

# Flanders AI Research Program webinar

# Reinforcement Learning: from the lab to the real world

December 3<sup>rd</sup>, 2020

Prof. Ann Nowé  
Dr. Denis Steckelmacher  
Dr. Leander Schietgat

 [ai.vub.ac.be](https://ai.vub.ac.be)

 @aibrussels

VUB **Artificial Intelligence** Lab



# Outline

- Introduction to the **Flanders AI Research Program** (Sabine Demey)
- The **VUB AI Lab** and its research in the program (Leander Schietgat)
- Introduction to **Reinforcement Learning** (Ann Nowé)
- Reinforcement Learning **in Practice** (Denis Steckelmacher)
- **Take-home messages** (Leander Schietgat)
- Q&A

# In which (industry) sector are you active?

Manufacturing

Energy

Logistics

Telecom

Health

Finance

Government

Consulting

ICT

Recruitment

Robotics

AI (technology provider)

Academia

Other

# Did you (or your company) already implement or integrate an AI solution in your business?

Yes

No, but considering  
it on the short term

No idea

No



# Vlaams AI Onderzoeksprogramma Flanders AI Research Program

Sabine Demey, Director

December 2020

.AGORIA

Vlaams  
netwerk van  
ondernemingen

sirris  
shaping industry by technology

imec  
embracing a better life

FLANDERS  
MAKE

VIB

vito

VUB  
VRIJE  
UNIVERSITEIT  
BRUSSEL

KU LEUVEN

Universiteit  
Antwerpen

UNIVERSITEIT  
GENT

UHASSELT

AGENTSCHAP  
INNOVEREN &  
ONDERNEMEN

DEPARTEMENT  
ECONOMIE  
WETENSCHAP &  
INNOVATIE

Vlaamse  
overheid

PUBLIC

# Flanders AI Program

Program Structure with 3 pillars, funded by the Flemish Government

1 FLANDERS AI RESEARCH PROGRAM

Start: July, 1<sup>st</sup> 2019

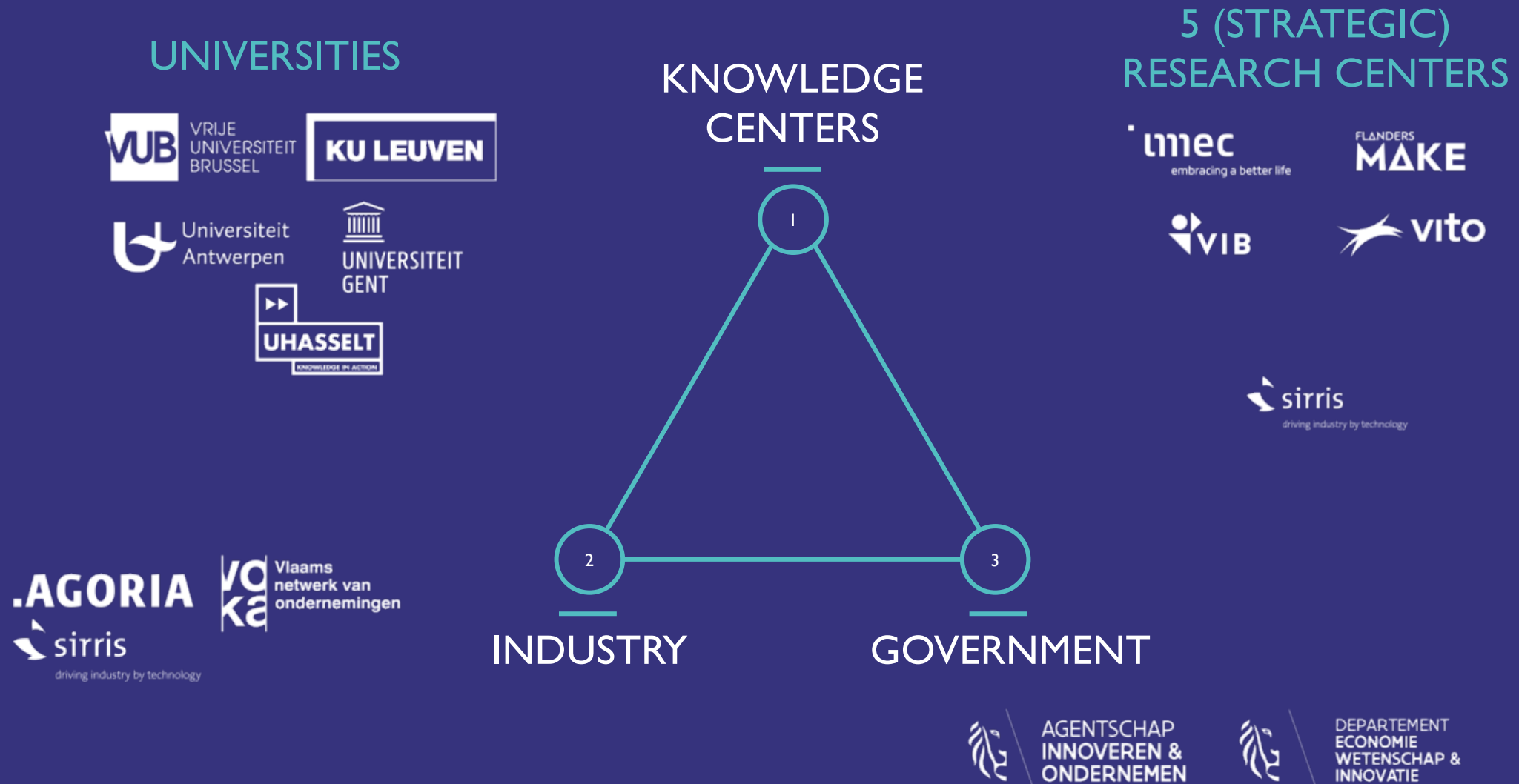
2 FLANDERS AI IMPLEMENTATION PROGRAM

3 FLANDERS AI SUPPORTING ACTIVITIES:  
ETHICS, EDUCATION AND TRAINING

 Knowledge Centre  
Data & Society

Flanders AI Academy

# The 'triple helix-model'



# Challenge-Based Research with Demand-Driven Impact



CHALLENGE  
BASED  
RESEARCH

AI-DRIVEN DATA  
SCIENCE



AI IN THE EDGE



MULTI-AGENT  
COLLABORATIVE AI



HUMAN-LIKE AI



PROOFS-OF-CONCEPTS (Demonstrators)

FOCUS DOMAINS and USE CASES



WITH  
DEMAND-DRIVEN IMPACT



# Flanders AI Research Program

## 4 Research Challenges

AI-DRIVEN DATA SCIENCE



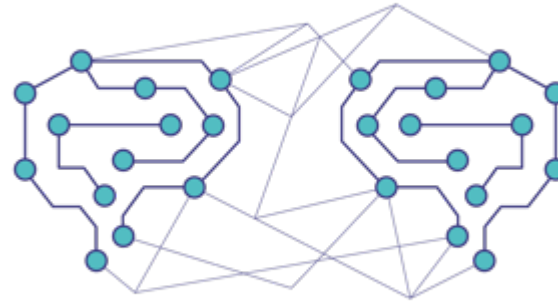
*Making Data Science  
Hybrid, Automated,  
Trusted and Actionable*

AI IN THE EDGE



*Real-Time and  
Power-Efficient AI in  
the Edge*

MULTI-AGENT  
COLLABORATIVE AI



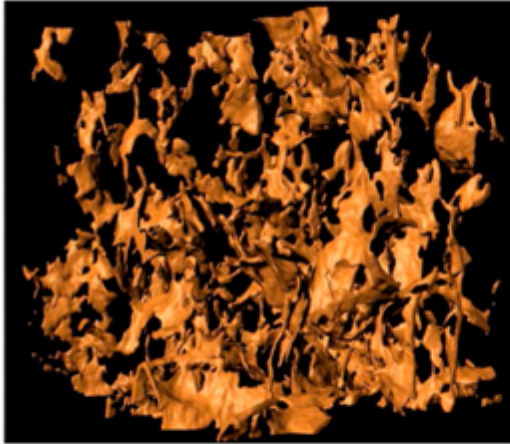
*Interact Autonomously  
with other Decision-  
Making Entities*

HUMAN-LIKE AI

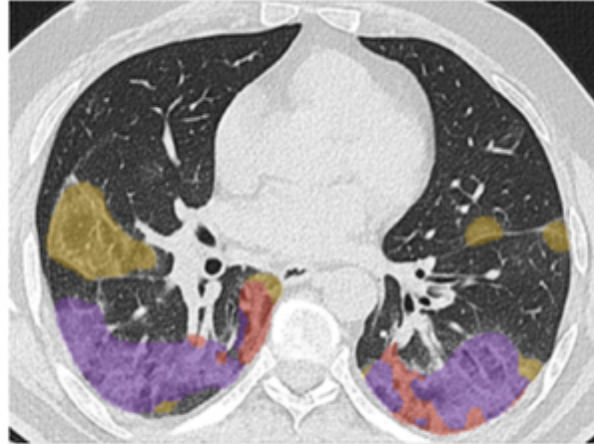


*Communicate and  
Collaborate Seamlessly  
with Humans*

# AI in health



Precision medicine



Clinical decision support



Monitoring and treatment

# Selected Applications in Health

PRECISION MEDICINE

CLINICAL DECISION SUPPORT

HOSPITAL  
DECISION  
SUPPORT

Single Cell  
Technologies

Multiple  
Sclerosis

Medical Imaging  
*Radiation  
Oncology and  
Radiology*

Epilepsy

Hospital  
Treatment  
Decisions

Personal Health  
Data  
Management

Visualisation of  
single cell data  
& Segmentation  
of 3D electron  
microscopy  
images

