

Wie is je AI held?

Frank van Harmelen over Alan Bundy

De Connectie
nummer 1, jaargang 4, Maart 2009

Joris de Ruiter, AI Masterstudent aan de VU
(jdruiter@few.vu.nl)

Alan Bundy is professor binnen de informatica-afdeling van de Universiteit van Edinburgh. Hij is voornamelijk bekend om zijn bijdragen aan *automated reasoning* waaronder *theorem proving* en *proof-planning*. Automated reasoning is het begrijpen en implementeren van wiskundig redeneren. Dit redeneren wordt weergegeven in logische en wiskundige symbolen en het vakgebied is dan ook onderdeel van de symbolische tak van AI. Subgebieden zijn *automated theorem proving* (het geautomatiseerd bewijzen van proposities) en *automated proof checking* (het checken of deze bewijzen kloppen). Dit laatste veld heeft een groot probleem: hoe te zoeken in de explosief grote zoekruimte van mogelijke proposities. Een oplossing hiervoor is proof planning. Dit geeft een plan om tot een bewijs te komen, waarna het plan wordt uitgevoerd en de propositie wordt bewezen. Hierbij wordt rekening gehouden met veelvoorkomende patronen en veelgemaakte fouten in bewijzen.

Alan's 40 pagina's tellend CV is indrukwekkend. Met 52 onderzoeksvoorstellen, 5 boeken en 190 papers op zijn naam, is hij volgens Frank "al heel lang heel goed". "Veel grote onderzoekers hebben een geniale periode en glijden daarna af, maar de resultaten van Alan zijn consistent goed".

Maar het is vooral op persoonlijk vlak dat Frank hem waardeert. Hij is standvastig, collegiaal, en staat altijd open voor discussie. De laatste twee eigenschappen zie je bij veel bèta-wetenschappers, maar standvastigheid is niet iets wat elke onderzoeker heeft. Veel wetenschappers rijden mee op de schommelende onderzoekstrein van wisselende overheids gelden. Verandert de focus van het publiek, - en daarmee de geldpot van de overheid - dan verandert hun onderzoeksterrein. Enerzijds is dit goed omdat wetenschap zo in dienst blijft staan van de samenleving, maar anderzijds is de standvastigheid van Alan Bundy iets wat Frank erg kan waarderen.

In de tijd dat Frank onderzoek deed binnen Alan's onderzoeksgroep aan de University of Edinburgh, heeft hij behalve over AI veel geleerd over zijn manier van leiding geven. "Hij geeft mensen de ruimte, maar zorgt ervoor dat ze toch binnen zijn onderzoeksgebied blijven". Ook laat hij zich niet voorstaan op zijn status en autoriteit en

De persoonlijke held van prof Frank van Harmelen is Alan Bundy. Niet de couch potato uit de Amerikaanse serie, niet de Amerikaanse seriemoordenaar, maar de Britse onderzoeker in logica en AI. Een gesprek over leiderschap en symbolisme.

is hij altijd bereid tot discussie. Slechte ideeën mogen gerust, er moet geen drempel zijn om met ideeën te komen. Samenwerking was heel belangrijk en werd gestimuleerd door samen te lunchen en door elke week een bespreking te houden waar-

in de onderzoekers vertelden over hun ervaringen. "Soms was er veel te vertellen, soms weinig, maar het zorgde in ieder geval voor groepsbinding". Door deze besprekingen kenden onderzoekers elkaars werk en was er meer ruimte voor samenwerking en het delen van gedachten. Volgens van Harmelen is dit lang niet overal zo. "Elders lunchen PhD's in hun kamer om maar geen tijd te verliezen".

Behalve de wekelijkse bespreking en het samen lunchen, was er nog een andere manier om samenwerking te bevorderen: het opschrijven van ideeën en gedachten in een gezamenlijk Blue Book wat alleen binnen de afdeling ter inzage was. Dit had als doel je gedachten snel op te schrijven voor je ze vergat, ze te delen met andere groepsleden, en ze te structureren door ze op te schrijven. Belangrijk bij het Blue Book is dat er geen ondergrens is aan kwaliteit. Halve oplossingen, problemen zonder oplossing en huidige gedachten, het mag allemaal.

