

## **Samenvatting Intelligence without representation**

Onderzoek naar kunstmatige intelligentie is tot halt gekomen doordat er geen goede centrale besturing is binnen een AI. In dit artikel wordt behandeld hoe de schrijver hiermee om is gegaan bij het bouwen van een zelfsturende AI.

AI startte met het doel om menselijke intelligentie na te maken. Dit bleek lastiger dan gedacht. Kleinere subdoelen werden wel behaald, maar de focus ligt tegenwoordig meer op “intelligent assistants”. Subdoelen als zicht of spraak wordt nog wel flink aan gewerkt. De schrijver denkt dat het niet mogelijk is om deze subdoelen te combineren en zo een menselijke AI te maken. Hij pleit voor de volgende manieren om AI's te maken:

- De gemaakte AI's moeten gehele systemen zijn, niet subonderdelen. Alleen dan weet je of ze echt goed zullen werken;

- Alle AI's moeten gemaakt en getest worden in de ‘echte’ wereld zodat ze ook daadwerkelijk werken ook buiten simulaties.

Volgens deze wijze heeft de schrijver verschillende autonome robot's gemaakt en ze zijn tot de volgende conclusies gekomen:

- Bij het maken van simpele intelligenties blijkt een centrale besturing niet te werken en een model van de wereld ook niet;

- Er wordt verwacht dat het niet helpt om een centrale besturing te gebruiken

We hebben al bewijs dat er intelligente systemen bestaan, namelijk de mens. Over 3.6 miljoen jaar evolutie is de mens zo langzaamaan ontstaan. Dit begon bij een-cellige organismen en eindigde uiteindelijk in de mens. Dit suggereert dat intelligente eigenschappen redelijk simpel zijn als een intelligentie eenmaal bestaat en kan reageren. De essentie hierbij is het overleven in een dynamische omgeving. Dit gedeelte van intelligentie kostte ook evolutie de meeste tijd. De basis is de mogelijkheid te bewegen, observeren en survival gerelateerde taken uit te voeren.

De schrijver vertelt in het volgende stuk een verhaal over vroege vliegtuig pioniers, de boodschap achter die verhaal is dat men wel een groot einddoel kan hebben maar als men de basis niet begrijpt maakt het niet uit hoe veel subonderdelen je wel beheerst, de basis is fundamenteel voor AI.

Een AI krijgt vaak niet zijn succes toegekend, als niemand weet hoe het probleem kan worden opgelost is het een AI probleem, echter als het probleem eenmaal is opgelost blijken het gewoon een aantal algorithmes te zijn en noemt men het geen AI meer.

Volgens de schrijver ligt dit aan de programmeurs zelf. De programmeurs in die tijd maakten systemen in een virtuele wereld en maakte het heel makkelijk en duidelijk voor deze systemen. De systemen werkte wel, maar in de echte wereld zouden ze alsnog falen. Wat de AI's niet heel goed maakte. Voor de AI's in die tijd was de wereld nog te abstract. In hun veilige virtuele omgeving hadden ze goede kans om te werken maar in real life waren ze nog niet goed genoeg. De abstractie is juist het moeilijkste gedeelte wat een werkelijk intelligent systeem wel kan oplossen. Een AI in die tijd kan enkel doen waar hij voor gemaakt is, maar ziet niet het grotere plaatje. Denk bijvoorbeeld aan MYCIN die enkel bacterie problemen kan oplossen maar niet als iemands aorta open ligt. De wereld wordt zo gecomprimeerd dat een AI deze net zo zal zien als de mens. Echter dat is natuurlijk niet perse de beste visie (Merkwelt). Onze perceptie is nou eenmaal ook niet foutloos en vanaf onze eigen dimensie. De mens ziet de wereld niet zoals bijvoorbeeld een muis dit zou doen. Het kan ook heel goed zijn dat wij andere clues gebruiken voor perceptie dan dat wij denken.

De schrijver wil ‘Creatures’ maken, hij bedoel hiermee een AI die samen met de mens op zichzelf in de echte wereld kan leven. Hij doet dit omdat hij ervan wil leren over AI's.

Hij wilde dat zijn Creature voldeed aan de volgende eisen:

- Een Creature moet in beperkte tijd op de goede manier met zijn omgeving om kunnen gaan;
- Een Creature moet robuust zijn, als zijn omgeving verandert moet zijn gedrag niet stoppen;
- Een Creature moet meerdere doelen kunnen bewerkstelligen, afhankelijk van zijn situatie;
- Een Creature moet een doel hebben om te bestaan.

Het Creature wordt gebouwd vanuit een engineering perspectief, zijn idee wordt opgesplitst in onderdelen, de onderdelen worden gebouwd en deze worden weer op de juiste manier samengesteld.