Improve MS  
treatment

Segmentation &  
classification for  
radiation  
oncology and  
radiology

Automated  
detection of  
epileptic  
seizures

Prediction of  
length of stay in  
hospitals

# AI in Industry



**MANUFACTURING**



**ENERGY**



**AGRICULTURE**

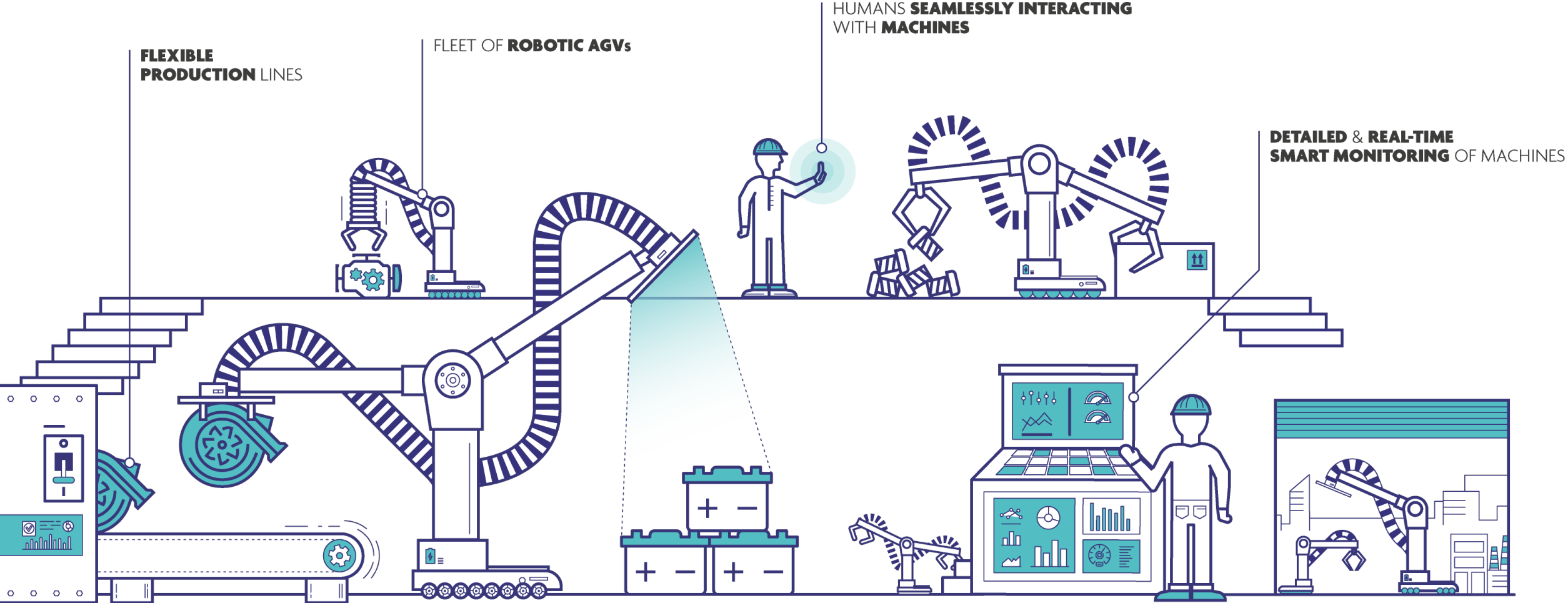


**LOGISTICS**



**RETAIL**

# A future manufacturing/logistics plant



# AI for Government & Citizens & in Smart Spaces

PUBLIC  
EMPLOYMENT  
SERVICES

PERSONAL DATA

## Challenges

- 1. Ethical & Trustworthy AI, Privacy, Bias & Fairness
- 2. Personalization
- 3. Intuitive interactions

CONVERSATIONAL  
AGENTS

RECOMMENDER  
SYSTEMS



# Flanders AI Research Program



Sabine Demey  
Program Director Flanders AI Research  
imec



Jo De Boeck  
Executive Representative imec

## AI-driven Data Science



Prof. Bart De Moor  
ESAT, KULeuven



Prof. Piet Demeester  
IDLab, Ghent University-imec

## Multi-agent Collaborative AI



Ann Nowé  
Professor AI Lab, VUB

## AI in the Edge



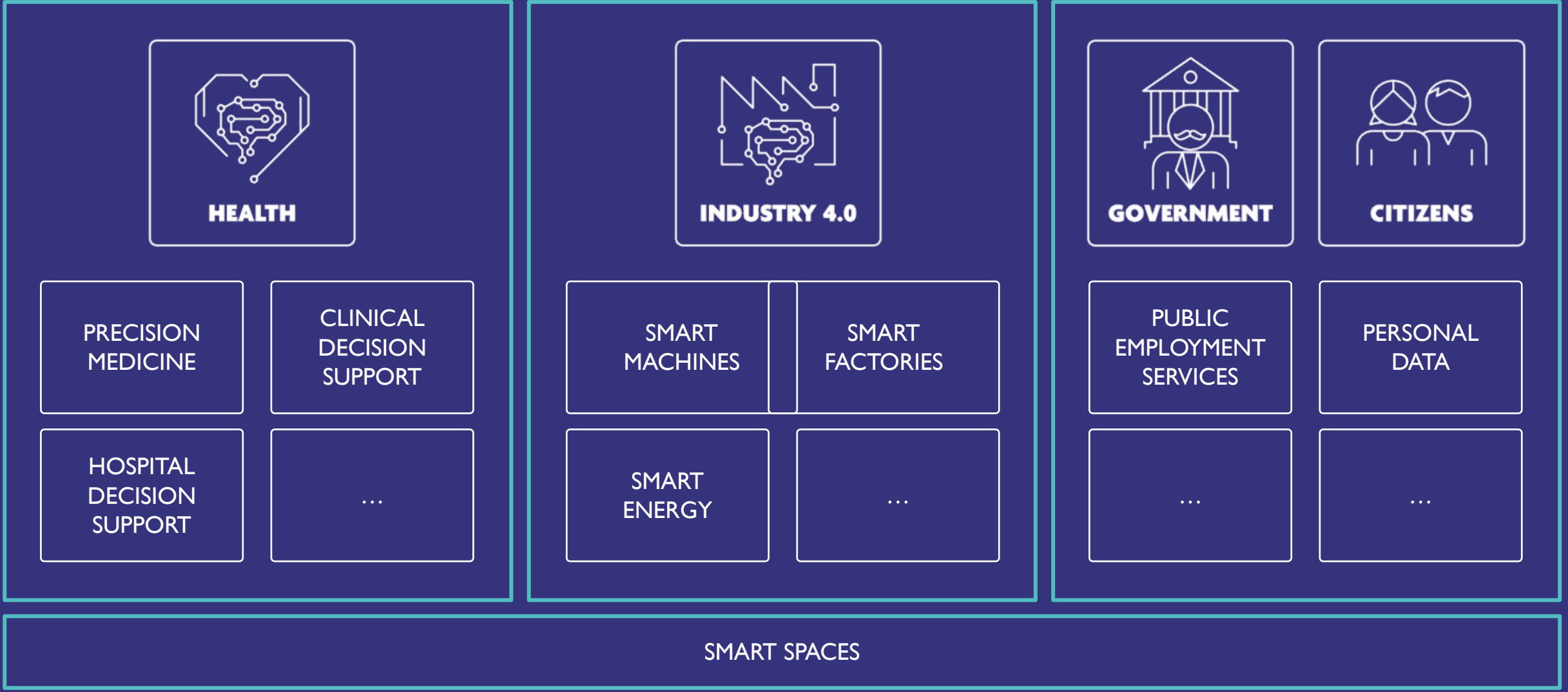
Mieke De Ketelaere  
Program Director AI, imec

## Human-Like AI



Prof. Steven Latré  
IDLab, University of Antwerp  
imec

# Domains and Applications





# Today's topic: Reinforcement Learning

## Relevant for multiple challenges

- AI systems are not standalone
  - they act in an **environment**
  - they **interact** with **humans**
- **Grand Challenge 3: Multi-agent collaborative AI** (led by VUB)
- **Grand Challenge 4: Human-like AI** (led by UAntwerpen)



*Agents in a smart grid model*

January 2021

Multi-agent systems webinar

# VUB Artificial Intelligence Lab

Strong heritage, leading position



**Founded in 1983**  
by Prof. Luc Steels

First AI Lab on the  
European mainland

37 years of  
**experience**



50 researchers  
from 22 countries  
**13 professors**



950 publications  
28 000 citations



**Current projects**  
9 EU projects  
20+ national projects  
15 industry funded projects



5 spin-offs  
former colleagues @ DeepMind,  
Prowler, MIT, CalTech, Collibra,  
Sony, ...

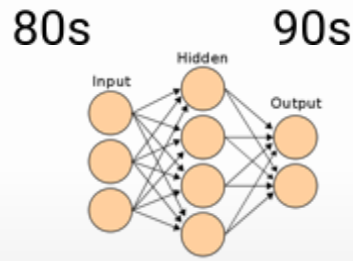
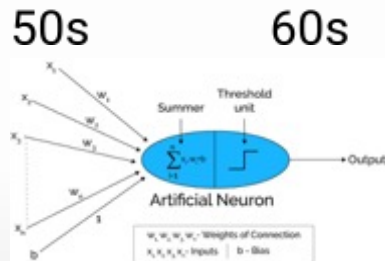
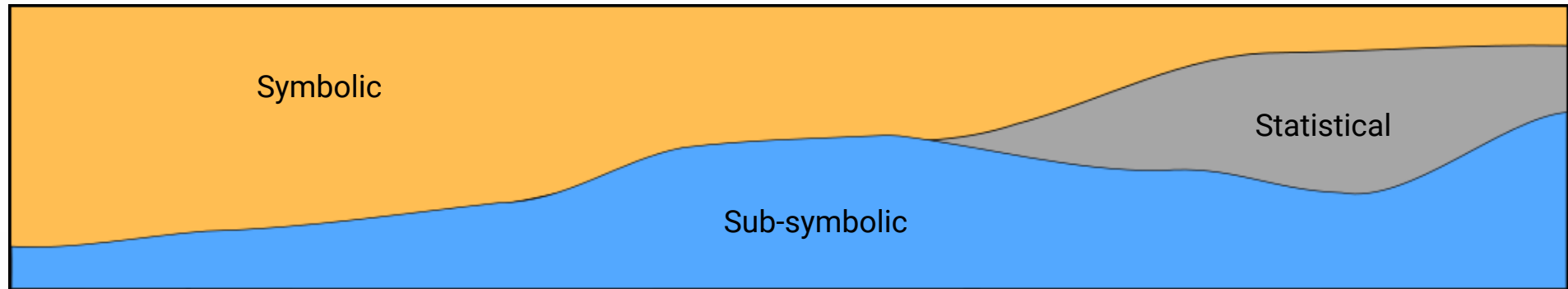
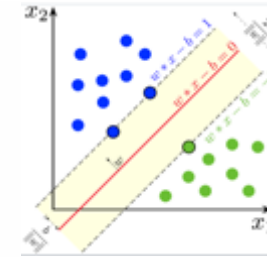
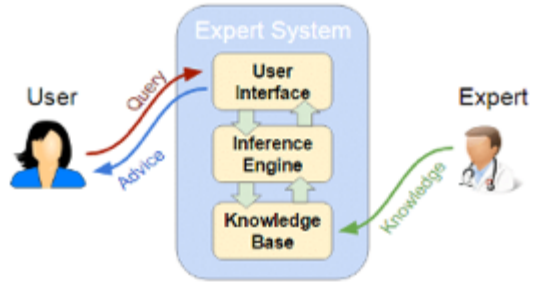
# Have you heard about the Flemish AI Impulse Program before?

