$ B waar ("1") is als of A waar is, of B waar is of beide waar zijn. In natuurlijke taal: de bewering "Het regent of ik wordt nat" is waar als het regent, als je nat wordt of als beide het geval zijn maar niet als het niet regent en je niet nat wordt.

De waarheidtabel bij het $\rightarrow$-symbool is:

A	**B**	**A $\rightarrow$ B**
1	1	1
1	0	0
0	1	1
0	0	1

Deze tabel is misschien moeilijker te begrijpen. De bewering "Als A dan B" is waar als A en B waar zijn, maar niet als A waar is en B niet. Immers, als A waar is, moet B ook waar zijn. Als A niet waar is, kan B waar of onwaar zijn. De bewering "Als A dan B" zegt tenslotte alleen maar iets over situaties waarin A waar is.
In natuurlijke taal: de bewering "Als het regent, dan wordt ik nat" is waar als het regent en je nat wordt, maar niet waar als het regent en je niet nat wordt. Als het niet regent, kan je nog altijd nat worden of niet. Dat doet niets af van de bewering dat als het regent, je nat wordt.

OPDRACHT

Opdracht 8.3
Maak de waarheidstabellen bij het $\neg$-symbool en het $\wedge$-symbool.

Aan de hand van een eenvoudig voorbeeld bekijken we hoe je waarheidstabellen kunt gebruiken om de geldigheid van een bewering aan te tonen.

Twee logische beweringen kunnen hetzelfde zeggen, gelijkwaardig zijn:

$A \rightarrow B$	Als het regent, dan wordt ik nat.
$\neg A \vee B$	Het regent niet of ik wordt nat.

Om te laten zien dat de twee beweringen gelijkwaardig zijn tonen we aan dat de waarheidstabellen bij beide beweringen gelijk aan elkaar zijn.

De waarheidstabel bij de eerste bewering ($A \rightarrow B$) hebben we als gezien:

A	**B**	**$A \rightarrow B$**
1	1	1
1	0	0
0	1	1
0	0	1

De waarheidstabel bij de tweede bewering ($\neg A \vee B$) is:

A	**$\neg A$**	**B**	**$\neg A \vee B$**
1	0	1	1
1	0	0	0
0	1	1	1
0	1	0	1

We zien dat de waarheidstabellen hetzelfde zijn. Dat wil dus zeggen dat de bewering $A \rightarrow B$ in dezelfde gevallen waar is als de bewering $\neg A \vee B$.

OPDRACHT

Opdracht 8.4
Toon met behulp van waarheidstabellen aan dat de beweringen

$\neg(A \wedge B)$ en $\neg A \vee \neg B$

gelijkwaardig zijn.

A	**B**	**$A \wedge B$**	**$\neg (A \wedge B)$**
1	1		
1	0		
0	1		
0	0		

A	**$\neg A$**	**B**	**$\neg B$**	**$\neg A \vee \neg B$**
1		1		
1		0		
0		1		
0		0		

8.5 Nieuwe symbolen

Je hebt nu een indruk gekregen hoe logica werkt. Aan de hand van beweringen kunnen we andere beweringen afleiden of de gelijkwaardigheid van beweringen aantonen.

In deze paragraaf ronden we deze kennismaking met het vakgebied logica af door nog twee symbolen aan de taal toe te voegen.

$\exists$	er is
$\forall$	alle

Deze symbolen worden gebruikt in dat deel van de logica dat zich bezig houdt met beweringen waarin verzamelingen, eigenschappen en relaties voorkomen. Dit soort beweringen wordt ook wel **predikaten** genoemd:

$\mathrm{M}(x)$	x is mens
$\mathrm{S}(x)$	x is sterfelijk

De bewering "alle mensen zijn sterfelijk" kan nu logisch worden genoteerd als:

$$\forall x\ (\ \mathrm{M}(x) \rightarrow \mathrm{S}(x)\ )$$

Dat wil zeggen: "voor alle elementen x geldt: als x in de verzameling mens zit, dan zit x ook in de verzameling sterfelijk".

Een ander voorbeeld: "er is een zwart schaap"

$\mathrm{S}(x)$	x is schaap
$\mathrm{Z}(x)$	x is zwart

$$\exists x\ (\ \mathrm{S}(x) \wedge \mathrm{Z}(x)\ )$$

Een laatste voorbeeld: "Alle logici kunnen logisch redeneren. Derhalve is Jan geen logicus of hij kan logisch redeneren"

$\mathrm{L}(x)$	x is logicus
$\mathrm{R}(x)$	x kan redeneren

$$\forall x\ (\ \mathrm{L}(x) \rightarrow \mathrm{R}(x)\ ) \qquad \rightarrow \qquad \neg\ \mathrm{L}(j) \vee \mathrm{R}(j)$$

OPDRACHT

Opdracht 9.5
Vertaal de zin: "Als ik niets begrijp, dan begrijp ik iets niet." in een logische zin.

8.6 Samenvatting

De Brit **Alan Turing** stelde de vraag "Kunnen machines denken" en beantwoordde deze aan de hand van het voorbeeld van de **Turing test**.

