

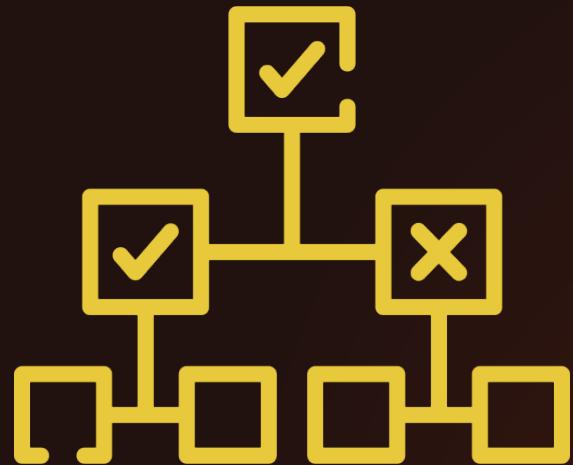
# LEARNING TO BEHAVE

REINFORCEMENT LEARNING  
IN HUMAN CONTEXTS

Floris den Hengst



AI = intelligente machines



serie van keuzes



intelligentie door te leren



taak-afhankelijke scores

A close-up portrait of Douglas Adams, the author. He has short, light-colored hair and is looking slightly to his left with a thoughtful expression. The lighting is soft, highlighting his eyes and forehead.

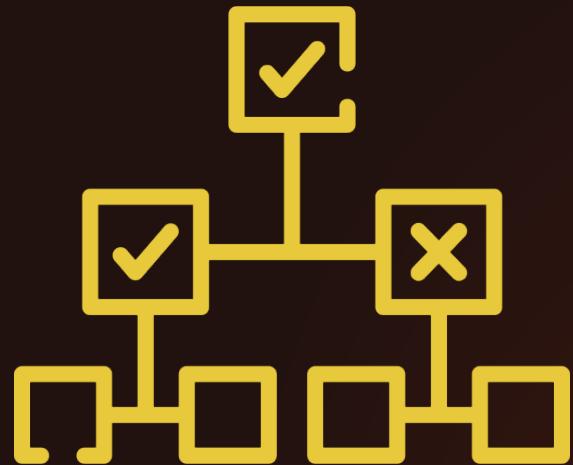
“We demand  
rigidly defined areas  
of doubt  
and uncertainty!”

- Douglas Adams -

schrijver van "A Hitchhiker's guide to the galaxy"



AI = intelligente machines



serie van keuzes



intelligentie door te leren



taak-afhankelijke scores

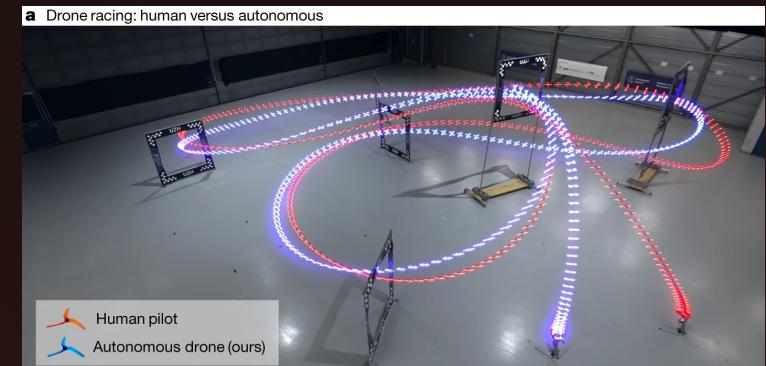
# Successen van reinforcement learning



wereldkampioen Go  
2017



ontwerp van chip in 6 uur  
2021



snelste rondje met een drone  
2023

The background of the slide features a small white humanoid robot sitting on a keyboard, with numerous colorful plastic building blocks scattered around it. The blocks are primarily red, blue, green, and yellow. The robot has a simple design with large blue eyes and a small mouth. The keyboard keys are visible in the foreground, with some having musical symbols like asterisks and dots. The overall theme is playful and educational.

# DEEL I

## TOEPASSINGEN VAN REINFORCEMENT LEARNING IN MENSELIJKE OMGEVINGEN

# Reinforcement learning voor personalisatie

overzicht van problemen & oplossingen

aantal publicaties neemt toe over tijd

toepassingsgebieden & algoritmen in kaart

I STARTED THE DAY WITH  
LOTS OF PROBLEMS.  
BUT NOW, AFTER HOURS  
AND HOURS OF WORK,  
I HAVE LOTS OF PROBLEMS  
IN A SPREADSHEET.