Yes

No

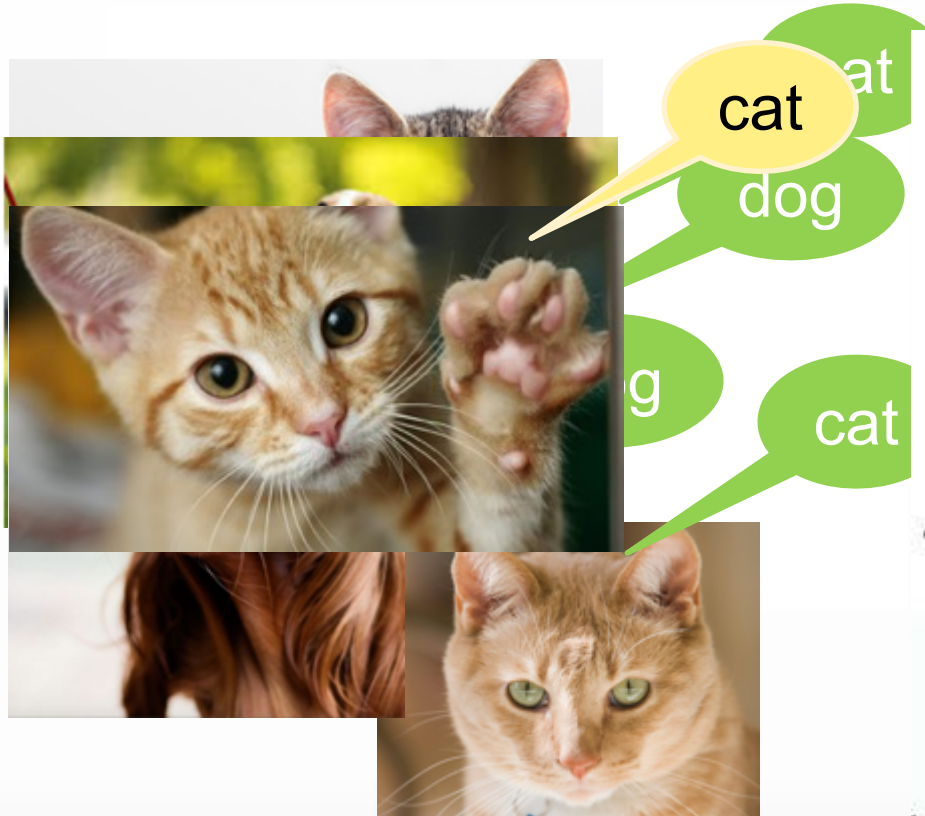
# Introduction to Reinforcement Learning

# Brief history of AI: the different movements



# Subfields of Machine Learning

Supervised



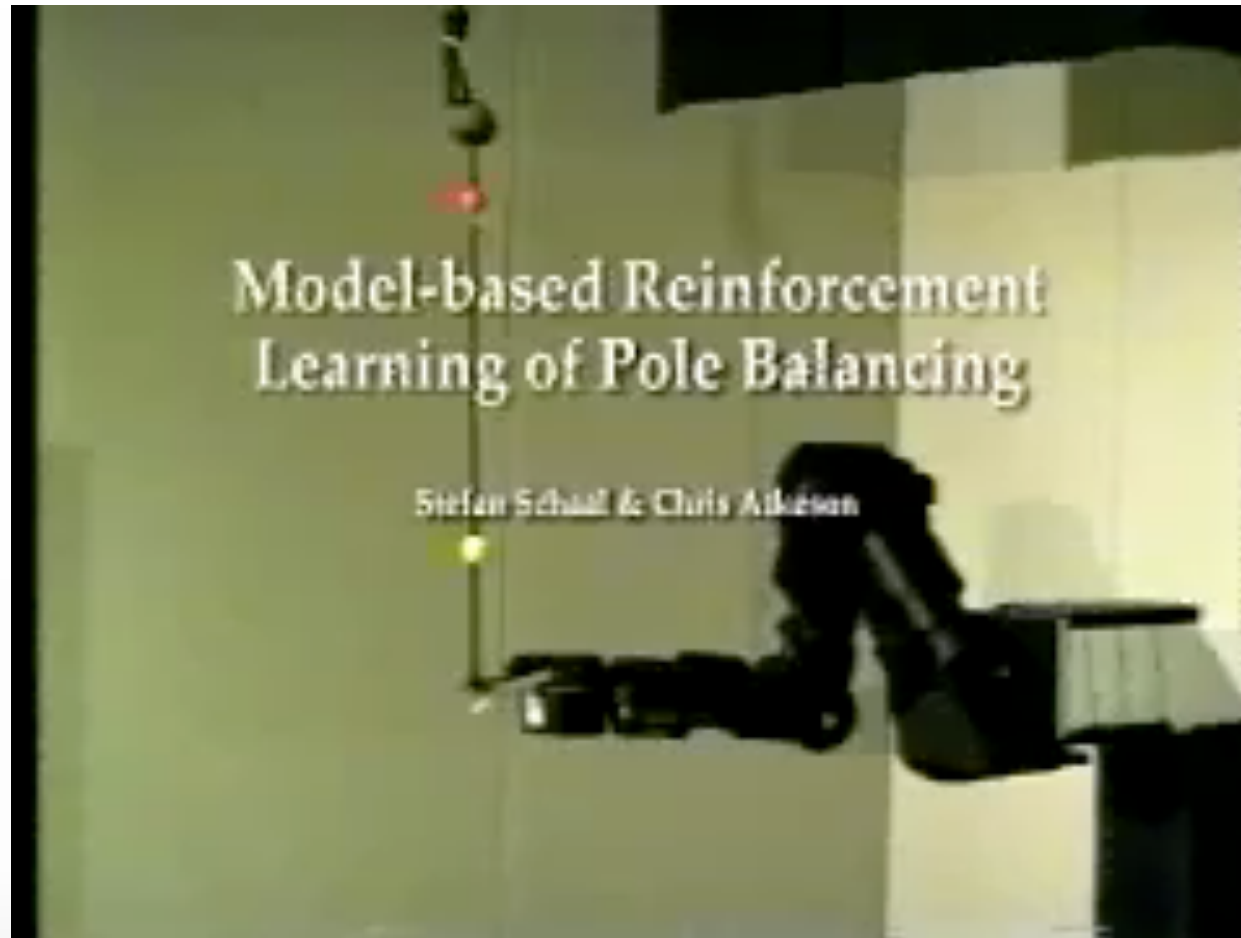
Reinforcement



Unsupervised



# Reinforcement learning



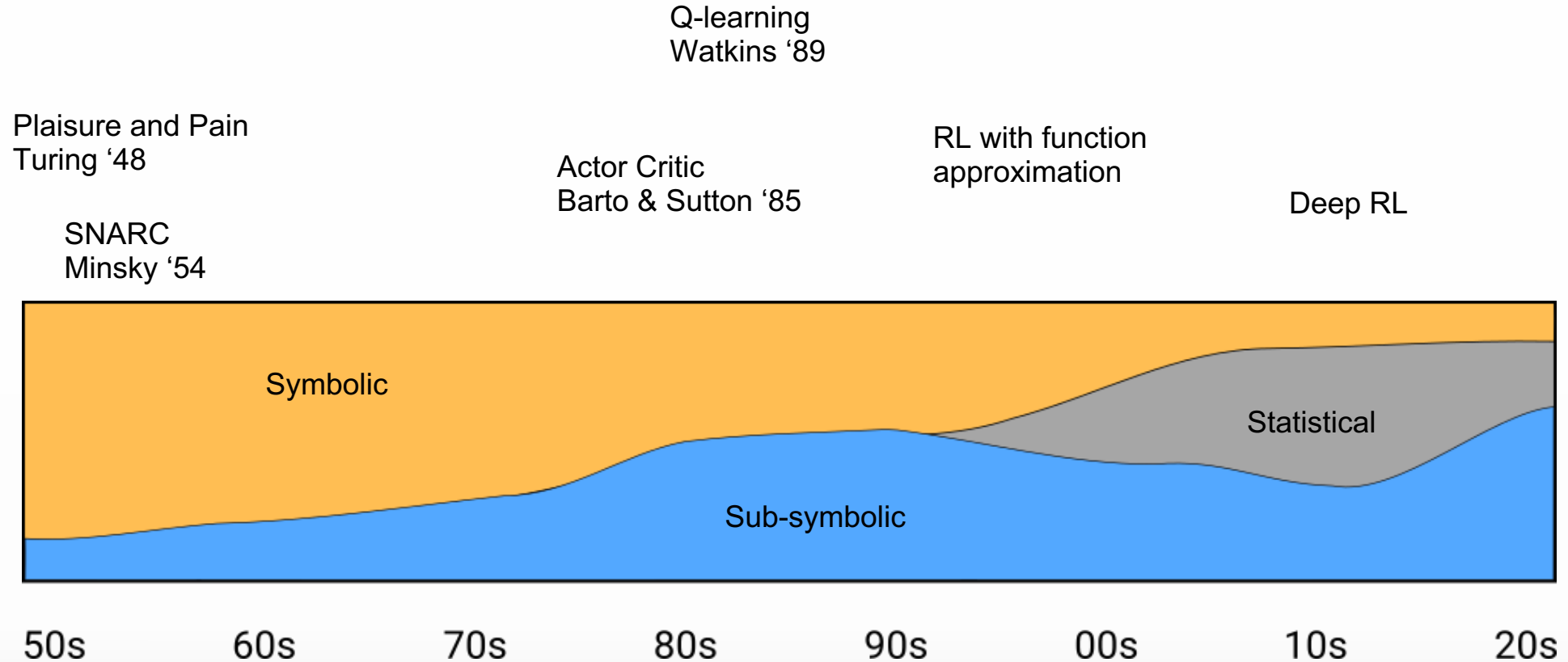
# In which scenarios is reinforcement learning the preferred choice?

