Inaugurale rede uitgesproken door

Frank van Harmelen

op 13 Februari 2003

bij het aanvaarden van het ambt van Hoogleraar

in het vakgebied

“Kennisrepresentatie en Redeneren”

aan de Vrije Universiteit

Mijnheer de Rector Magnificus,
Dames en Heren,

Het is voor mij een grote eer om hier te staan, een eer waarvan ik sterk voel dat ik die niet alleen maar aan mezelf te danken heb. Vandaar dat ik, misschien tegen de traditie in, wil beginnen met wat dankwoorden uit te spreken.

Allereerst dank ik het bestuur van de Vereniging voor christelijk wetenschappelijk onderwijs en het College van Bestuur van deze Universiteit voor mijn benoeming.

Mijn ouders verdienen een zeer oprecht gemeend dankwoord: bedankt voor alle aanmoediging die jullie mij over de jaren heen hebben gegeven, en voor het vertrouwen dat jullie in mij hadden, vaak aanzienlijk meer vertrouwen dan ik zelf had (en dat geldt tot op de dag van vandaag). Het moet voor jullie een hele opluchting zijn om nu toch eindelijk jullie andere zoon ook in toga te zien staan :-).

Lynda (of moet ik zeggen: Prof. Hardman, want jij mocht al een half jaar eerder dan ik een toga dragen), bedankt voor alles wat we samen hebben gedeeld, en nog zullen delen in de toekomst.

Hooggeleerde Van Vliet, beste Hans, dank dat je 20 jaar geleden mijn eerste leermeester in het onderzoek was. Jij hebt me aangespoord om in de wetenschap mijn toekomst te zoeken.

Hooggeleerde Van Mill en Treur, beste Jan en Jan, dank dat jullie je zo hebben ingezet voor het verwezenlijken van deze leerstoel. Ik zal mijn best doen om ervoor te zorgen dat jullie er geen spijt van krijgen.

En tot slot een oprecht dankwoord aan de Hooggeleerde Wielinga, beste Bob, en aan Professor Alan Bundy in Schotland. Ik beschouw jullie beiden als mijn belangrijkste leermeesters. Van jullie heb ik heel veel geleerd, veel over de technische inhoud van ons vak, maar ook nog heel veel meer dan dat. Ik zal trots zijn als ik ooit nog maar eens half zo goed zal worden als jullie in het leiden en inspireren van een groep onderzoekers om mij heen.

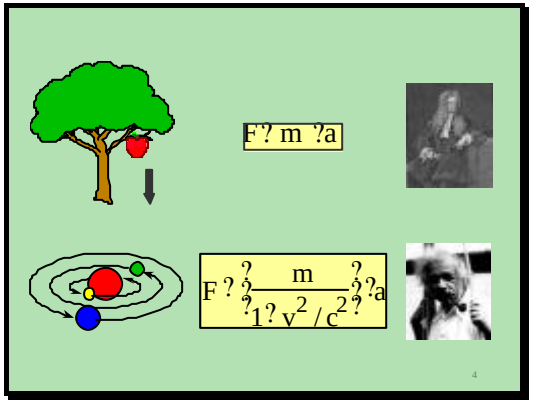
Maar goed, genoeg dankwoorden, het moet vandaag gaan over wetenschap, en wel over “Kennisrepresentatie en Redeneren”. Daartoe wil ik eerst iets zeggen over wetenschap in het algemeen.

"Wetenschap is een methode voor het omgaan met onzekerheid; ze geeft ons betere modellen, geen eeuwige waarheden".

Dit is de tekst van een poster die al sinds jaar en dag de deur siert van een kast in mijn werkkamer hier op de Universiteit.

Wetenschap gaat over het bouwen van modellen,niet over het formuleren van eeuwige waarhedenIk hoop maar dat dit voor mijn collega's geen nieuws is, maar het bestrijdt wel een wijd verbreid misverstand in de populaire opv atting over wetenschap.

Laat ik dit illustreren met een voorbeeld uit de Natuurkunde (dat mag vast wel van mijn vakbroeders, nu we sinds kort met alle exacte wetenschappen zijn verenigd in één faculteit):



We observeren een fenomeen in de wereld om ons heen, en een wetenschapper, probeert van dat fenomeen een model op te stellen.

Dat model stelt ons in staat om verschijnselen te voorspellen en als we zaken maar vaak genoeg precies genoeg kunnen voorspellen, dan zeggen we dat we die verschijnselen begrijpen.

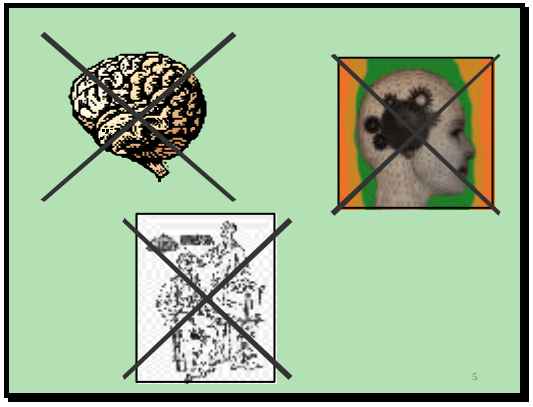
En soms klopt dat model voor bepaalde fenomemen wel (zoals Newton's model klopt voor uit de boom vallende appels), maar gaat het niet meer op als we het toepassen op andere fenomenen, bijv sommige details van de beweging van de planeet Mercurius), en was het aan een andere wetenschapper (in dit geval Einstein), om een verbeterde versie van het model op te stellen.

En de modellen van Einstein en Newton hebben natuurlijk met elkaar te maken: Einsteins model is een verfijning van het model van Newton: als de snelheid *v* maar klein is ten opzichte van de lichtsnelheid *c*, dan wordt Einstein's model gelijk aan dat van Newton.

Goed: als wetenschap bestaat uit het opstellen en testen van modellen, en als Kennis Representatie dan een tak van wetenschap is, waar maken we in de Kennis Representatie dan modellen van? Of, om het in gewoon Nederlands te zeggen:

"*waar gaat dat vak eigenlijk over?*"

Laat ik eerst een paar woorden zeggen over waar we in Kennisrepresentatie en Redeneren geen modellen over maken, met andere woorden: waar het niet over gaat:



We houden ons niet bezig met de fysieke structuur van de menselijke hersenen. Dat is een heel interessant en actief vakgebied, misschien wel een van de grote onontgonnen gebieden, maar daarvoor moet u bij een heel andere hoogleraar zijn.