Veel van de manieren en houding van Alan Bundy zie je terug in van Harmelen. Ook Frank is collegiaal en heeft geen moeite met discussie. De deur naar zijn kamer staat altijd open en studenten mogen hem gerust met 'je' aanspreken. Doelbewust heeft hij de manieren van Alan Bundy in zijn eigen onderzoeksgroep in praktijk gebracht. De wekelijkse bespreking, het samen lunchen en het creëren van samenwerking is gelukt, het Blue Book niet. De reden hierachter is van Harmelen nog steeds niet duidelijk.

De invloed van Alan op Frank is duidelijk, maar invloed stroomt de laatste tijd ook de andere kant op. Sinds 1998 is van Harmelen pionier op het Semantic web, en langzaam begint ook Alan Bundy hier aan te werken.

Symbolisme en statistiek:

Gedurende zijn carrière heeft Alan de overstap gemaakt van logica naar AI. Hetzelfde werd gedaan door veel andere onderzoekers van die tijd, maar er waren ook velen die overstapten naar de statistische kant, de 'nieuwe AI', het veld van machine learning, neurale netwerken en evo-

lutionaire algoritmen. Het heeft geleid tot een scherpe deling van onderzoeksvelden. De symbolische en statistische AI-kant, met elk aparte termen, succes-criteria, wiskundige gereedschappen, paradigma's, toepassingen, afdelingen en tijdschriften. Daarmee is het niet alleen een inhoudelijke scheiding geworden, maar ook een culturele.

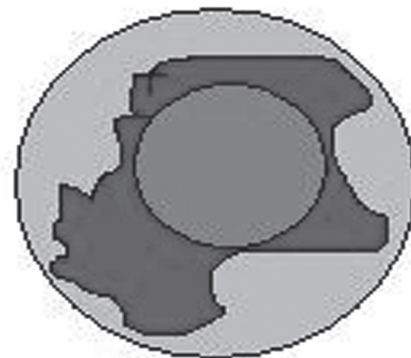
De scheiding werd persoonlijk voor van Harmelen toen collega Maarten de Rijke overstapte van de symbolische naar de statistische kant. Nog altijd discussiëren ze over welke kant beter is. Een goede plek hiervoor zijn workshops waarin sprekers worden uitgenodigd om hun visie te geven op het AI-veld. Naast het aanprijzen van eigen onderzoek kunnen sommigen het niet laten ook even te steken naar de tegenpartij. Wat is er immers frustrerender dan collega-onderzoekers die hopeloos de 'verkeerde' kant op werken?

Frank zelf is symbolist gebleven. "De beste manier van modelleren is via wiskunde. De beste manier om menselijk redeneren te modelleren is via logica". "Statistische methoden als neurale netwerken zijn een black box. Ze werken, maar je weet niet hoe ze werken. Ook kunnen ze niet goed omgaan met kleine aantallen". "Een mens weet al na één auto-ongeluk dat je uit moet kijken voor auto's; neurale modellen moet je eerst tien keer trainen". Omgekeerd beschuldigt de statistische kant het symbolisme ervan niet goed om te kunnen gaan met grote aantallen en met onzekerheid in data.

Hier zien we precies de sterkten en zwakten van beide methoden. In een beperkt domein kan logica helder de relaties weergeven, maar bij grote hoeveelheden feiten wordt het systeem langzaam en is de kans op inconsistenties groter. Statistische modellen daarentegen werken juist goed bij grote aantallen. Hoe groter de dataset, hoe zekerder de uitkomst. Dit geldt voor datamining, neurale netwerken en evolutionaire algoritmen.

Hoewel Frank zijn heil blijft zoeken in symbolisme, geeft hij toe dat logica te smal is om alle problemen te dekken. Hij biedt drie oplossingen: (1) het uitbreiden van logica zodat het om kan gaan met onzekerheden. Dit bestaat al in de vorm van *fuzzy logic* waarin iets voor 1/3e waar kan zijn (in tegenstelling tot booleaanse logica waarin iets alleen waar of onwaar kan zijn), maar dit maakt gebruik van getallen en waarschijnlijkheden, en waar haal je die vandaan? Een beter idee lijkt hem het gebruik van approximate reasoning en rough logic. Hierbij worden alle uitspraken opgedeeld in verschillende ruimtes: een ruimte waar de uitspraak zeker waar is, een ruimte waarin de uitspraak zeker onwaar is en een onzekere ruimte.

