

Home Correspondenten 'Robotica en AI'
(/plus) (<http://www.standaard.be/plus/tag/robotica-ai>)

DOORBRAAK IN 'MACHINE LEARNING' BIJ GOOGLE-FILIAAL DEEPMIND

Computer leerde zichzelf Atari 2600-spelletjes

26 FEBRUARI 2015 | Dominique Deckmyn (<http://www.standaard.be/auteur/dominique-deckmyn>)

Een computer van DeepMind heeft zichzelf 49 spelletjes op de klassieke Atari 2600-console aangeleerd, en speelt ze nu even goed of beter dan een mens. Kinderspel? Allerm minst. Misschien is het zelfs een doorbraak in artificiële intelligentie.



Onderzoekers van DeepMind, een filiaal van Google, publiceren deze week een opmerkelijk artikel in het blad *Nature* over hun computerprogramma DQN. Dat

programma draait op een simpele pc die wordt ‘gevoed’ met de beelden van een oude Atari 2600-spelconsole, waarvan hij de joystick (virtueel) bedient. Zo ‘leerde’ DQN 49 klassieke games spelen. In ongeveer de helft daarvan deed hij het beter dan een professionele games-tester (ja, zulke mensen bestaan). DQN had daar wel, per game, enkele dagen tot enkele weken non-stop training voor nodig.

Wat is hier zo bijzonder aan? Speelde de computer Deep Blue van IBM niet al in 1996 wereldkampioen schaken Gary Kasparov naar huis? Ja, maar Deep Blue was daarvoor – en alleen daarvoor – geprogrammeerd. Bovendien is schaak een spel dat een computer heel goed ligt: het bord kan gemakkelijk gereduceerd worden tot een model in zijn geheugen en het spelverloop is perfect voorspelbaar, als je maar genoeg rekenkracht in huis hebt om voldoende zetten vooruit de ‘denken’.

Een computerspel als *Space invaders* daarentegen, is heel wat chaotischer. De speler moet signalen verwerken waar veel ‘ruis’ op zit: het beeldscherm van een spelconsole waarop allerlei flikkerende figuurtjes en lijnen bewegen. Vervolgens moet hij leren welke handeling (het raketje van Space invaders naar links of naar rechts bewegen, een schot afvuren) het beste resultaat opleveren in elke situatie. En DQN pakte een aantal heel verschillende games aan, met naast *Space invaders* bijvoorbeeld ook een race-spel (*Enduro*) en een virtuele boksmatch (*Boxing*).

Oude en nieuwe trucs

Opmerkelijk is ook dat de mensen van DeepMind dit klaarspeelden door twee bestaande technieken op een slimme manier aan elkaar te koppelen: een neuraal netwerk en *reinforcement learning*. Die eerste component, het neuraal netwerk, wordt al decennia gebruikt. Het is een computersysteem dat de werking van menselijke neuronen imiteert. Het wordt meestal ingezet om patronen te herkennen, bijvoorbeeld om te zien of een verzameling stipjes op een blad papier de letter ‘n’ of de letter ‘h’ voorstelt. Of om te herkennen of je het woord ‘poot’ of ‘boot’ hebt uitgesproken. Zo’n neuraal netwerk kun je iets aanleren door het heel geduldig een eindeloze reeks voorgekauwde voorbeelden aan te reiken van letters ‘n’ en ‘h’.

DQN werkt met een andere benadering, *reinforcement learning* genaamd. Daarbij wordt een programma ‘beloond’ voor een goede handeling, zoals je een hond kunstjes leert. In dit geval wordt het programma beloond met een hoge score in het spel. Geleidelijk – na het equivalent van enkele weken continu spelen – leert het programma zo de beste manier om te reageren op verschillende spelsituaties. Opmerkelijk is dat de computer daarbij ook vooruit probeert te kijken: DQN maakt op elk moment een schatting van welke handeling – naar links bewegen, bijvoorbeeld – hem uiteindelijk tot de beste eindscore zal leiden.