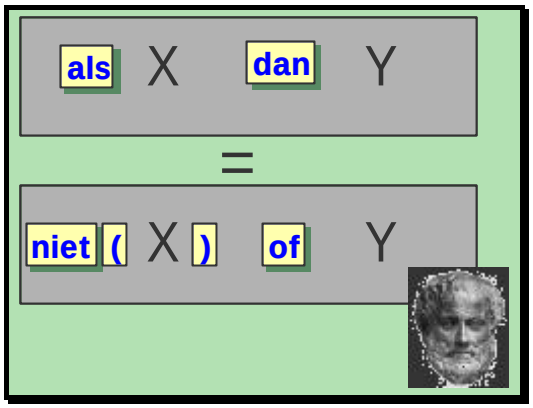
We houden ons zelfs niet bezig met de processen die zich afspelen in de menselijke hersenen. Ook dat is een heel interessant vakgebied, maar dat heet cognitieleer, en ook daarvoor moet u bij een heel andere hoogleraar zijn (een goede hoogleraar op dat gebied bevindt zich overigens in de zaal, heb ik gezien). Gevolg is dat we binnen de kennisrepresentatie ook geen nare dingen aan proefpersonen hoeven te doen, dus dat is een hele opluchting.

Waar gaat het binnen het vakgebied Kennisrepresentatie en Redeneren dan wel om?. In een algemene formulering zouden we kunnen zeggen dat het ons bij Kennisrepresentatie en Redeneren gaat om het beschrijven van de structuur van kennis en patronen van redeneren Ik zal met wat voorbeelden duidelijk maken wat hiermee bedoeld wordt, maar voordat ik dat doe een opmerking over de gebruikte termen “kennis” en “redeneren”.

Het vakgebied “Kunstmatige Intelligentie”, waar Kennisrepresentatie en Redeneren een onderdeel van is, lijdt al vanaf haar geboorte (zo eind jaren ‘50) aan de neiging om metaphoren over de menselijke ingelligentie te gebruiken (en vaak: misbruiken) als terminologie. Dat lijdt vaak tot allerlei misverstanden, zowel binnen als buiten het vak: door het gebruik van een term als “kunstmatige intelligentie” worden verkeerde verwachtingen geschapen over de doelstellingen en de resultaten van het vak.

Dat geldt ook voor termen als “kennis” en “redeneren”. Het liefst zag ik dat mijn vakgebied veel neutralere termen zou gebruiken, zoals: “symbolische informatie” in plaats van “kennis”, en “Informatie verwerking “in plaats van “redeneren”, maar ik heb in de loop van de jaren geleerd dat het geen zin heeft om me tegen het geldende taalgebruik te verzetten, en zal dus ook praten in termen van “kennis” en “redeneren”, ook al heeft mijn vakgebied weinig met de gangbare betekenis van deze termen van doen.

Ondertussen heb ik u nog steeds niet verteld waar dat vak dan wel over gaat. Laat ik dat proberen te doen aan de hand van een aantal voorbeelden.



Eigenlijk verdient Aristoteles de eer om de eerste onderzoeker in Kennisrepresentatie en Redeneren te zijn. Hij was de eerste die standaard redeneerpatronen vast legde in zijn syllogismen. Een aantal daarvan kennen we allemaal:

ALS iemand een toga draagt DAN is hij hoogleraar

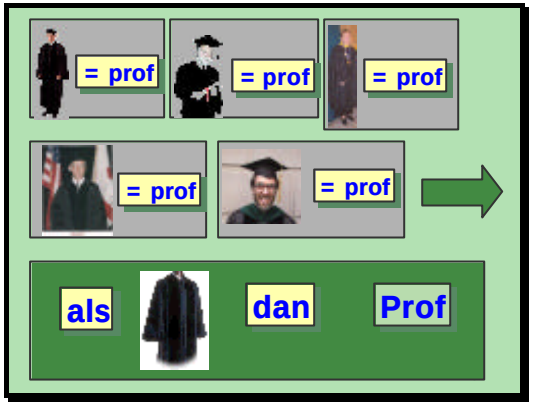
en een tweede voorbeeld:

iemand draagt geen toga OF hij is hoogleraar

En sterker nog: sinds lange tijd weten we dat beide uitspraken eigenlijk hetzelfde betekenen.

Wat hier gebeurt is dat er een structuur aan kennis gegeven wordt (bijv: conditie enerzijds & conclusie anderzijds), en dat er algemene redeneerpatronen worden geidentificeerd waarmee we met die kennis kunnen manipuleren. Met andere woorden: Kennisrepresentie en Redeneren.

Nu zult u zich wellicht afvragen of we sinds de dagen van Aristoteles (in de 3e eeuw voor Christus) we niet een beetje opgeschoten zijn. En het antwoord is natuurlijk ja. Ik geef u wat voorbeelden van recentere datum:

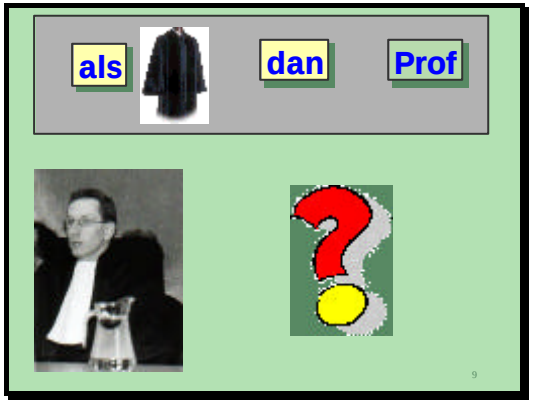


Als we een groot aantal keren iemand tegenkomen in een toga, en elke keer blijkt dat een hoogleraar te zijn, dan kunnen we wellicht de uitspraak uit de voorgaande transparant afleiden.

Ruwweg is dit waar het in Machinaal Leren om gaat. Alweer wordt hier een structuur in de kennis onderscheiden (nl: concrete voorbeelden enerzijds en een algemene regel anderzijds), en probeert men te bepalen hoe we met zulke kennistructuren mogen manipuleren.

Nog een voorbeeld: stel dat we de algemene kennisregel van de vorige slide als waar aanvaarden:

Maar nu komen we iemand tegen die weliswaar een toga draagt, maar waarvan we uit persoonlijke ervaring weten dat ie zeker geen hoogleraar is. Wat moeten we dan?