Predict  
traffic jams

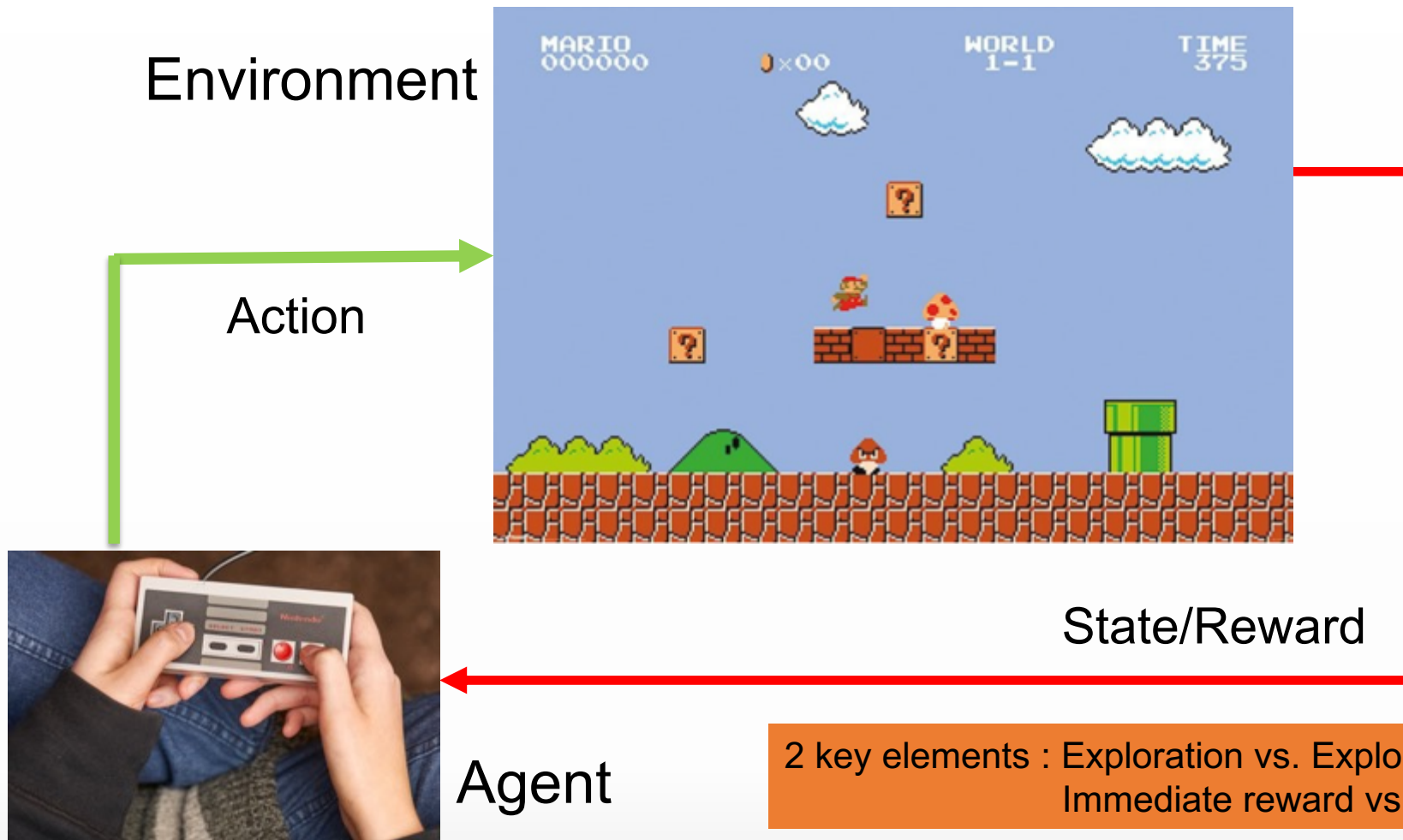
Prevent  
traffic jams



# Reinforcement Learning



# Learning from interactions



2 key elements : Exploration vs. Exploitation  
Immediate reward vs. Long term reward.

# Exploration versus exploitation: bandits

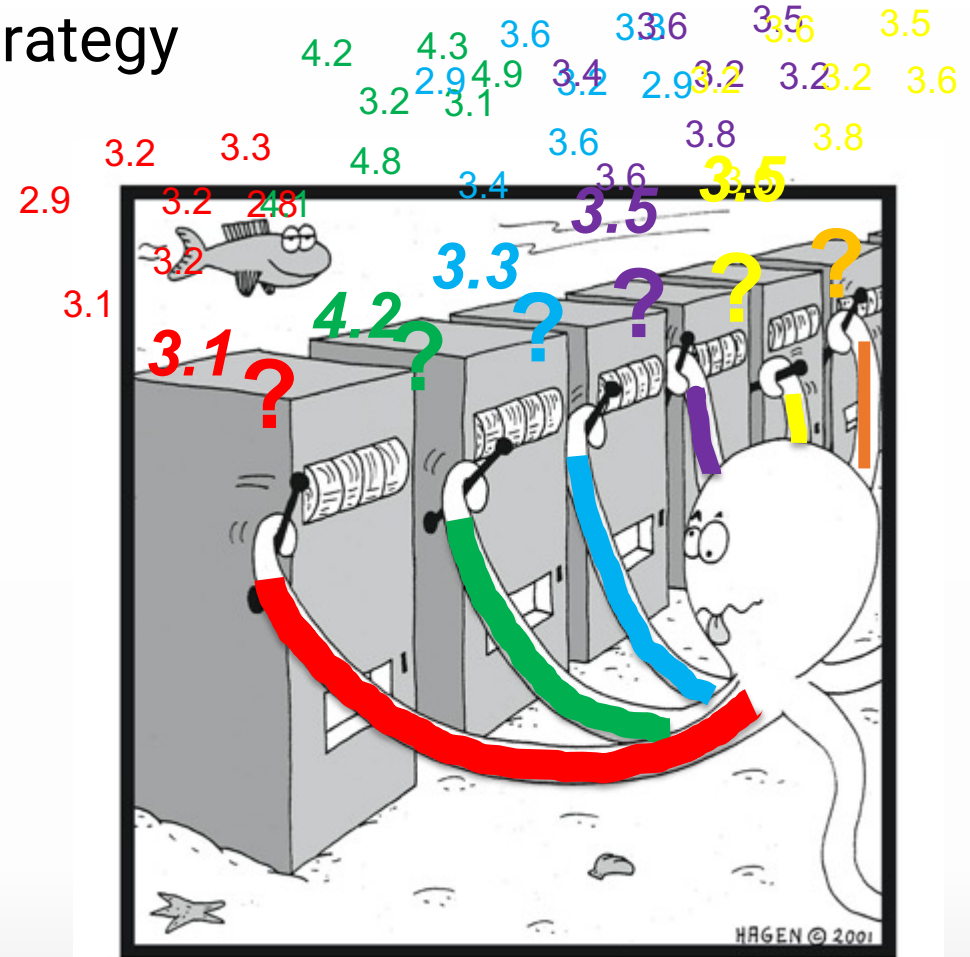
Finding the best epidemic mitigation strategy

Close school

Close restaurants

Close non-essential shops

Close shops & restaurants



# Exploration versus exploitation: bandits

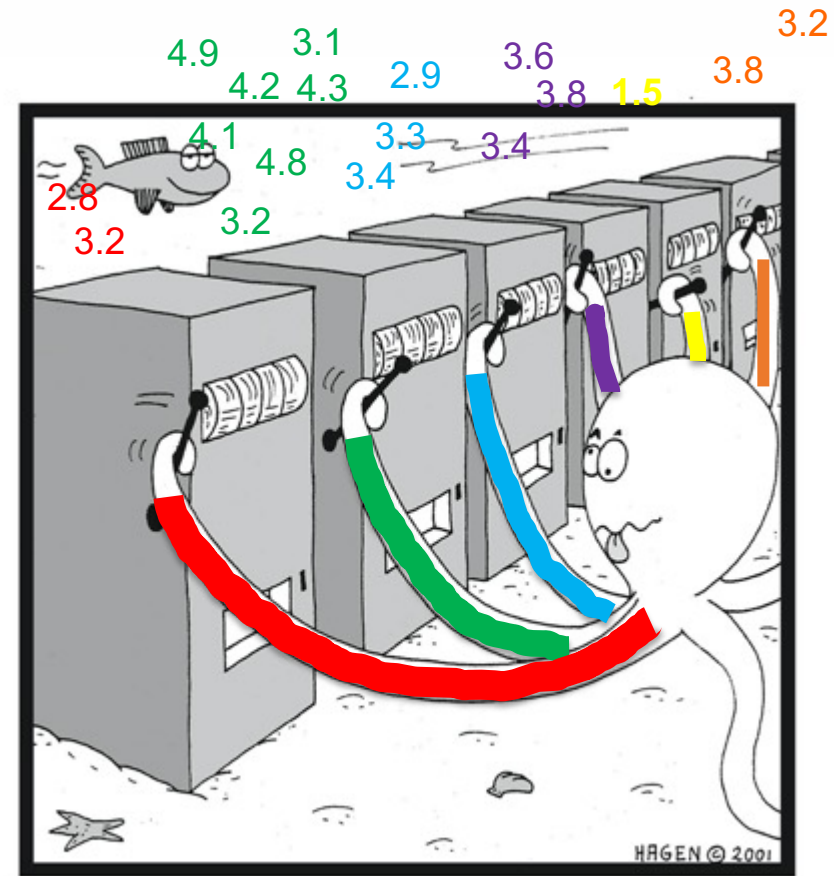
Finding the best epidemic mitigation strategy