Figuur 1. Rough logic. Het grijze gebied in het midden geeft de ruimte van mogelijke oplossingen. Het donkergrijze gebied geeft de oplossingen aan die 'zeker waar' zijn; het lichtgrijze gebied de oplossingen die zeker 'niet waar' zijn.



Hoe kleiner de onzekere ruimte, hoe preciezer het antwoord. Een tweede oplossing (2) is het gebruik van logica in onlogische toepassingsgebieden. "Een baksteen gebruikt geen differentiaalvergelijkingen om te vallen, maar toch kun je differentiaalvergelijkingen prima gebruiken om de valbeweging te modelleren. Net zo goed zijn emoties niet logisch, maar kun je wel logica gebruiken om ze te beschrijven. Prof John-Jules Meyer uit Utrecht heeft dit gedaan door emoties als 'angst' en 'woede' op te nemen als logische termen en hiermee te redeneren.

Een laatste oplossing (3) is een integratie van symbolische en statistische methoden. Mensen aan beide kanten van het debat beseffen ook best dat de waarheid in het midden ligt, maar toch is dit midden nog nauwelijks onderzocht. Pogingen worden wel ondernomen, maar er zijn nog geen doorbraken. Van Harmelen zou niet weten waar te beginnen. "De onderzoeksgebieden zijn zo verschillend, het is een grote uitdaging".

Toekomstplannen

Frank van Harmelen werkt graag verder aan het semantisch web, een laag bovenop het huidige web wat computers webpagina's moet laten begrijpen zoals mensen ze begrijpen. Een probleem hierbij is de grootte en complexiteit. Miljarden pagina's met incomplete beschrijvingen, hoe zoek je daarin? En hoe ga je om met onzekerheden en tegenstrijdigheden? Een opzet hiertoe wordt gedaan met LarKC1, een platform voor het op grote schaal redeneren op het semantische web. Met een *computercluster* en een '*computing at home*'-programma wordt gepoogd het schaalbaarheidsprobleem op te lossen en met methoden als *information retrieval*, *machine learning*, *information theory*, *databases*, en *probabilistic reasoning* wordt geprobeerd om te gaan met incomplete kennis. Onbewust wordt hiermee meteen een combinatie getoond van zowel symbolische als statistische methoden. ∅

Over Frank van Harmelen

Frank van Harmelen is professor in de *Kennisrepresentatie & Redeneren* groep binnen de AI-afdeling van de Vrije Universiteit. Na het studeren van wiskunde en informatica in Amsterdam, verhuisde hij naar Edinburgh, waar hij samenwerkte met Alan Bundy aan *proofplanning* en een PhD ontving voor zijn werk naar meta-level redeneren. Hierna werkte hij bij de UvA aan reflectie in expert systemen, en voegde zich in 1995 bij de AI-afdeling van de Vrije Universiteit. Momenteel is hij daar professor en leider van de onderzoeksgroep Kennisrepresentatie & Redeneren. Zijn interesse gaat uit naar Approximate reasoning, het Semantic Web, en het representeren van medische kennis. Momenteel werkt hij aan *LarkC¹* (Large Knowledge Collider), een platform voor het op grote schaal redeneren op het semantisch web.



Frank van Harmelen

Over Alan Bundy

Alan Bundy is al 37 jaar werkzaam aan de Universiteit van Edinburgh. Eerst in de wiskunde-onderzoekseenheid, wat veranderde in Computational Logic, Artificial Intelligence en uiteindelijk in School of Informatics. Zijn carrière volgde precies hetzelfde pad. Een bachelor in wiskunde, een PhD in wiskundige logica, en daarna een onderzoekscarrière binnen de Kunstmatige Intelligentie. Naast zijn vele boeken, papers en onderzoeksvorstellen, wordt hij gezien als *founding fellow* van AAAI en ECCAI. Zijn onderzoek richt zich op automated reasoning waaronder theorem proving en proofplanning. Ook doet hij onderzoek naar redeneerprocessen als proof methods, decision procedures, analogical reasoning, abstraction, en redeneren met onzekerheid. Dit werk heeft hij vervolgens toegepast in gebieden als mathematische inductie, analyse, mechanische problemen, het bouwen van ecologische modellen, game playing en cognitieve modellering.



Alan Bundy

1. The Large Knowledge Collider, <http://www.larkc.eu>