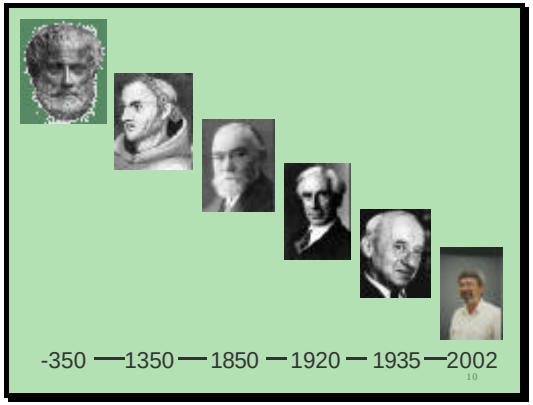
Op een of andere manier moeten we dan onze bestaande kennis herzien, zodat we eerdere geldige conclusies kunnen behouden, maar de nieuwe ongewenste conclusie kunnen laten vallen. Dat probleem heet “belief revision” of “knowledge update”, en is nog lang niet goed opgelost.

In elk van deze gevallen doen we min of meer hetzelfde:

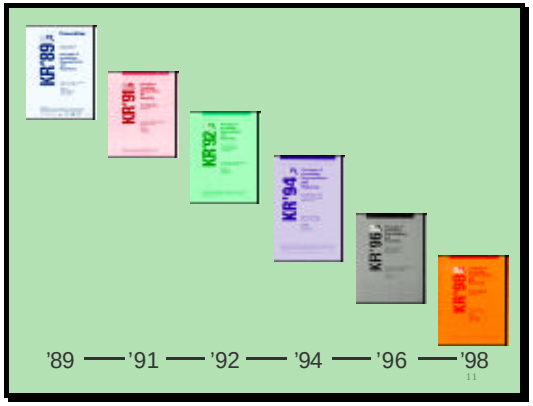
we identificeren structuur in de aanwezige kennis (de “representatie” van de kennis),

en

we bepalen de formele regels volgens welke we mogen manipuleren met die kennis (het “redeneren”).



Zoals al eerder aangegeven is deze manier van denken niet bijzonder nieuw. Ze stamt al uit de tijd van de Oude Grieken, en heeft een eerbiedwaardige geschiedenis, die loopt via de middeleeuwse filosofen, denkers uit de Verlichting, ontwikkelingen in de wiskunde van de 19e eeuw, en de 20ste eeuwse filosofie, en beleeft een grote bloei in de recente decenia dankzij een vruchtbare ontmoeting met de informatica, kortom: De Logica.



De op logica gebaseerde aanpak van het vak Kennis Representatie en Redeneren is zeer dominant binnen mijn vakgebied. Als we de conferentie verslagen van de grootste KennisRepresentatie Conferentie van de laatste 10 jaar er op naslaan, dan blijkt al snel dat vrijwel al het werk binnen de KennisRepresentatie gestoeld is op de logica.

Het heeft natuurlijk grote voordelen om een vakgebied te baseren op zulke stevige fundamenten. Na 2500 jaar logica onderzoek hebben we hele sterke resultaten om ons vakgebied op te bouwen: er bestaat een sterk begrippen apparaat: modeltheorie, bewijstheorie, standaard noties als compleetheid en correctheid; er bestaat veel kennis over de eigenschappen van de verschillende logische formalismen: uitdrukkingskracht, beslisbaarheid, complexiteitseigenschappen, en (mede dankzij de vruchtbare ontmoeting tussen logica en informatica van de laatste decennia) bestaan er ook veel kennis over methoden om logische formalismen op een computer te implementeren: resolutie-bewijzers, tableaux-machines, local-search methoden, etc.

Is er dan alleen maar goed nieuws te melden over het gebruik van logica als het fundament van de KennisRepresentatie?

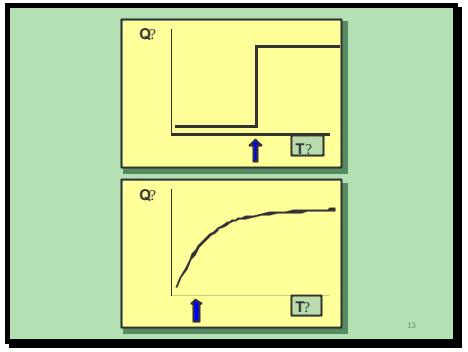
Een belangrijke boodschap van mijn lezing vandaag zal zijn dat het antwoord hierop “NEE” moet zijn: het gebruik van logica als fundament van de KennisRepresentatie heeft weliswaar veel voordelen (zoals hiervoor geschetst), maar heeft ook een belangrijk nadeel.

En dat nadeel kan eigenlijk in één zin worden samengevat:

Logica maakt modellen van foutloze redeneervormen onder ideale omstandigheden.

Logica houdt zich niet bezig met welke consequenties eventuele gemaakte fouten hebben op de uiteindelijke conclusie, en ook niet methoe we zo goed mogelijke conclusies moeten trekken als we niet in een perfecte wereld leven, bijv: wat te doen als er niet genoeg tijd is om een redenering af te maken, of wat te doen als niet alle benodigde kennis beschikbaar is, of wat te doen als (nog erger) niet alle beschikbare kennis correct is.

Het zal duidelijk zijn dat het weliswaar heel nuttig is om modellen te maken van zulk perfect, geidealiseerd redeneren, maar dat het ook een zware beperking is.Een aantal van deze beperkingen kunnen worden samengevat in het volgende diagram:



Als we een computer programma nemen dat gebaseerd is op de momenteel gangbare vormen van kennisrepresentatie, die sterk gestoeld zijn op de logica, en we stellen dit programma een vraag op basis van de in het programma gerepresenteerde kennis, dan vertoont dat programma het volgende gedrag:

Aanvankelijk (en vaak: een hele lange tijd) wordt er door het programma ijverig geredeneerd, en vertelt het programma ons helemaal niks, totdat we opeens van het programm het perfecte antwoord op onze vraag krijgen. Het is natuurlijk prettig dat we er zeker van kunnen zijn dat het programma ons het perfecte antwoord zal geven (op basis van de inzichten in de onderliggende logica), maar het is wel vervelend dat het programma ons in de tussentijd helemaal niks laat weten. In veel omstandigheden zou het veel nuttiger zijn als kennisrepresentaties het gedrag zouden vertonen als in het onderste diagram. In dit geval doet het programma zijn best om ons zo snel mogelijk in ieder geval een gedeeltelijk antwoord te geven, en naarmate er meer tijd beschikbaar is wordt de kwaliteit van het antwoord beter, totdat uiteindelijk het perfecte antwoord gegeven wordt.