**Close school**

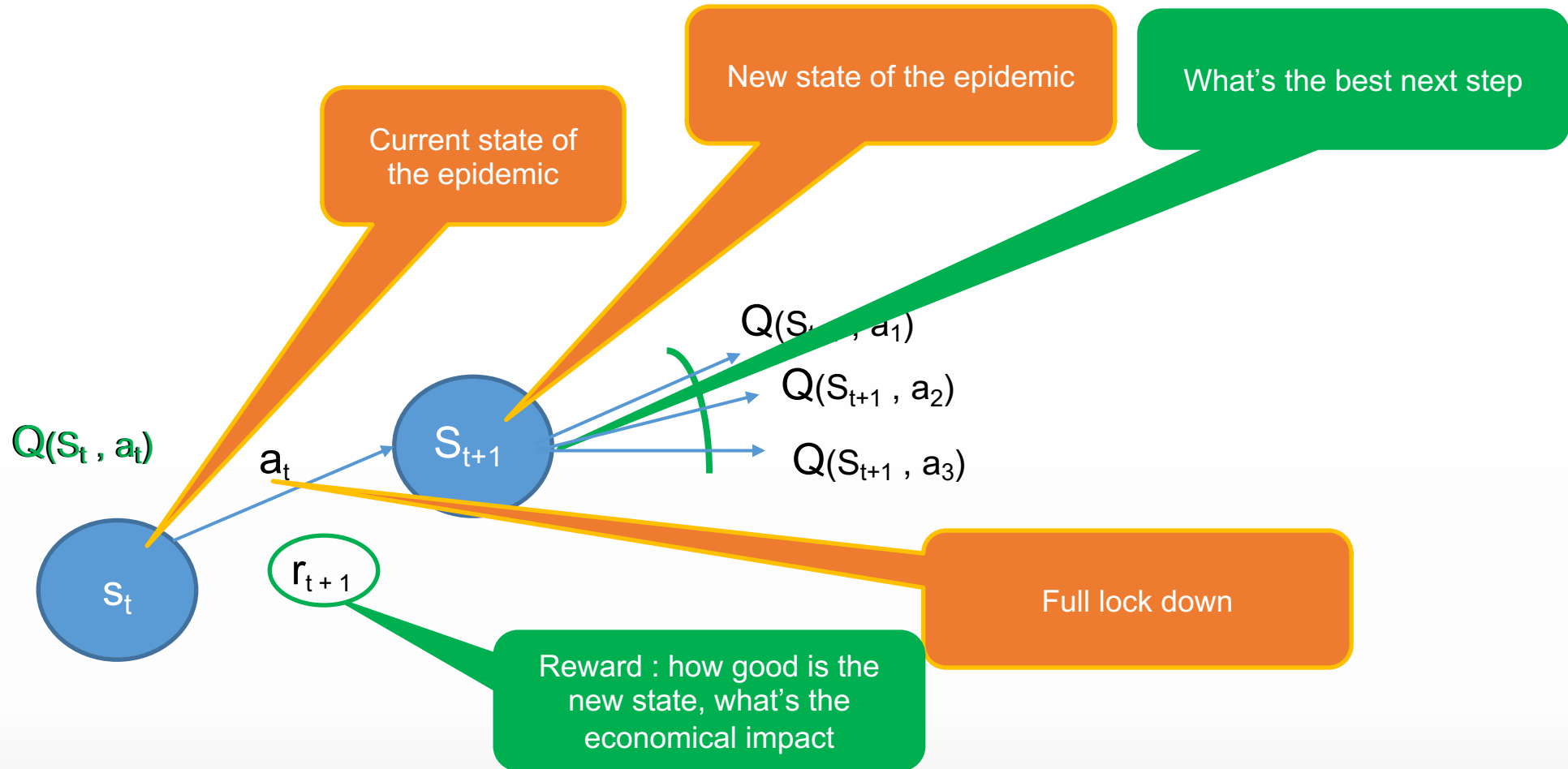
**Close restaurants**

**Close non-essential shops**

**Close shops & restaurants**

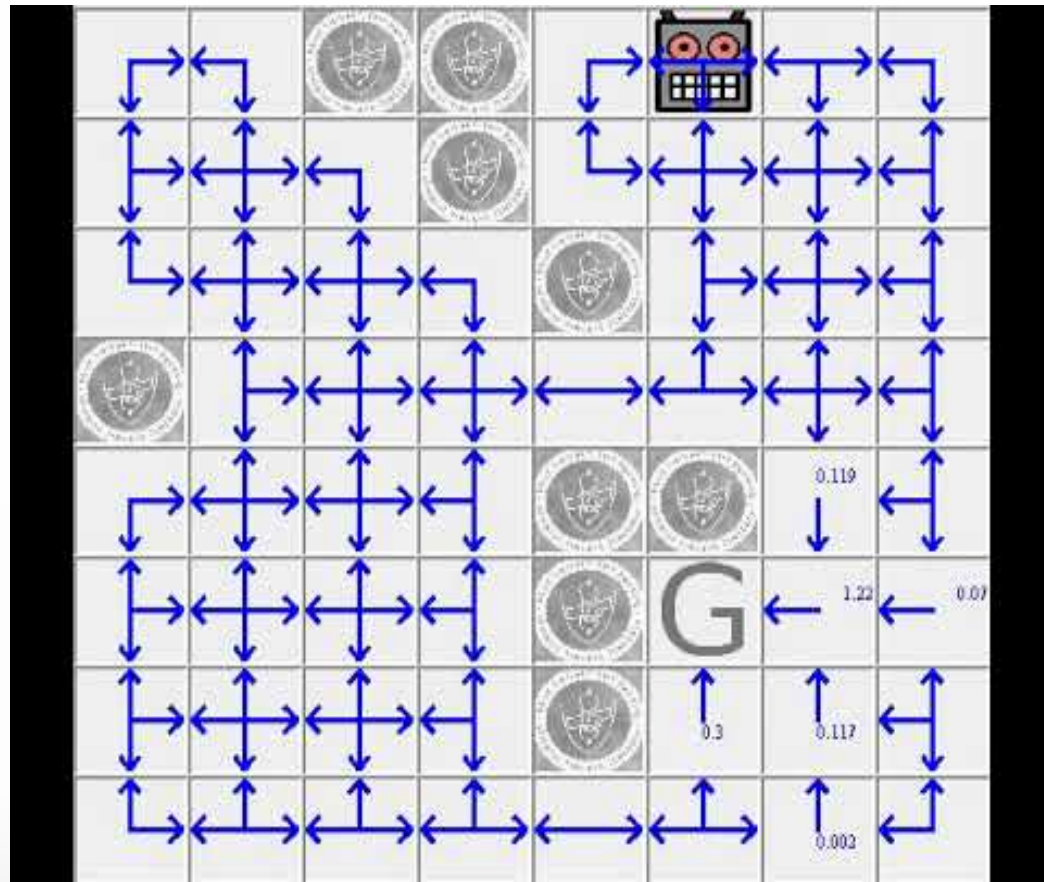


# RL: immediate and long-term reward

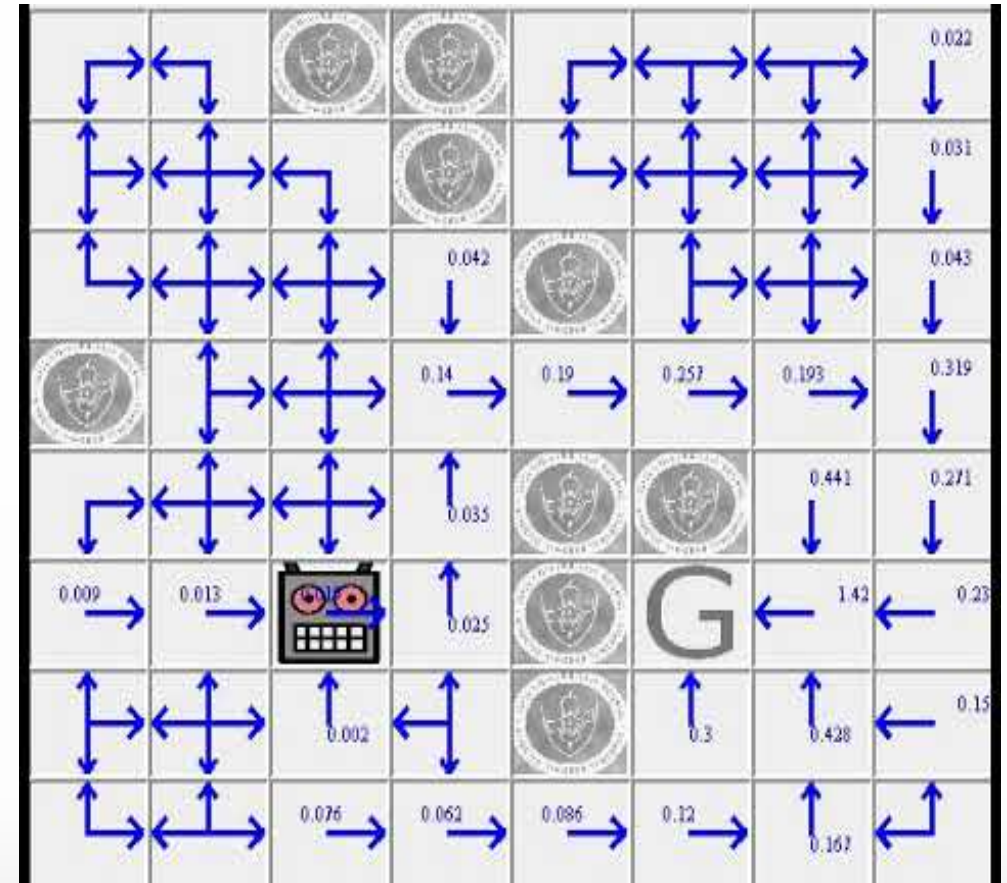


# How does RL work?

Q-Learning



Q(lambda)



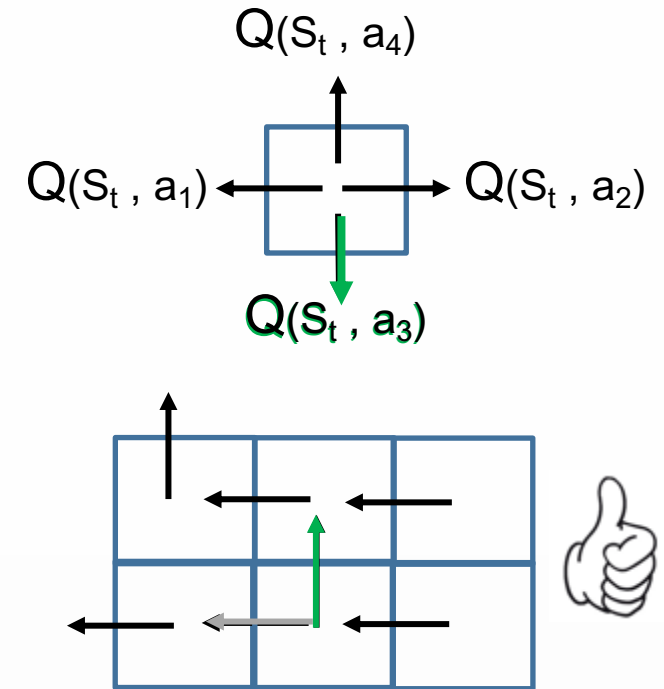
# Two main approaches

- Learn how good it is to apply a certain action in a given state (**value-based approaches**)

Q-learning → DQN, Rainbow

- Evaluate a given policy, and try to improve the policy (**actor critic methods**)

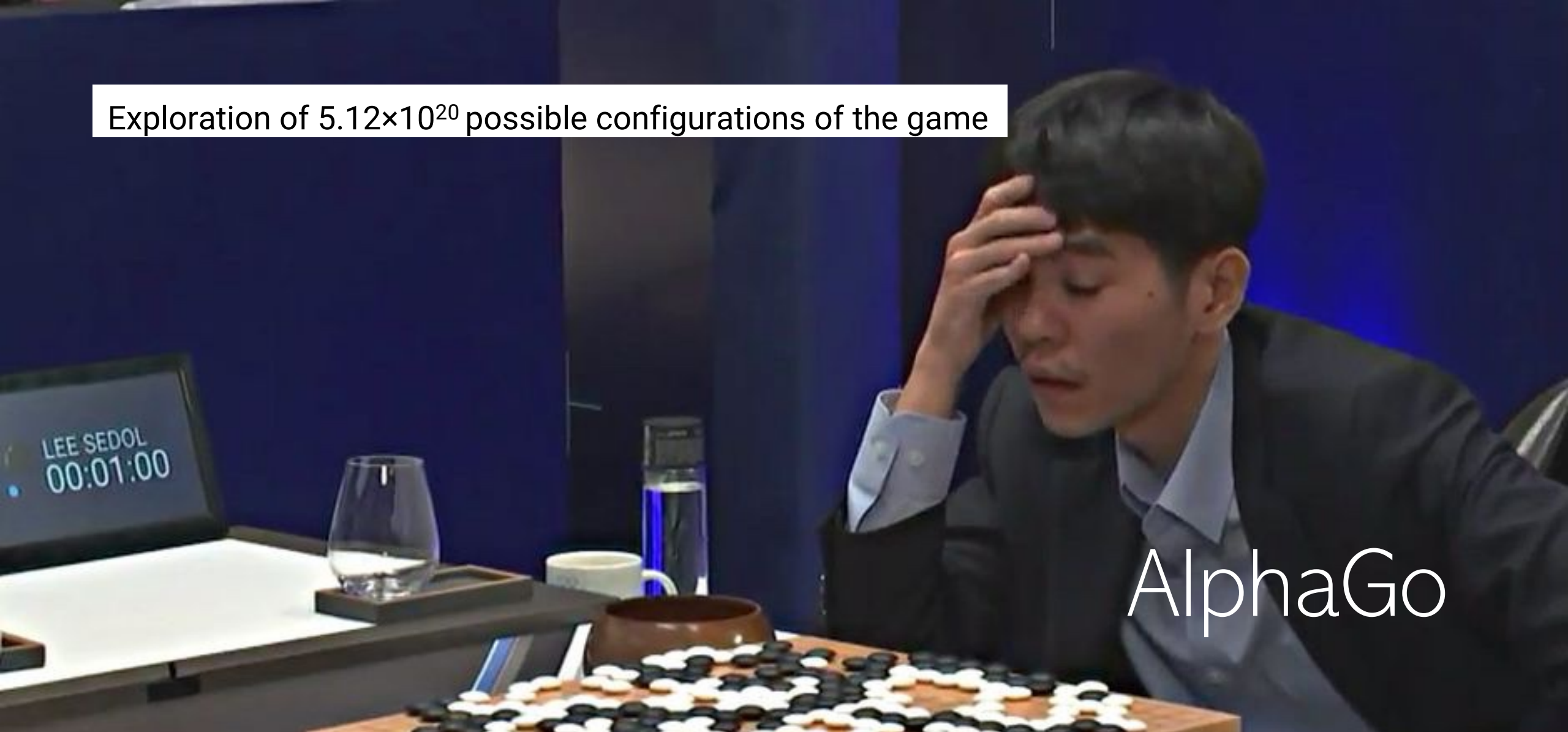
Policy Gradient → DDPG,  
Soft Actor-Critic