AI's werden tot dan toe het meest gebouwd met een centraal besturingssysteem. De schrijver denkt dat dit geen goede oplossing is en stapt hier vanaf. Hij denkt dat het te ingewikkeld is om zo'n systeem te maken.

Hij is meer geïnteresseerd in actieve subsystemen die elkaar allemaal aansturen. Deze werkt dan in verschillende lagen die ieder een activiteit produceren. Een activiteit is een patroon van interacties met de wereld en word ook wel een skill genoemd. Het is een stuk makkelijker om zo'n intelligent systeem te maken aangezien je elke keer een klein stukje kan toevoegen aan het al bestaande systeem. De lagen kunnen parallel aan elkaar werken waardoor de robot prima autonoom zijn taak kan vervullen.

Door meerdere lagen te gebruiken in de robot is het niet nodig voor de robot om 'het grote plaatje' van de wereld te zien. Wat de robot beter maakt om te werken in elke situatie. Door het simpel te houden konden de Creatures veel metingen doen en waren ze altijd 'up to date' en konden ze op tijd de juiste handeling uitvoeren. Dit zou niet het geval zijn als er met een complexe representatie gewerkt zou worden. Door geen centraal besturingssysteem te hebben en het zicht niet te verwerken zal de Creature een grotere kans hebben om in een andere omgeving te kunnen werken dan een AI met een centrale besturing. Elke laag in de Creature heeft zijn eigen doel, met eigen subdoelen. Door de echte wereld te gebruiken voor het bepalen of de doelen te behalen zijn werkt de Creature beter in de echte wereld dan ieder ander AI. De Creature werkt simpeler maar toch beter.

Voor een AI hoeft het niet nodig te zijn om een representatie te hebben van ofwel de wereld ofwel zijn doel om dit doel om intelligent gedrag te kunnen vertonen. De omstandigheden in de wereld bepalen wat de Creature doet, niet een centraal besturingssysteem.

Het bouwen van zo'n creature kan volgens de schrijver alleen in de echte wereld en zodat deze niet van een versimpelde omstandigheid af kan hangen. Per laag moet er uitvoerig getest worden of deze werkt en daarna pas kan de volgende laag erop gebouwd worden. Hierdoor kunnen zogenaamde 'bugs' opgespoord worden.

De Creatures zijn daadwerkelijk gebouwd op deze wijze en werkten naar behoren. Door middel van sonar sensoren en kleine 'computational units' in elke laag konden acties worden uitgevoerd en doelen worden behaald.

De schrijver wil ook duidelijk maken wat zijn creature niet is, namelijk:

- Connectionism, hierbij wordt geprobeerd om een netwerk te maken van meerdere processoren en deze te laten samenwerken. Tot dan toe was dat echter nog niet in 'real life' gelukte en de creatures bestaan enkel uit finite state machines.
- Een neurale netwerk, de Creature werkt niet hetzelfde als een neurale netwerk door een tekort aan onderlinge verbindingen van finite state machines.

- Productie regels, bepaalde systemen konden een actie uitvoeren door omstandigheden te matchen met bepaalde variabelen en aan de hand hiervan een actie uitvoeren. Ook dit is de creature niet want deze handelt niet indirect maar direct door de omstandigheden waarin hij zich vindt.
- Een 'blackboard' systeem, de units in de creatures geven alleen de juiste informatie aan elkaar door en beschikken niet allemaal over alle informatie.
- De Creature is niet gebaseerd op de filosofie van de Duitse filosoof Heidegger.

De schrijver beweert dat zijn creature een van de meest vooruitstrevende is en dat deze al de intelligentie van een insect heeft bereikt. Hij stelt ook kritische vragen over de aanpak die hij gebruikt heeft, namelijk:

- Hoeveel meer lagen kunnen er in een creature gebouwd worden?
- Hoe complex kan het gedrag van een creature worden zonder dat deze gebruik maakt van centrale representatie?
- Kunnen processen als leren ook bereikt worden met vaste verbindingen van finite state machines?

Hierop geeft hij de volgende antwoorden:

- Tot nu toe in het echt 3 en virtueel 6, het telkens bijvoegen van 1 laag schijnt tot nu toe goed te werken;
- De nieuwe creature waar de schrijver aan werkt is iets meer complex, deze kan lege blikjes opzoeken op een tafel en deze op de juiste plek deponeren;
- Het is mogelijk om dingen te leren door een afzonderlijke leerunit in te bouwen, er kan dan echter alleen geleerd worden over 1 specifiek ding en maar tot een bepaalde hoeveelheid.