Het zal duidelijk zijn dat in dit informele diagram de kwaliteit van het berekende antwoord (Q) uitgezet is tegen de beschikbare rekentijd (T). Maar we kunnen hetzelfde beeld schetsen voor andere aspecten van kennisrepresentatie, bijvoorbeeld de beschikbare kennis, zeg: “K”. Alweer zal een traditionele, op logica gebaseerde kennisrepresentatie ons slechts een antwoord geven als alle benodigde kennis beschikbaar is, en als niet alle kennis beschikbaar is horen we helemaal niets. Het zou (alweer) veel aantrekkelijker zijn als ook al bij een klein beetje beschikbare kennis ons

programma ons al een benadering van het uiteindelijke antwoord zou kunnen geven, en als die benadering geleidelijk beter zou worden naarmate er meer kennis beschikbaar komt.

Om het te zeggen in termen van een metafoor: de traditionele logica, met haar gedrag zoals in het bovenste diagram, geeft ons alleen het "laatste oordeel": perfect, en volmaakt, het oordeel van de perfecte beslisser onder ideale omstandigheden, terwijl wat we eigenlijk nodig hebben is een "eerste oordeel": weliswaar onvolmaakt, en slechts een benadering, maar wel op tijd beschikbaar, en geleidelijk aan verbeterend als we er meer moeite in steken. Of, om het in goed Nederlands te zeggen: "Het eerste oordeel is een daalder waard"

Laat ik met wat voorbeelden illustreren waarom zo'n "eerste oordeel" vaak zo belangrijk is:

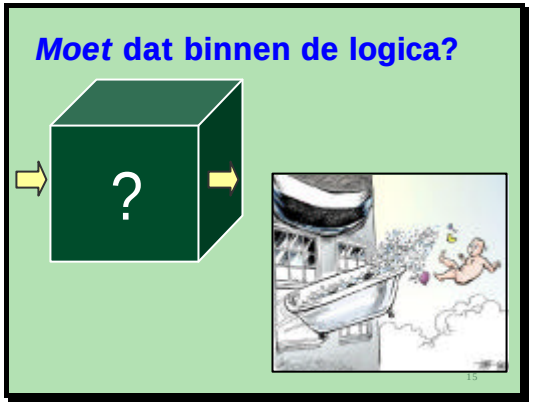
Het meest voor de hand liggende voorbeeld is redeneren onder tijdsdruk. Overal ter wereld worden kennissystemen gebruikt om te zorgen dat complexe industriele processen goed blijven verlopen. En ook bij bijv. onze eigen Hoogovens, hier niet zo heel ver vandaan, is dat het geval. Als er in een industrieel proces iets mis dreigt te gaan is er vaak slechts een kort tijdsinterval waarbinnen een beslissing genomen moet worden. Dan is een redelijk antwoord binnen de vereiste tijd een hoop nuttiger dan het perfecte antwoord achteraf, als de boel al in de luch gevlogen is.

Een iets minder voor de hand liggende reden om graag een goed eerste oordeel te krijgen is bevat in het volgende plaatje. Bevriende artsen hebben me wel eens verteld dat ze met gemak altijd een juiste diagnose konden stellen, als ze maar zouden beschikken over alle benodigde gegevens. Maar juist die gegevens zijn vaak moeilijk te verkrijgen. Een huisarts moet een diagnose stellen op basis van zeer beperkte informatie, en zelfs een specialist zou graag vaak meer meetgegevens hebben. Maar ja, als je voor een aantal van die extra observaties de patient eerst moet opensnijden, dan is het toch ook wel prettig om een goed eerste oordeel te kunnen vellen op basis van beperkte observaties.

Soms heeft het ook helemaal geen zin om door te redeneren tot er een precies "laatste oordeel" gevonden is. Een mooi voorbeeld daarvan is te vinden in het repareren van computers. Computers bestaan weliswaar uit heel duizenden individuele kleine onderdelen, maar in de praktijk kunnen we die individuele kleine onderdelen helemaal niet vervangen. In plaats daarvan is een computer opgebouwd uit een dozijn of wat grote componenten die wel vervangen kunnen worden (geluidskaart, videokaart, etc). Als na een korte periode van diagnose de fout in eerste oordeel herleid kan worden tot één van die kaarten, dan wordt simpelweg die kaart vervangen. Het eerste oordeel dat de schuldige vervangbare kaart aanwijst is dan voldoende; het heeft geen enkele zin om precies het verantwoordelijke onderdeel op de kaart aan te wijzen, want dat uiteindelijke preciese oordeel heeft geen enkele invloed op de actie die we gaan ondernemen (nl. het vervangen van de hele kaart).

Als we dan nu gezien hebben wat de tekortkomingen zijn van de huidige, op logica gebaseerde kennisrepresentaties, dan is natuurlijk de volgende vraag wat we daar aan kunnen doen. En om precieser te zijn, moeten we twee vragen beantwoorden:

1. *Kan* zo'n "eerste oordeel" verkregen worden binnen de logica (dus: kunnen we binnen de grenzen van de logica blijven, en toch het gedrag als in ons 2e diagram verkrijgen), en:
2. *Moet* dat binnen de logica
   (oftewel: waarom zouden we niet de logica helemaal overboord zetten, en geheel andere vormen van kennis-representatie gaan beoefenen).



Eerst maar eens de vraag: "Moet dat binnen de logica"? Immers, er zijn binnen de Kunstmatige Intelligentie ook talrijke andere technieken voorhanden, zoals bijv. Neurale Netwerken, Genetische Algorithmen, statistische methoden, etc. En deze methoden hebben vaak precies het juiste geleidelijke, benaderende gedrag dat we zo missen in de logica. Dus: waarom niet de logica overboord gooien, en deze andere methoden als basis voor de Kennisrepresentatie nemen?