Machines kunnen denken zolang het gaat om voorgeprogrammeerde acties in een **gesloten redeneersysteem**. We noemen deze operationele vorm van intelligentie ook wel **kunstmatige intelligentie**.

In de wiskunde en filosofie was al voor het ontstaan van de computer een vakgebied ontstaan dat zich bezig hield met het redeneren, de **logica**.

De logica stelde zich tot doel om een taal te ontwikkelen zonder onzuiverheden zoals **paradoxen**, zichzelf tegensprekende beweringen.

In de logica worden beweringen (**proposities**) aan elkaar gekoppeld door symbolen (**connectieven**). In de propositielogica onderscheiden we de volgende symbolen:

$\neg$	niet
$\wedge$	en
$\vee$	of
$\rightarrow$	als …. dan

Met behulp van deze symbolen kunnen uit basisbeweringen nieuwe beweringen worden samengesteld en afgeleid.

De waarheid van een bewering kan worden nagetrokken met behulp van **waarheidstabellen**.

Een ander deel van de logica houdt zich bezig met beweringen waarin verzamelingen, eigenschappen en relaties voorkomen. Dit soort beweringen wordt ook wel **predikaten** genoemd. In de predikaatlogica worden de volgende symbolen toegevoegd:

$\exists$	er is
$\forall$	alle

ANTWOORDEN

Opdracht 8.1
Met de zoektermen “paradox” en “voorbeeld” kom je vanzelf een aantal bruikbare links tegen.

Een voorbeeld van een paradox in natuurlijke taal:

<u>De barbier van Sevilla.</u>
In de middeleeuwen had de barbier van Sevilla een uithangbord: “ik scheer alle mannen die zichzelf niet scheren”. Een voorbijganger vroeg: “Scheert u zichzelf of niet ?”.

Een voorbeeld van een wiskundige paradox:

<u>Achilles en de schildpad (paradox van Zeno)</u>
De snelvoetige Achilles gaat een wedstrijd aan met een schildpad. De laatste krijgt een voorsprong. Wanneer Achilles het punt A bereikt, waar de schildpad kort tevoren was, is de schildpad intussen in punt B aangekomen. Arriveert Achilles in dit punt B, dan is de schildpad intussen aangekomen in punt C, enzovoorts.
Conclusie: de achterstand wordt kleiner, maar Achilles haalt de schildpad nooit in, totdat de schildpad niet meer verder kan en neervalt.

Als toegift een filosofische paradox:

> Een vrachtwagenchauffeur rijdt op een smalle weg en ziet voor zich aan zijn kant van de weg twee wandelaars en aan de overkant een fietser. Hij kan er niet tussendoor dus hij remt, maar de remmen weigeren. Als hij niets doet rijdt hij twee mensen dood. Als hij naar links stuurt slechts een. Wat moet hij doen?

Veel mensen zijn het er over eens dat hij naar links moet sturen. Dan wordt de situatie veranderd.

> Een arts loopt in de gang van zijn ziekenhuis. Hij heeft twee patiënten met een zeer zeldzame bloedgroep. De ene heeft een hart nodig, de ander een lever. Maar die zijn niet voorradig en de patiënten zullen sterven. Dan komt hem een verpleegster tegemoet waarvan hij weet dat zij de zelfde zeldzame bloedgroep heeft. De arts heeft hier ook de keuze. Doet hij niets, dan vallen er twee doden, maar hij kan ook ingrijpen en een verpleegster opofferen.

In dit geval is vrijwel iedereen het erover eens dat de verpleegster niet mag worden ingezet. Toch zijn in essentie beide situaties gelijk.

Opdracht 8.2
De vertaling van de zinnen is:

a) Ik ga voetballen en ik ga niet naar de stad.
b) Als ik ga voetballen, dan regent het.
c) $\neg X \rightarrow (Y \vee Z)$

ANTWOORDEN

Opdracht 8.3
De waarheidstabellen bij het $\neg$-symbool en het $\wedge$-symbool zijn:

A	**$\neg$A**
1	0
0	1

A	**B**	**A $\wedge$ B**
1	1	1
1	0	0
0	1	0
0	0	0

Opdracht 8.4
De beweringen $\neg$(A $\wedge$ B) en $\neg$A $\vee$ $\neg$B zijn gelijkwaardig:

A	**B**	**A $\wedge$ B**	**$\neg$ (A $\wedge$ B)**
1	1	1	0
1	0	0	1
0	1	0	1
0	0	0	1

A	**$\neg$ A**	**B**	**$\neg$ B**	**$\neg$ A $\vee$ $\neg$ B**
1	0	1	0	0
1	0	0	1	1
0	1	1	0	1
0	1	0	1	1

Opdracht 8.5
De vertaling van de zin “Als ik niets begrijp, dan begrijp ik iets niet.” is:

B(x) ik begrijp x

$$\forall x \ \neg B(x) \rightarrow \exists x \ \neg B(x)$$

Logisch, toch ?