# Personalisatie van praatprogramma's

personalisatie van aanbevelingen

twee varianten op basis van reinforcement learning

beter dan bestaande oplossingen

beter bestand tegen onzekerheid





"This call is monitored for quality. Try to  
keep up your end of the conversation."

Progress: 0%

1 out of 30

**TASK 1:** If the chatbot has a breakdown, select that first message

**Breakdown:** When a response does not match the goal of the person, and the chatbot does not overcome this mistake.

Find a restaurant with Welsh food please.

Person

I am sorry, but there is no restaurant with Welsh food.

Chatbot

Gastro pub?

Person

- The conversation breaks down from here -

I am sorry, but there is no restaurant with Welsh food.

Chatbot

Italian?

Person

END OF THE CONVERSATION

**TASK 2:** Rate the user satisfaction of the conversation with the chatbot

**User Satisfaction:** How the person experienced their goals or desires being accomplished by the chatbot.

Very dissatisfied

Dissatisfied

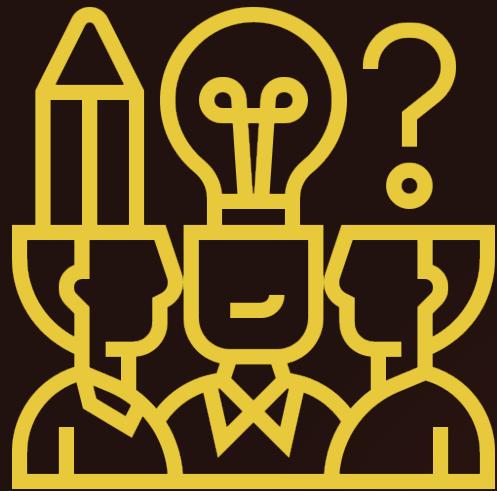
Unsure

Satisfied

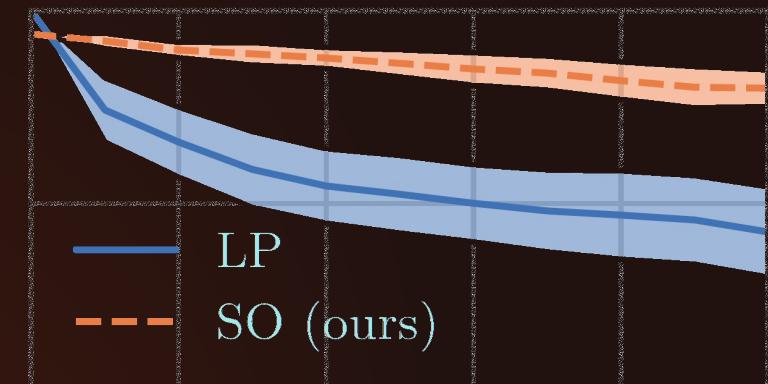
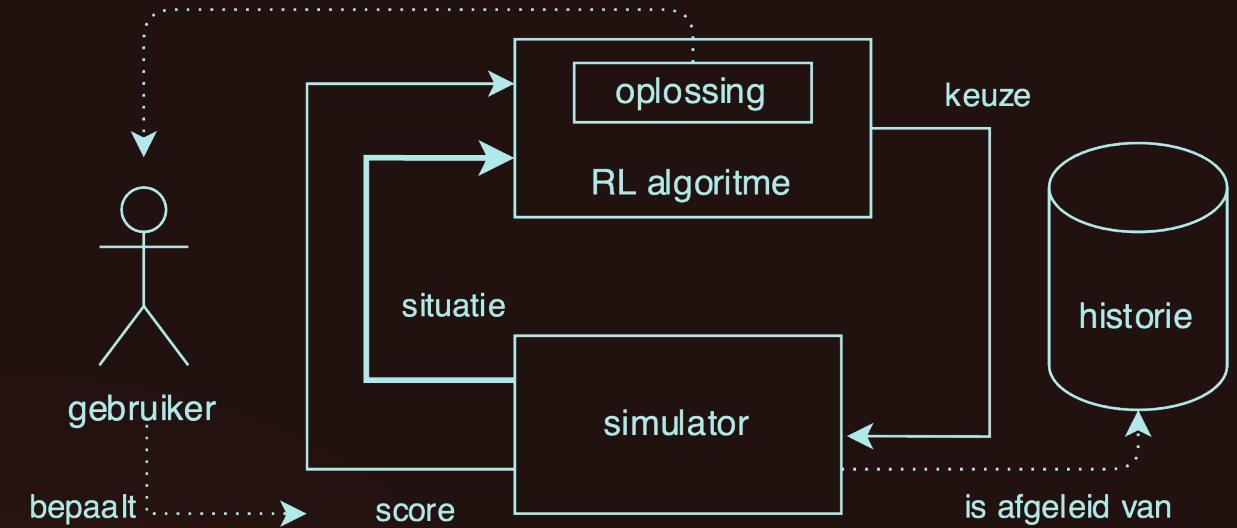
Very satisfied