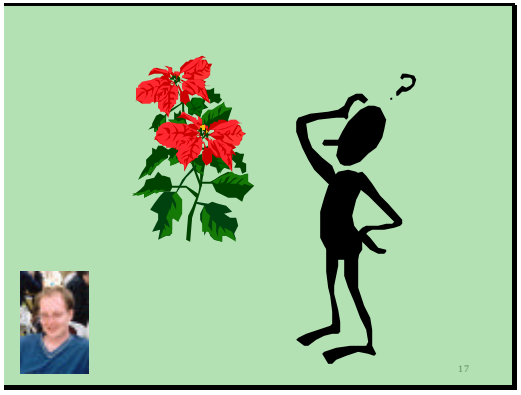
Echter (en met wat ik nu ga zeggen zal ik op de tenen van een hoop van mijn collega's gaan staan), deze alternatieve methoden hebben een belangrijke nadeel. Om dat in te zien moeten we even teruggaan naar ons uitgangspunt: wetenschap = modellen bouwen. Wat willen we van die modellen? We willen niet alleen dat die modellen ons de juiste voorspelllingen geven (= de juiste antwoorden), maar we willen ook graag dat die modellen een zekere verklarende werking hebben, ze moeten ons helpen de beschreven verschijnselen te begrijpen. En juist op dit verklarende aspect van modelvorming is de logica zo sterk. Sommige andere methoden (bijv. Neurale Netwerken) hebben een beetje een "zwarte doos" effect: je stopt er een vraag in, het goede antwoord komt er uit, maar waarom dat antwoord er nu uitkomt,

wat de interne structuur is van de kennis die van de vraag tot het antwoord leidt, daar laten die modellen zich niet zozeer over uit.

De logica daarentegen geeft ons een helder, wiskundig inzicht in de structuur van de onderliggende kennis, en is dus een veel “transparanter” model, met een grotere verklarende werking. We moeten dus proberen om die sterke kant van de logica te behouden in onze zoektocht naar meer geleidelijke, benaderende vormen van redeneren. Met andere woorden, door de logica geheel te laten vallen zouden we de baby met het badwater weggooien. Laten we dus proberen dat niet te doen.

Nu de vraag “*moet dat binnen de logica*” bevestigend is beantwoord rest ons nog de tweede vraag: “*Kan dat binnen de logica*”?

In het resterende deel van deze lezing hoop ik u ervan te overtuigen dat werk wat ik de afgelopen jaren samen met een aantal collega’s heb gedaan laat zien dat het antwoord hierop een luidkeels “JA” is: het is mogelijk om binnen de logica te blijven (met andere woorden: de baby niet weg te gooien), en toch het gewenste geleidelijke en benaderende gedrag te verkrijgen dat ontbreekt in de traditionele logica. We zullen daarvoor kijken naar experimenten die er door mijn collega’s zijn uitgevoerd aan bestaande systemen, we zullen kort de formele wiskundige onderbouwingen noemen, en we zullen twee toepassing schetsen. Laten we maar eens kijken naar wat experimenten, die zijn uitgevoerd door een van mijn promotie studenten, Perry Groot.



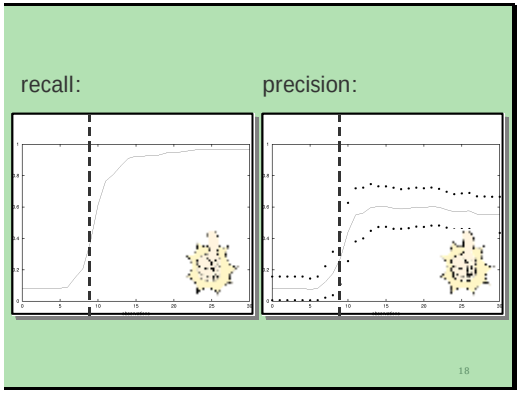
Het betreft hier een classificatie systeem voor planten zoals die worden aangetroffen in Zuid Duitland. Op basis van een groot aantal kenmerken van een plant, zoals de lengte van de steel, de vorm van het blad, de kleur van de bloem, de vochtigheidsgraad van de bodem etc bepaalt het systeem de soort plant waar we mee te maken hebben. Het archetypische voorbeeld van een classificatie systeem.

Het betreffende systeem is gebouwd door een aantal Duitse collega’s, onder leiding van Prof. Frank Puppe, aan de Universiteit van Wurzburg. Op basis van zo’n 40 verschillende eigenschappen (bladvorm, bloemkleur, bloeitijd etc) gebruikt het systeem zo’n dikke 7500 classificatie regels om een groot aantal planten goed te kunnen classificeren.

Als onderdeel van zijn promotie onderzoek vroeg Perry zich af hoe het systeem zich zou gedragen als we eens niet alle observaties in één keer ter beschikking zouden stellen. Met andere woorden: hoe goed zou het systeem werken als het systeem een “eerste oordeel” moest vellen over een plant op basis van slechts een beperkt aantal kenmerken.

Dit is overigens een realistische situatie: het classificatie systeem van Frank Puppe wordt gebruikt door medewerkers van de Universiteit van Wurzburg om vragen te beantwoorden van wandelaars die menen een zeldzame plant gevonden te hebben. Vaak hebben die wandelaars zo’n plant in het weekend gevonden, en bellen ze op Maandag de Universiteit op met hun vraag over de aard van de plant. Tegen die tijd is het natuurlijk niet meer mogelijk om alle benodigde informatie te achterhalen:

Was het een droge of een natte plek waar de plant stond? Oeps, daar had de wandelaar niet op gelet. Hoe lang was de stengel? Oeps, de wandelaar heeft alleen een bloemetje meegenomen naar huis, en niet een hele plant. Is het blad sterk gekarteld of slechts weinig gekarteld? Onze wandelaar is ook maar een leek, en weet vaak het verschil niet eens. Kortom: het is wel degelijk van belang dat het systeem in staat is om een goed “eerste oordeel” te geven op basis van zeer onvolledige informatie.
Wat Perry gemeten heeft zijn twee maten van de kwaliteit van het systeem, geheten “recall” en “precision”.



Wat die nou precies betekenen is niet zo belangrijk, maar ruwweg zeggen ze hoeveel van de goede antwoorden het systeem vindt (de recall) en hoeveel onzin antwoorden het systeem daarnaast nog vindt (de precision). We wilden

dus eigenlijk weten hoe deze twee kwaliteitsmaten zouden groeien naarmate er meer observaties aan het systeem beschikbaar werden gesteld.

En wat bleek, nogal tot onze verbazing: het systeem bleek verbazend robuust te zijn voor ontbrekende informatie. Zoals we in deze grafieken kunnen zien is er een eerste periode waarin het systeem te weinig informatie heeft om veel verstandigs te kunnen zeggen, maar reeds vanaf ongeveer 8 observaties breekt er een periode aan waarin de kwaliteit van de geleverde antwoorden volgens beide kwaliteitsmaten toeneemt tot het uiteindelijke maximum haalbare. En nog verbazender is het feit dat het systeem nauwelijks baat blijkt te heben bij meer dan 15 observaties (uit een maximaal aantal van zo'n 30 relevante observaties per plant). Blijkbaar is het systeem dus zeer wel in staat om een "eerste oordeel" te vellen over de aard van de gevonden plant op basis van slechts een deel van de in principe benodigde informatie.