# Applications of Reinforcement Learning



Exploration of  $5.12 \times 10^{20}$  possible configurations of the game



AlphaGo

Difficult non-linear controller, higher optimality with RL than classical control

# Space Shuttle attitude control (1993)

Exploration of the impact of decisions on long time-scales

# Datacenter cooling

Material- and temperature-dependent complex fluid dynamics (air)

# Weaving Machine

Many constraints and multiple criteria

# EVs as Smart Batteries



Non-stationary dynamic system, unexpected events



# Wind Turbines

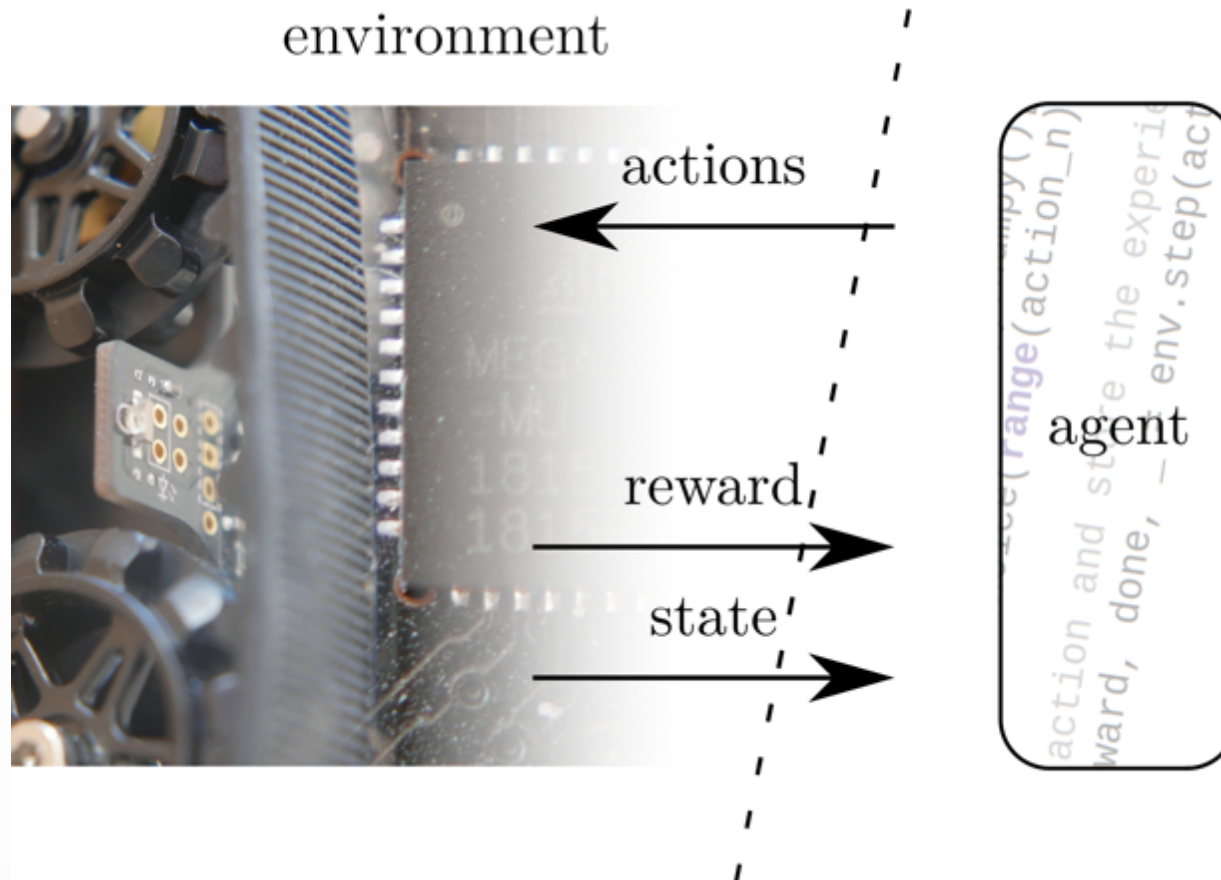
# Tackling RL challenges at the AI Lab



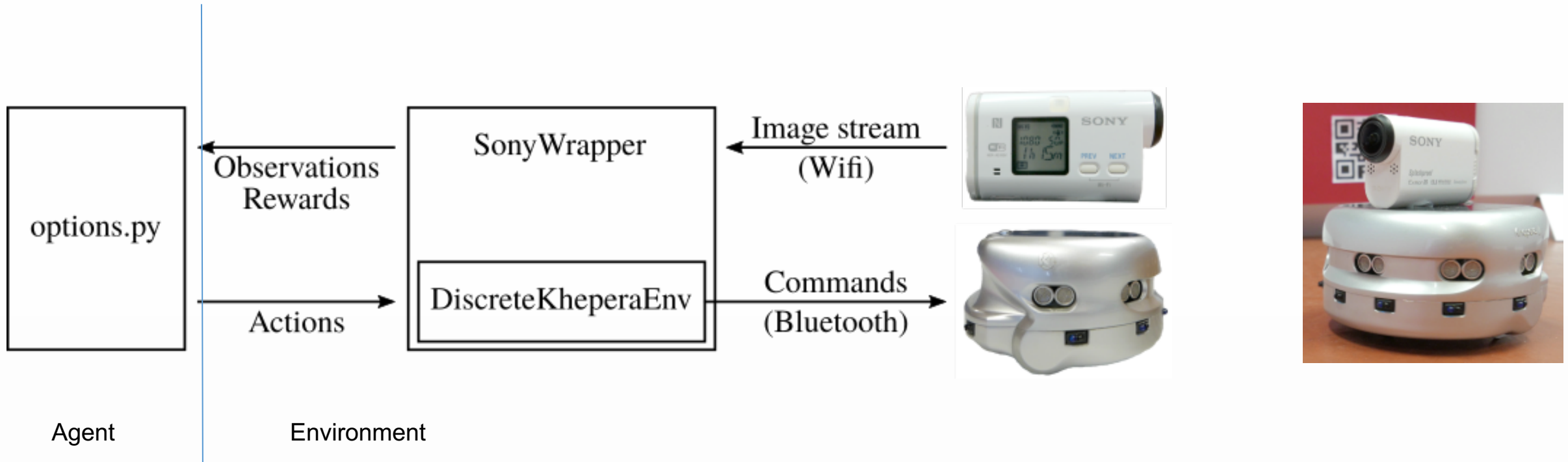
# Our motorized wheelchair



# Real-world RL

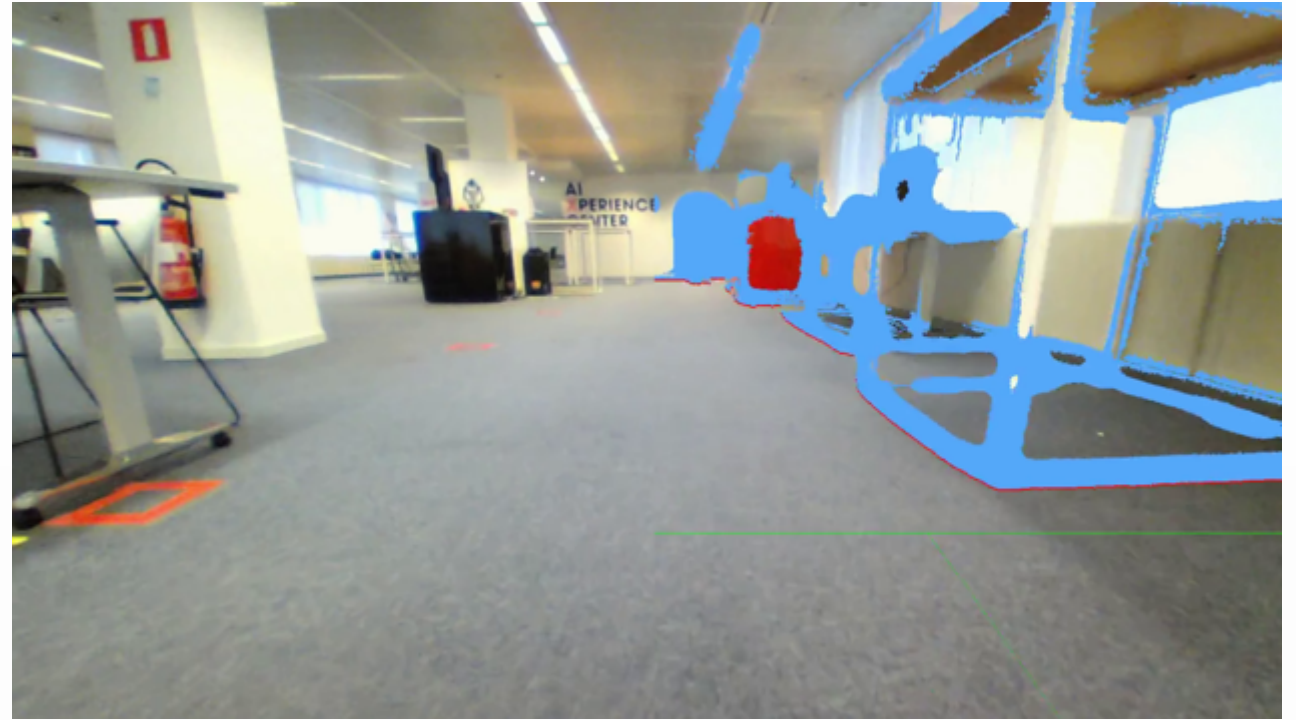


# RL Environments



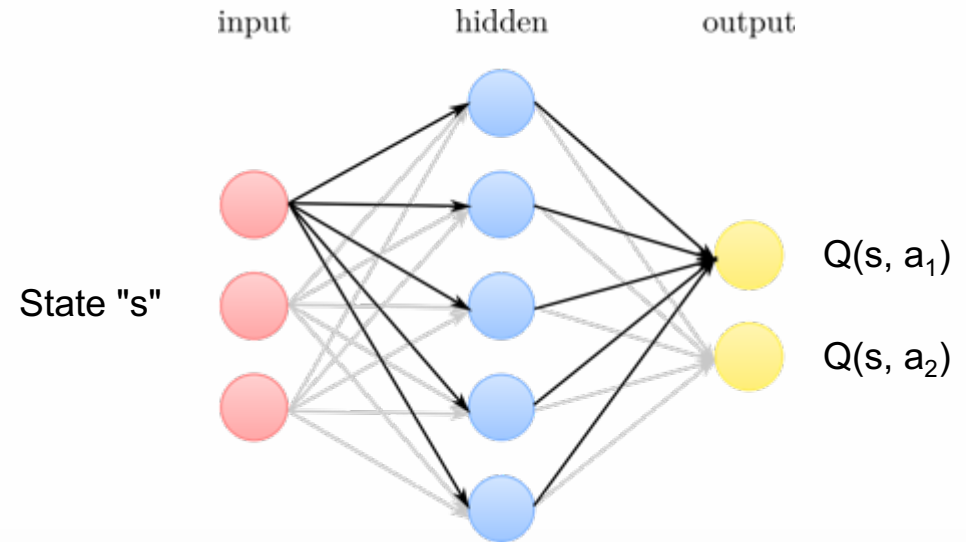