SAVE & GO TO NEXT PAGE

# Een optimaal personeelsbestand



de mensen  
op de juiste plek  
op het moment





# DEEL II

## SUBSYMBOLISCHE REINFORCEMENT LEARNING & SYMBOLISCHE KENNIS

Go



# Universum



meer dan  $10^{176}$  situaties

$10^{82}$  atomen

1

voor elk atoom meer dan

# Reinforcement learning zonder symbolen

- ✓ makkelijk te trainen
- ✓ complexe problemen
- ✓ bewezen in de praktijk
- ✗ veiligheidsgaranties
- ✗ benodigd aantal ervaringen
- ✗ directe manipulatie

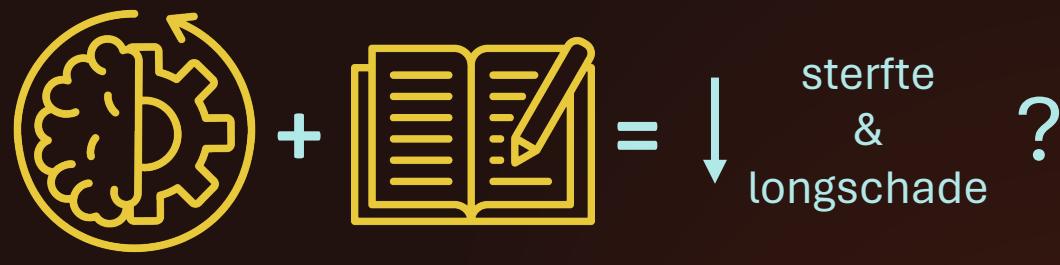
# Reinforcement learning met symbolen?

- ✓ makkelijk te trainen
- ✓ complexe problemen
- ✓ bewezen in de praktijk
- ✓ veiligheidsgaranties
- ✓ benodigd aantal ervaringen
- ✓ directe manipulatie

# Optimaal beademmen op de intensive care



gevarieerdere instellingen



leren

symbolische  
medische  
richtlijn

100% volgens richtlijn  
~55% bij arts

mogelijk minder sterfte

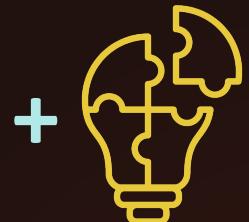
# Een veilig praatprogramma



leren



symbolische  
financiële  
richtlijn



expliciet  
redeneren



---

score

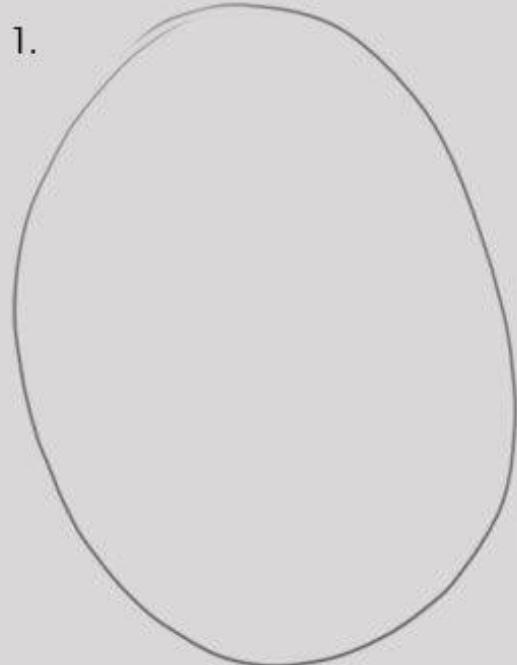


veiligheid



## Hoe je een hoofd tekent

1.



2.



1. teken een ovaal

2. teken nu de rest van het hoofd

# Leren met symbolische instructies



leren



symbolische  
instructies

sneller leren



hergebruik van deeloplossingen

# LEARNING TO BEHAVE

REINFORCEMENT LEARNING  
IN HUMAN CONTEXTS

Floris den Hengst