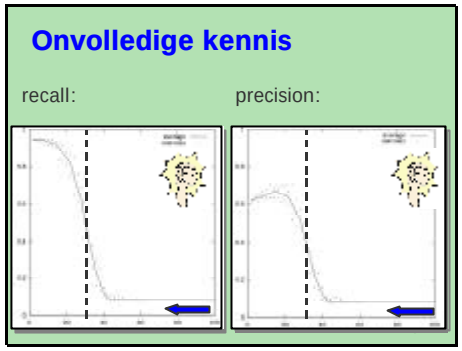
Ik wil hierbij benadrukken dat het systeem van Puppe was helemaal niet met opzet gemaakt om goed te kunnen werken bij ontbrekende input. Het was een fenomeen waar Puppe en zijn collega's helemaal niet zo aan dachten toen ze het systeem ontwierpen. Maar bij experimentatie bleek dat het systeem zich onverwacht robuust gedroeg, immers, het systeem heeft al aan 8-10 observaties genoeg om een goed eerste oordeel te kunnen vellen.

Misschien is de situatie dus helemaal niet zo slecht, en hebben veel van de systemen die we in de kennisrepresentatie ontwerpen en bestuderen al onbedoeld het benaderende en geleidelijke gedrag waar we zo naar op zoek zijn!

Echter, niet altijd is de situatie zo onverwacht gunstig, en dat bleek in een tweede experiment van Perry. U herinnert zich wellicht nog dat we in het algemeen willen bestuderen hoe de kwaliteit van de antwoorden van een systeem verandert als functie van verschillende parameters. Dat kon zijn rekentijd, of beschikbaarheid van gegevens, maar ook de beschikbaarheid van kennis.

Wat zou het gedrag van het systeem zijn als de kennis niet helemaal volledig was? Zou het in staat zijn om een redelijk "eerste oordeel" te vellen op basis van onvolledige kennis?

Ook hier heeft Perry naar gekeken, maar helaas met minder gunstige resultaten.

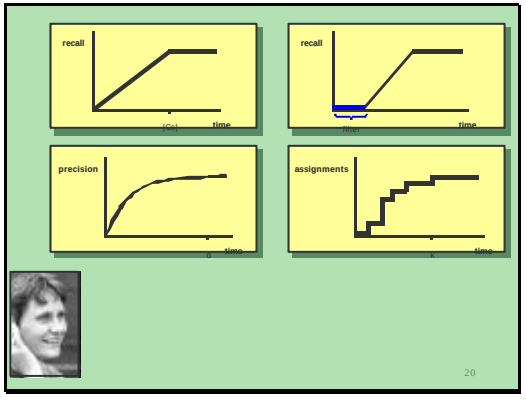


Deze grafieken laten zich het beste van rechts naar links lezen: helemaal rechts is 100% van de kennisbank verwijderd, en levert het systeem vanzelfsprekend zeer slechte antwoorden. Als we geleidelijk aan naar links schuiven is een steeds kleiner percentage van de kennisbank verwijderd (of, zo u wilt: wordt er steeds meer kennis aan de kennisbank toegevoegd), met als gevolg dat uiteindelijk het systeem goede kwaliteit antwoorden gaat leveren.

Het slechte nieuws in dit geval is dat het "eerste oordeel" wel erg lang op zich laat wachten: de recall en precision nemen pas redelijke waarden aan als de kennisbank voor 80% of meer compleet is, en dan is ook al bijna de tijd van het uiteindelijke perfecte oordeel aangebroken, dus is de winst minimaal.

Perry heeft dus in detail naar één specifiek systeem gekeken, en de grafieken die hij heeft bepaald heten ook wel "kwaliteitsprofielen".

Samen met Annette ten Teije hebben we nog een groot aantal andere systemen op hun kwaliteits-profielen geanalyseerd.



Weliswaar ontbrak ons de tijd, geduld en menskracht om voor al die systemen even preciese kwantitatieve experimenten te doen als Perry gedaan heeft, maar een kwalitatieve analyse van een aantal systemen leverde toch verbazende inzichten op. Zoals u ziet komt er “in de natuur” een verbazende varieteit aan kwaliteitsprofielen voor. En alweer is het opmerkelijk dat geen van de geanalyseerde systemen met opzet ontworpen was om zo’n mooi profiel te hebben, maar bleek dat veel van deze systemen als het ware “per ongeluk” al in staat waren om een goed “eerste oordeel” te geven op weg naar het uiteindelijke perfecte antwoord.

Na het uitvoeren van een aantal experimenten is het als onderzoekers onze taak om te proberen een model op te stellen dat deze gegevens kan verklaren, en vervolgens op basis van voorspellingen nieuwe experimenten kan suggereren. Het is nu dus tijd om eens te kijken naar zulke formele modellen.

Ook dat heb ik uitgebreid gedaan samen met Annette.

$$\vdash_3^{\emptyset} \Rightarrow \vdash_3^{S} \Rightarrow \vdash_3^{S'} \Rightarrow \vdash_2 \Rightarrow \vdash_1^{S'} \Rightarrow \vdash_1^{S} \Rightarrow \vdash_1^{\emptyset}$$
$$\nvdash_2 \Leftarrow \nvdash_1^{S'} \Leftarrow \nvdash_1^{S} \Leftarrow \nvdash_1^{\emptyset}$$

$$\emptyset = \{ABD_1^{\emptyset}\} \subseteq \{ABD_1^{S}\} \subsetneq \{ABD^2\}$$
$$\{ABD^2\} \supsetneq \{ABD_3^{S}\} \supsetneq \{ABD_3^{\emptyset}\} = \emptyset$$

$$0 = |\{ABD_1^{\emptyset}\}| \leq |\{ABD_1^{S}\}| \leq |\{ABD_2\}|$$
$$|\{ABD_2\}| \geq |\{ABD_3^{S}\}| \geq |\{ABD_3^{\emptyset}\}| = 0$$

We hebben, voortbordurend op werk van onze Italiaanse collega’s Marco Schaerf en Marco Cadoli, een aantal mooie afschattingsstellingen bewezen die precies karakteriseren hoe voor bepaalde problemen de kwaliteit van de oplossing toeneemt als functie van de rekentijd, allemaal netjes binnen een kader dat maar een heel klein beetje anders is dan de traditionele logica van Frege en Tarski.