# States



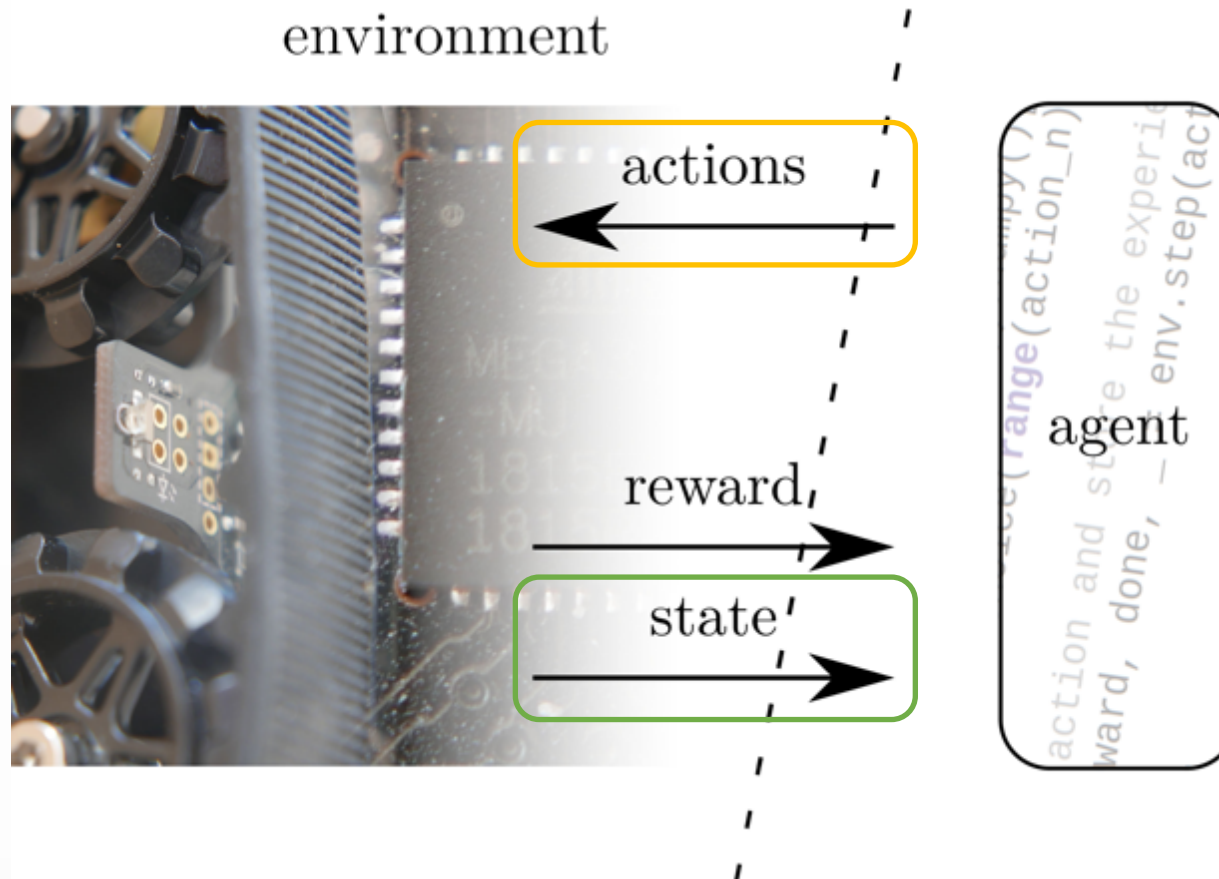
# Continuous state-spaces

State	$a_1$	$a_2$
$s_1$	$Q(s_1, a_1)$	$Q(s_1, a_2)$
$s_2$	$Q(s_2, a_1)$	$Q(s_2, a_2)$
$s_3$	$Q(s_3, a_1)$	$Q(s_3, a_2)$



- Approximation errors cause problems
- Clipped DQN, Proximal Policy Optimization

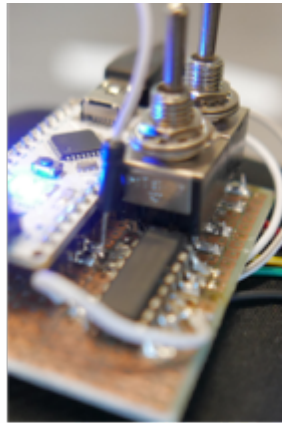
# The wheelchair: actions



# Actions



**Joystick**  
 $\pm X, \pm Y$   
 "2.5, 2.5; 0.7, 4.3"



**DAC**  
 MAX534

**Arduino**  
 Nano

**Webcam**  
 320x240 RGB  
 20 fps

USB



**Raspberry Pi**  
 Pi 3 model B

Serial USB  
 "X0.2 Y1.0"



**Threadripper**  
 2990WX  
 32x 3.8 Ghz  
 32GB RAM

Wi-Fi  
 Webcam images →  
 ← Joystick X-Y

(actual voltages)

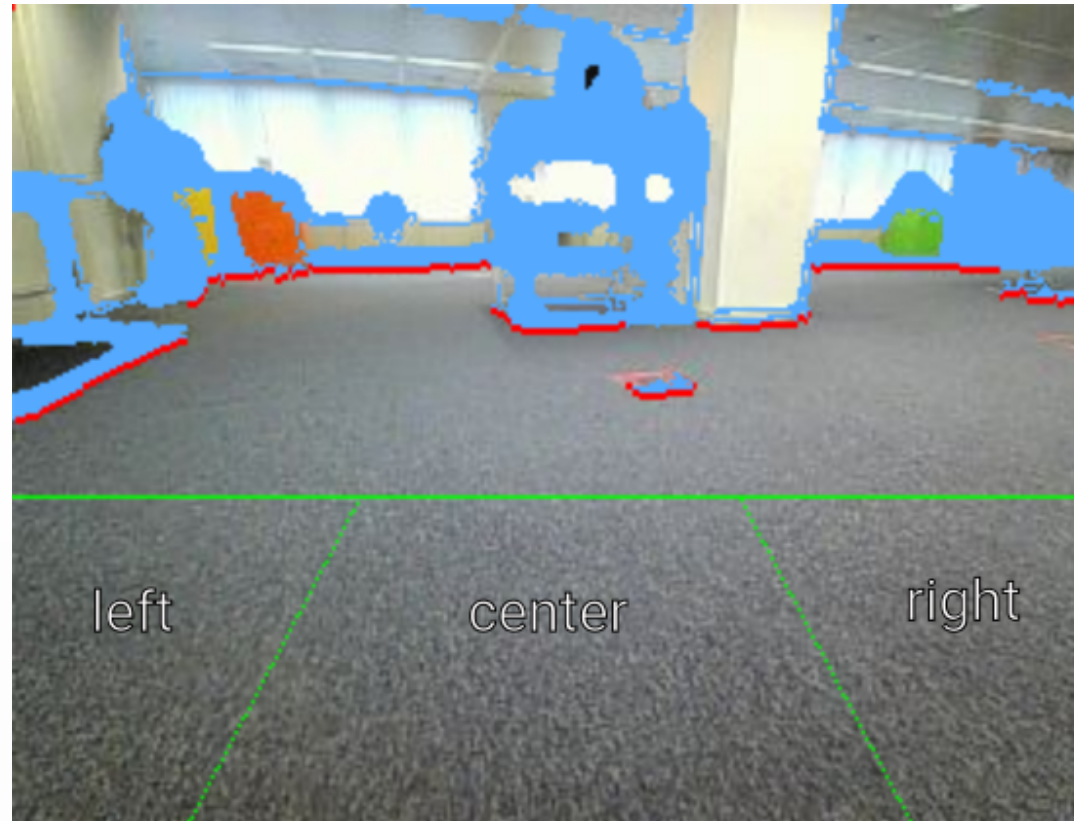
2.5V, 2.5V, 5V, 0V  
 0V, 5V, 2.5V, 2.5V  
 5V, 0V, 2.5V, 2.5V