De makers van DQN voegen aan die slimme mix van bestaande technieken ook een nieuw idee toe: DQN zal geregeld het spelverloop van eerder gespeelde games opnieuw bekijken. Dat blijkt het leervermogen sterk te verbeteren. En het is, volgens DeepMind, gebaseerd op de manier waarop de hippocampus in onze eigen hersenen eerdere ervaringen opnieuw afspeelt. DQN is dus geen precieze imitatie van onze hersenen, maar haalt er wel inspiratie.

Heilige graal

‘Het spelen van Atari-spelletjes is maar een demonstratie, deze aanpak is aanpasbaar aan vele verschillende taken’, zei Koray Kavukcuoglu van DeepMind dinsdag tijdens een telefonische persconferentie. Wat dan zoal? ‘We kijken nu games uit de jaren ’90, 3D-games, racing games, waarvan de moeilijkheidsgraad veel hoger is. Je moet navigeren in drie dimensies, wat betekent dat je meer complexe zintuiglijke informatie moet verwerken en dat je je een voorstelling moet maken van een virtuele ruimte.’

Een topspel van vandaag, zoals *Grand theft auto 5*, zou haalbaar kunnen zijn over een jaar of vijf, zeggen de mensen van DeepMind. En als je in een complexe simulatie als *Grand theft auto* met een wagen kunt rondrijden, kan DQN dan misschien ooit zelf een auto leren besturen? Volgens DeepMind wel. ‘Dit kan worden toegepast op elk sequentieel beslissingsprobleem’.

Dé heilige graal van de artificiële intelligentie is wat men ‘algemene’ of ‘sterke’ artificiële intelligentie noemt. Een computersysteem dat geen ‘vakidoot’ is, maar intelligentie kan vertonen op vele verschillende domeinen. Dat kan eigenlijk alleen maar als de computer zelf kan leren uit wat hij observeert. Zoals DQN al een beetje doet. En volgens Demis Hassabis, één van de oprichters van DeepMind, is dat ook het einddoel van het bedrijf. ‘Dit is de eerste trede op een ladder die ons zal brengen naar een algemeen leersysteem’, zegt hij. Volgens Hassabis ligt zo’n systeem, dat het leervermogen van de mens evenaart, wel nog ‘vele decennia’ in onze toekomst.

Uiteraard zijn de mensen van DeepMind niet de enigen die daaraan werken. In januari demonstreerden onderzoekers van de Universiteit van Maryland nog een robot die kookfilmpjes op YouTube bekijkt en daaruit kan leren hoe hij bepaalde keukeninstrumenten moet gebruiken. Dat onderzoek wordt gefinancierd door Darpa, de Amerikaanse overheidsorganisatie die indertijd het internet ontwikkelde.

Een computer slaagt er nog lang niet in om, zoals een mens, alle informatie uit diverse zintuigen te verwerken en daaruit een abstracte voorstelling van de wereld te maken. Zelfs binnen zijn beperkte domein heeft DQN vandaag nog massa’s gebreken. Games waarin tactisch wat verder vooruit moet worden gedacht, gaan hem nog niet zo goed af.

En wat het programma niét kan, is zijn ervaringen uit één spel toepassen op een ander, gelijkaardig spel. Die vaardigheid hopen zijn makers er later wel aan toe te voegen.

Dominique Deckmyn

Correspondent

Robotica en artificiële intelligentie



DOMINIQUE DECKMYN

Als correspondent 'Robotica en artificiële intelligentie' onderzoekt hij samen met Nico Tanghe de voordelen en de risico's van de vooruitgang.

Meer artikels van Dominique Deckmyn 

(<http://www.standaard.be/auteur/dominique-deckmyn>)

Een topspel van vandaag zoals *Grand theft auto 5* zou haalbaar kunnen zijn over een jaar of vijf