Deze stellingen zeggen dat de klassieke oplossing (steeds gelabelled met subscript “2”) van twee kanten ingesloten kan worden door een reeks steeds nauwkeuriger niet-klassieke oplossingen, zowel door een incomplete benadering (die dus te weinig, of te kleine oplossingen berekent) en door een overcomplete benadering (die te veel, of te grote oplossingen berekent).

Ruwweg, waar deze stellingen op neerkomen is dat we de traditionele logica uitgerust hebben met een soort “stelknop”. Die stelknop bepaalt de “kwaliteit” van de antwoorden die onze logica geeft, met de klassieke logica als ideale maatstaf. Hoe hoger we de stelknop instellen, hoe beter onze logica wordt (hoe meer hij gaat lijken op de klassieke logica). En natuurlijk dicteert de “wet van behoud van ellende” dat zulke betere resultaten komen met als de prijs een verhoogde rekentijd. Onze stellingen karakteriseren precies de gevolgend van de verschillende standen van de stelknop voor de kwaliteit van de antwoorden.

Tegenwoordig is een wetenschapper niet meer klaar als hij goede experimenten heeft uitgevoerd en mooie formele modellen heeft opgesteld om die de experimenten verklaren. Vervolgens eist de maatschappij (in de vorm van de subsidiegevers) dat die resultaten ook daadwerkelijk toepassing vinden.

Hiervoor wil ik werk aanhalen van een collega in mijn groep, Heiner Stuckenschmidt, en dat werk betreft het Semantic Web. Ruwweg komt het Semantic Web neer op het idee dat computers veel meer van de inhoud van web-pagina’s zullen gaan begrijpen dan ze nu doen, en dat ze ons daardoor veel beter kunnen ondersteunen bij
allerlei taken op het Web die we nu nog zelf moeten doen, zoals web-pagina’s opzoeken, informatie uit verschillende pagina’s combineren, prijzen van artikelen uit verschillende Webwinkels vergelijken, etc.

Laat ik u daarvan een voorbeeld geven.

Hier zien we de webpagina van een van de beste zoekmachines op het huidige Web: Google, waarop ik de ego-test heb uitgevoerd. Google blijkt maar liefst 21.200 antwoorden te kunnen vinden over “Harmelen”. Tot mijn grote tevredenheid is mijn eigen homepagina de het eerste antwoord dat Google teruggeeft. Maar ietwat tot onze verbazing zien we dat de volgende drie antwoorden helemaal niet over mij gaan, maar over verschillende aspecten van het dorpje “Harmelen”, dat te vinden is ergens niet ver van Utrecht. Hier is dus iets eigenaardigs gebeurd: ik bedoelde met mijn zoekopdracht de *persoon* “Harmelen”, maar ik kreeg ook antwoorden terug over het *dorpje* “Harmelen. Met andere woorden, de zoekmachine heeft in zekere zin de bedoelde “betekenis” van mijn zoekopdracht niet echt goed begrepen.

De droom van het Semantic Web is om het Web zo in te richten dat het bevolkt zal raken door programma’s die elkaar wel goed zullen begrijpen, en goed zullen kunnen samenwerken, bijvoorbeeld een samenwerking tussen mijn persoonlijke programma en een zoekmachine, (zoals hier), of een samenwerking tussen mijn persoonlijke programma en een webwinkel, of een samenwerking tussen mijn persoonlijke programma en een soortgelijk persoonlijk programma van een collega elders ter wereld, etc.

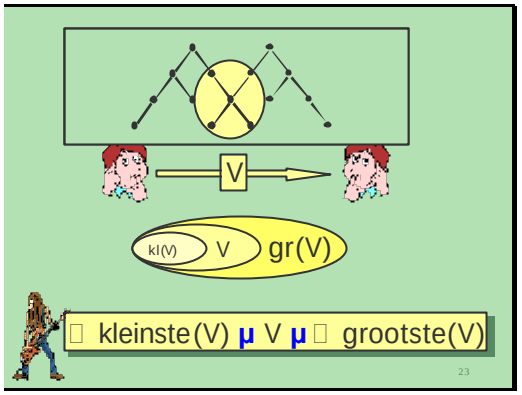
Om die samenwerking tussen programma’s goed te laten verlopen, is het niet anders gesteld dan met samenwerkingen tussen mensen: een eerste vereiste is dat je dezelfde woorden gebruikt, of in elk geval tot op zekere hoogte.

Welnu, het huidige Web is enorm divers, en het Semantic Web zal niet minder divers zal zijn. Een direct gevolg daarvan is dat het onwaarschijnlijk is dat twee programma’s die elkaar tegenkomen op het Web ook inderdaad dezelfde woorden zullen gebruiken.

Hoe moet dan nou met de communicatie van deze twee programma’s? Het inzicht van Heiner is dat, alhoewel die twee programma’s weliswaar niet perfect dezelfde taal zullen spreken, ze wellicht wel voor een deel dezelfde taal zullen spreken.

Immers, ik deel bepaalde interesses met de collega elders ter wereld, ik ben geinteresseerd in de producten van de Web winkel, etc. Kortom: een deel van het vocabulair van mijn persoonlijke programma en het programma van de andere partij zullen wel degelijk overlappen. Een als die twee programma’s een deel van elkaars woordenschat begrijpen, misschien kunnen ze dan nuttig met elkaar communiceren door gebruik te maken van alleen de overlappende delen van hun woordenschat.

Wat er nu in de theorie van Heiner gebeurt is als volgt:

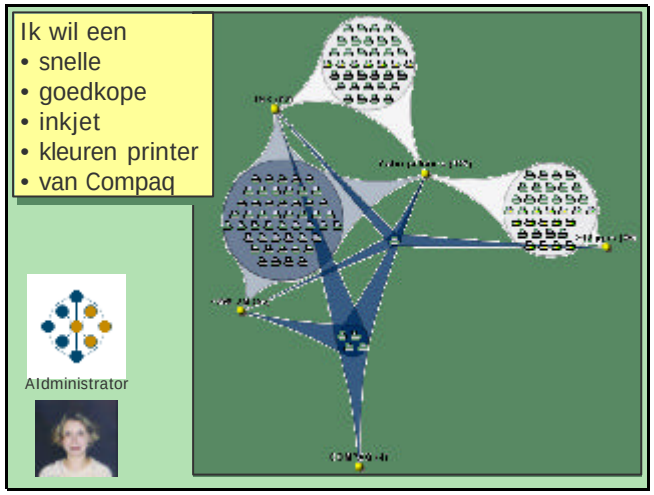


Als het ene programma, de “vragensteller” een vraag *V* wil stellen aan het andere programma, de “beantwoorder”, dan zal het vragende programma de vraag eerst moeten uitdrukken gebruik maken van alleen het gedeelde vocabulair. Anders kan het andere programma de vraag immers niet begrijpen, laat staan beantwoorden. Maar in het algemeen zal het niet lukken om de oorspronkelijke vraag *V* precies in het gedeelde vocabulair uit te drukken. Immers: slechts een deel van het vocabulair overlapt met dat van het andere programma. De oorspronkelijke vraag *V* zal dus door de vragensteller vertaald moeten worden in een vraag die (losjes gesproken), iets “kleiner” is dan de oorspronkelijke vraag *V*, of in een vraag die iets “groter” is dan de oorspronkelijke vraag *V*.

