

# LEARNING TO BEHAVE

REINFORCEMENT LEARNING  
IN HUMAN CONTEXTS

Floris den Hengst

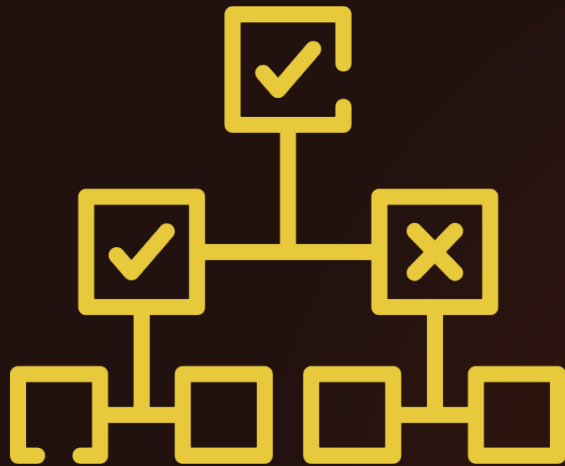




AI = intelligente machines




intelligentie door te leren



serie van keuzes



taak-afhankelijke scores



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

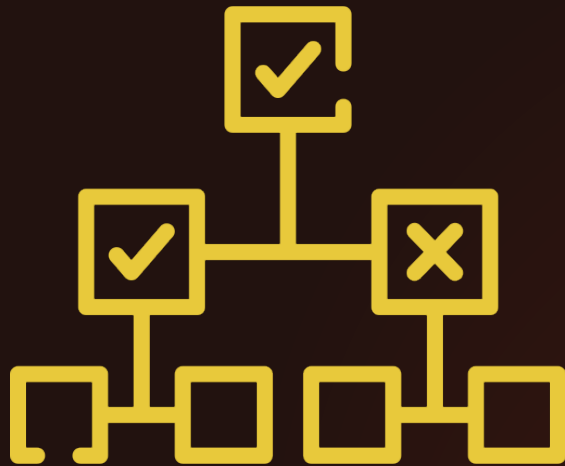
- Douglas Adams -  
schrijver van "A Hitchhiker's guide to the galaxy"