Beide varianten (zowel de kleine vraag als de grote vraag) zijn benaderingen van de oorspronkelijke vraag *V*. We zouden nu één van beide benaderingen (of zelfs: allebei de benaderingen) aan de beantwoorder kunnen sturen. Die kan een antwoord geven op de benadering van de oorspronkelijke vraag. Het antwoord op deze benaderende vraag is hopelijk ook een goede benadering van het antwoord op de oorspronkelijke vraag, waarmee een succesvolle benaderende communicatie is volbracht.

Feitelijk is dit niets anders dan wat er in ons eigen alledaagse sociale leven gebeurt: we delen geen van allen precies dezelfde woordenschat, maar door onze communicatie uit te drukken in het gezamelijk gedeelde vocabulair lukt het ons toch om redelijk te communiceren.

En ook hier is het weer mogelijk om preciese wiskundige karakteriseringen te geven van hoe goed de benadering van de oorspronkelijke vraag aan weerszijden is.



Deze lezing zou niet compleet zijn zonder een voorbeeld uit het werk van mijn goede collega's en vrienden bij Aidministrator, in samenwerking met Marta Sabou. Ik noemde net al "webwinkels" als een toepassing waar "benaderend" redeneren een rol zou kunnen spelen. Laten we als laatste nog naar dat voorbeeld kijken. Bij AIdministrator heeft men in het verleden gewerkt voor een webwinkel in die printers verkocht. Op de website van die winkel kon je aangeven in wat voor printer je geinteresseerd was. Laten we eens een gemiddelde bestelling beschouwen:

"Ik wil een snelle, goedkope, inkjet kleuren printer van Compaq".

Er is echter een probleem met deze bestelling. De kenners onder u zullen het al wel gezien hebben: zulke printers bestaan helemaal niet. Zoals zoveel klanten hebben ook wij een beetje teveel wensen op ons lijstje: we kunnen wel snelle kleuren printers krijgen, maar dan zijn ze niet goedkoop, en Compaq maakt al helemaal geen snelle printers. Als we deze bestelling plaatsen bij de gemiddelde webwinkel is het antwoord dan ook eenvoudigweg "nee": "we hebben geen geschikt product voor u". Klant teleurgesteld, winkel niks verkocht.
In een gewone winkel zou dat anders gaan: dan werd ons geen "NEE" verkocht, maar was er een verkoper, die uitlegt dat ie weliswaar niet precies heeft wat we zoeken, maar wel iets wat daar op lijkt, en zouden we daar misschien in geinteresseerd zijn?

Bij AIdministrator zijn methodes ontwikkeld om net zo'n effect te verkrijgen, maar dan voor webwinkels. Als we onze printer-bestelling intypen krijgen we niet zomaar "NEE" te horen, maar wordt ons het bovenstaande plaatje getoond. Dit figuur geeft aan wat onze oorspronkelijke eisen zijn: snel (meer dan 12 pagina's per minuut); goedkoop (minder dan 700 DM, het voorbeeld is enigzins gedateerd:-); inkjet; kleuren printer; van Compaq.
We zien direct dat er geen enkele printer is die aan alle vijf de eisen voldoet. Maar interessanter is dat ons ook direct getoond wordt dat er wel een printer is die aan 4 van de 5 eisen voldoet, als we de eis maar laten vallen dat het een Compaq moet zijn. En als we perse een Compaq willen, zijn er 4 printers die aan alle eisen voldoen behalve de snelheid. En als we nog meer water bij de wijn doen, en maar 3 van de 5 eisen vervuld willen hebben, dan is er opeens een grote keuze aan printers.

Kortom: het preciese antwoord op de bestelling was "NEE, zo'n printer bestaat niet", maar ook hier blijkt een benaderend antwoord veel nuttiger te zijn dan het preciese antwoord.

We bereiken zo langzamerhand het eind van deze rede. Ik herinner u nogmaals aan wat ik heb geprobeerd te doen:

- we hebben gezien dat het vakgebied Kennisrepresentatie stevig is gestoeld op de logica
- we hebben gezien dat dit stevige fundament ook een aantal belangrijke nadelen met zich meebracht,
- en met name het gebrek aan geleidelijk en benaderend redeneergedrag was een belangrijke beperking
- we hebben ons afgevraagd of het mogelijk zou zijn om, zonder de baby met het badwater weg te gooien, dus: met behoud van de mooie eigenschappen van logica, deze beperkingen op te heffen, en wel een mooi geleidelijk en benaderend redeneergedrag te verkrijgen.
- daartoe hebben we gekeken naar een aantal experimenten aan bestaande systemen, en die bleken zich netter te gedragen dan verwacht,
- ik heb gemeld dat er nette wiskundige formaliseringen mogelijk zijn van dit gedrag, gestoeld op de klassieke logica,
- en ik heb laten zien dat er in concrete toepassingen, zoals bijvoorbeeld in de wereld van het Semantic Web en e-commerce, behoefte is aan deze benaderende technieken.

Tot slot:

Ik heb eerder de metafoor van het eerste en het laatste oordeel gebruikt, om het kontrast weer te geven tussen wat logica ons geeft, en wat kennisrepresentatie eigenlijk nodig heeft.

Het lijkt me daarom dan ook alleen maar passend om af te sluiten met het volgende beeld:

het “*Laatste Oordeel*” geschilderd door Michelangelo op het plafond van de Sixtijnse kapel, en daarbij op te merken dat het blijkbaar in de kennisrepresentatie niet anders is dan elders in het leven, namelijk:

> dat het laatste oordeel weliswaar volmaakt zal zijn,
> maar dat het te laat komt om er nog iets aan te hebben,
> en dat we daarom beter af zijn met een eerste oordeel
> dat nog op tijd komt om er iets nuttigs mee te kunnen doen.

Ik heb gezegd.