AI = intelligente machines



intelligentie door te leren



serie van keuzes

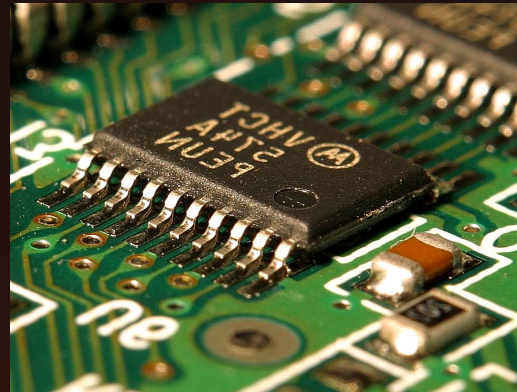


taak-afhankelijke scores

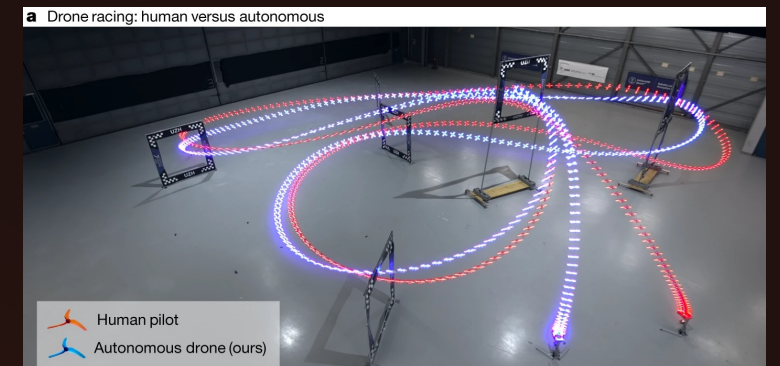
# Successen van reinforcement learning



wereldkampioen Go  
2017



ontwerp van chip in 6 uur  
2021



snelste rondje met een drone  
2023

# 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



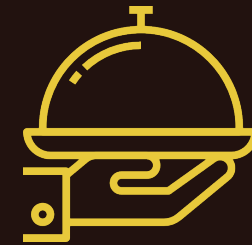
# Personalisatie van praatprogramma's

personalisatie van aanbevelingen

twee varianten op basis van reinforcement learning

beter dan bestaande oplossingen

beter bestand tegen onzekerheid







WILDT

"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

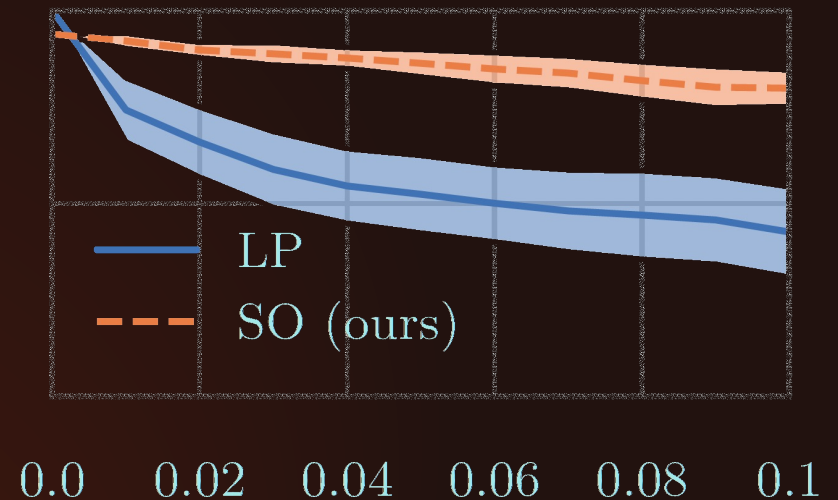
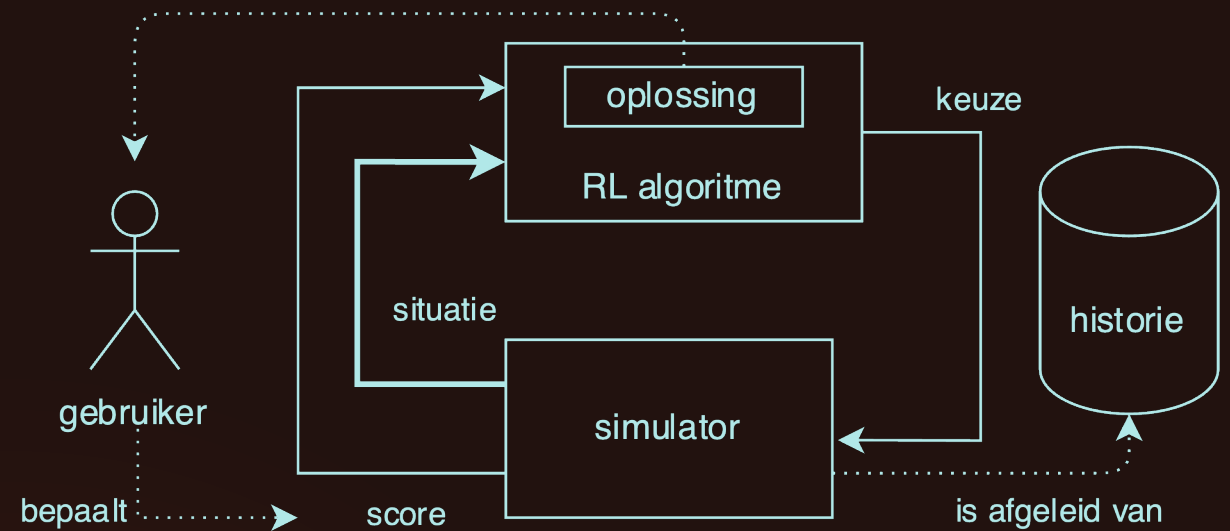
Very satisfied

SAVE & GO TO NEXT PAGE

# Een optimaal personeelsbestand



de op de op het **juiste** mensen plek moment



# DEEL II

## SUBSYMBOLISCHE REINFORCEMENT LEARNING & SYMBOLISCHE KENNIS





# 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

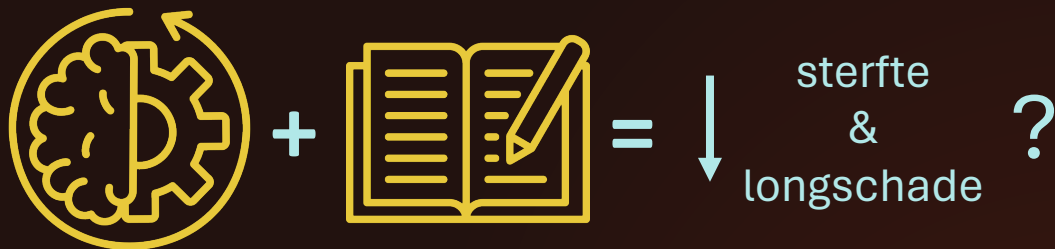
# Optimaal beademen op de intensive care



gevarieerdere instellingen

100% volgens richtlijn

~55% bij arts



leren

symbolische  
medische  
richtlijn

sterfte  
&  
longschade

?

mogelijk minder sterfte



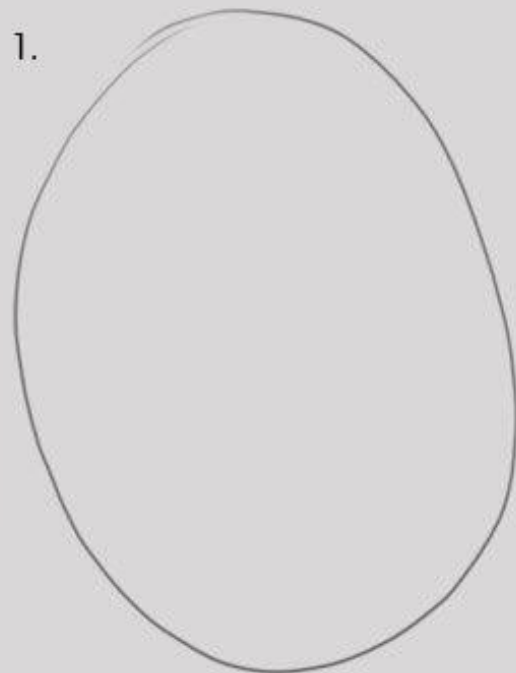
# Een veilig praatprogramma



	score	veiligheid
	★ ★ ★	★ ☆ ☆
	☆ ☆ ☆	★ ★ ★
	★ ★ ★	★ ★ ★

## Hoe je een hoofd tekent

1.



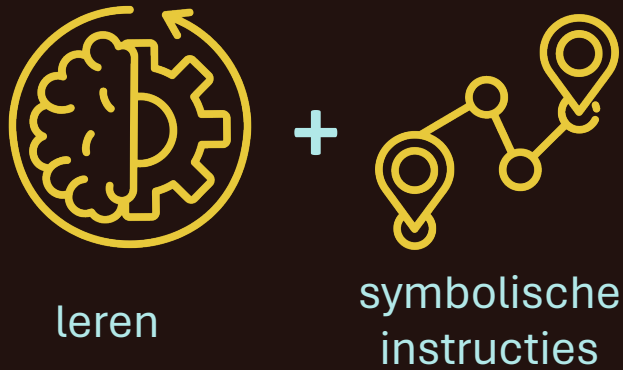
1. teken een ovaal

2.



2. teken nu de rest van het hoofd

# Leren met symbolische instructies



sneller leren



hergebruik van deeloplossingen

# LEARNING TO BEHAVE

REINFORCEMENT LEARNING  
IN HUMAN CONTEXTS

Floris den Hengst