X0 Y1  
 X-1 Y0  
 X1 Y0

0  
 1  
 2

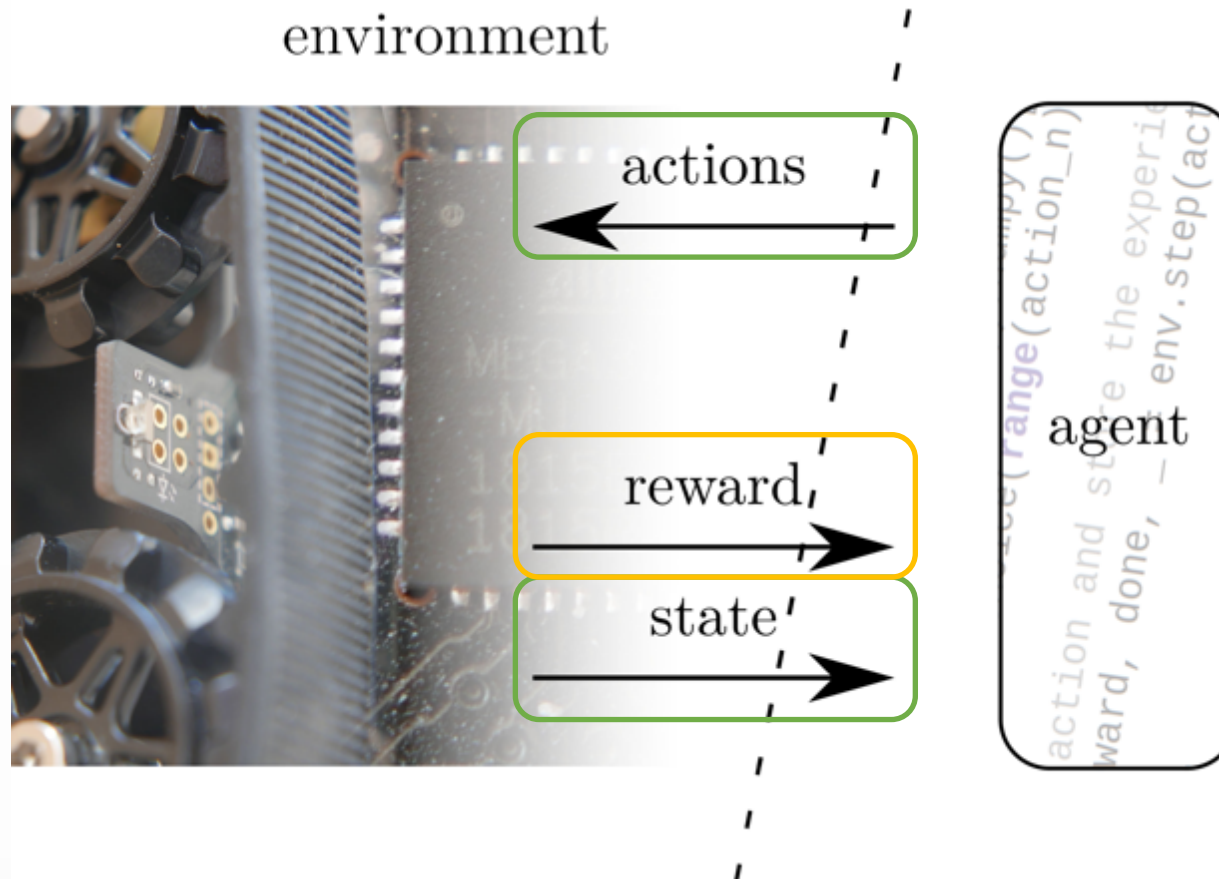
0: Forward  
 1: Left  
 2: Right

# Preventing mistakes: backup policies

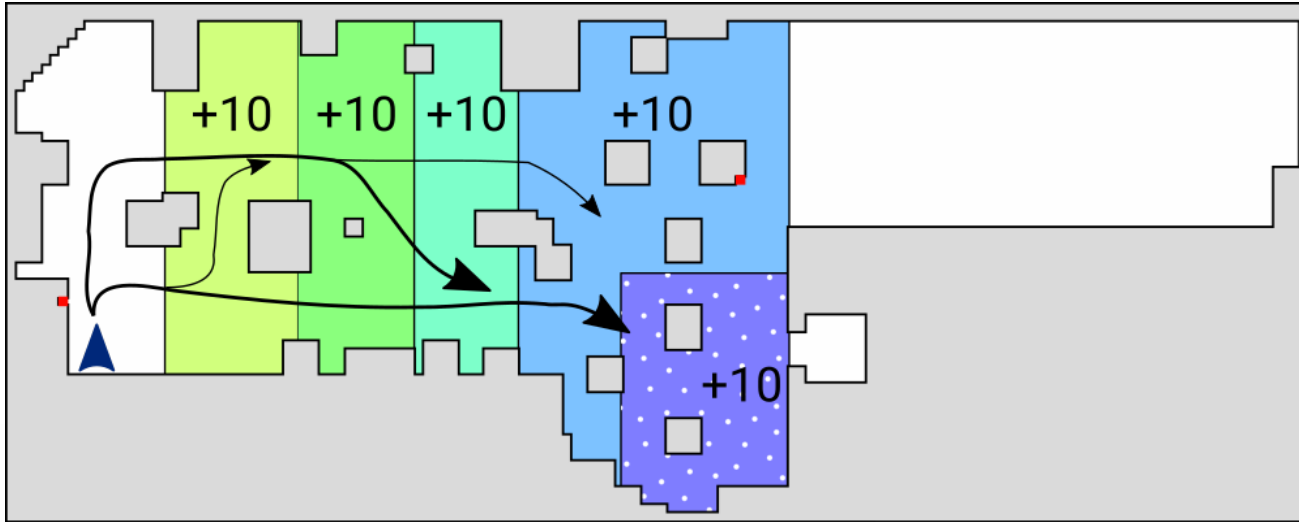




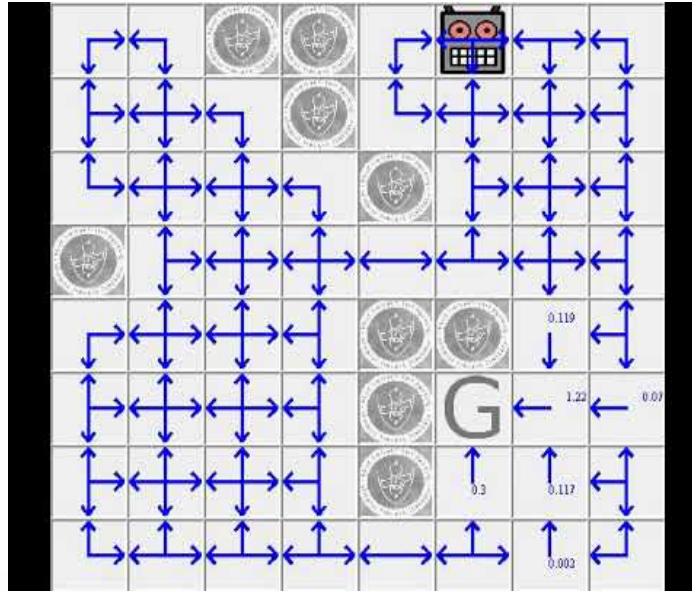
# The wheelchair: rewards



# Reward function design

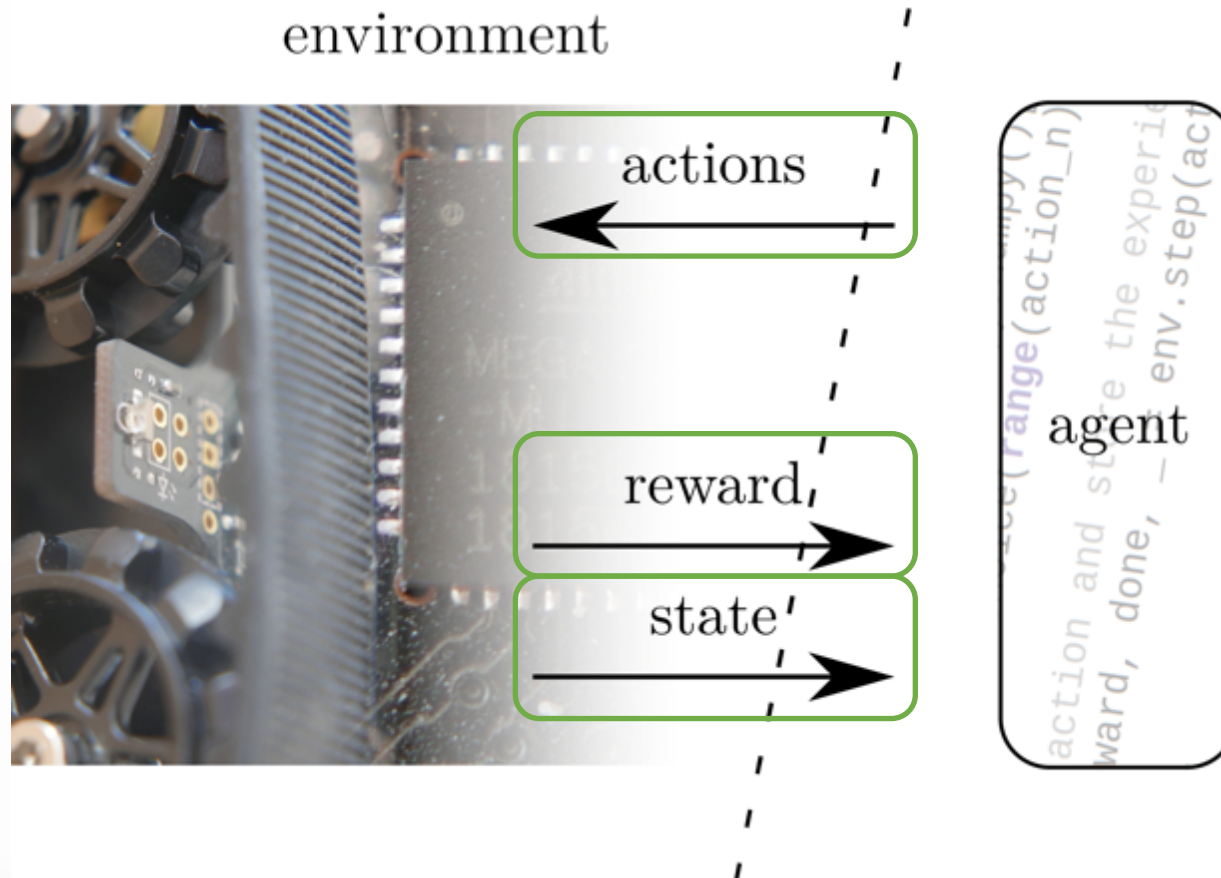


# Reward shaping



- R is the **sparse** environment reward
- F is a **shaping function**
- $R_F = R + F$  is given to the agent

# The wheelchair setting now complete

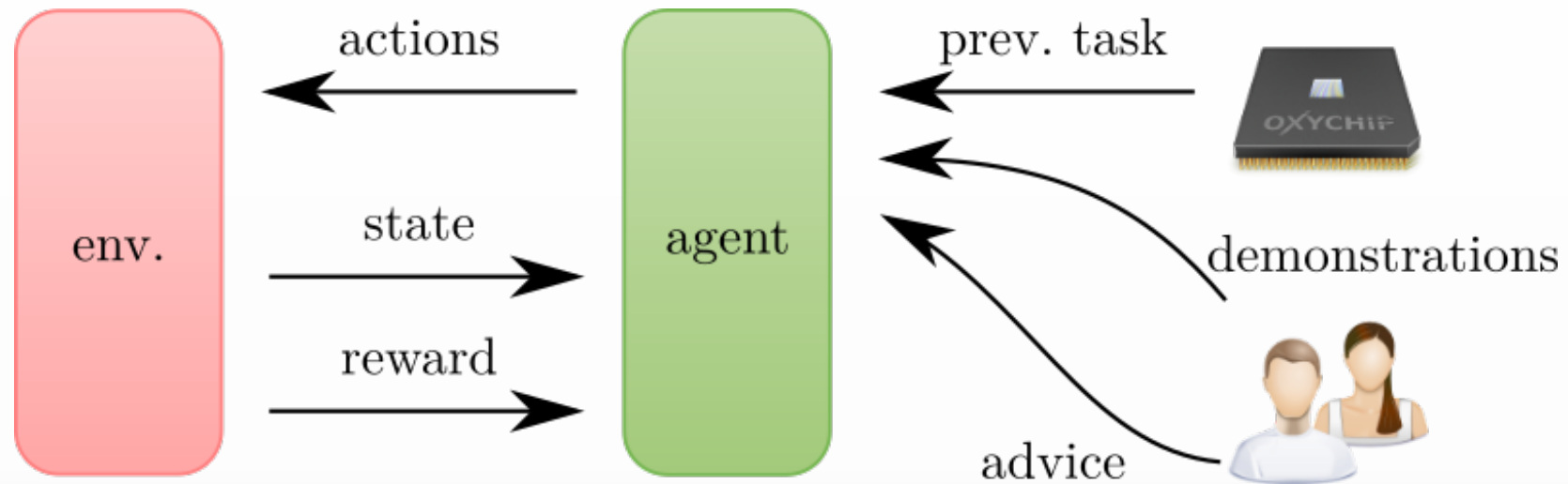


# Sample-efficiency

- Bootstrapped Dual Policy Iteration
- <https://github.com/vub-ai-lab/bdpi>

# Transfer Learning

- Quick-start learning
- The Actor-Advisor

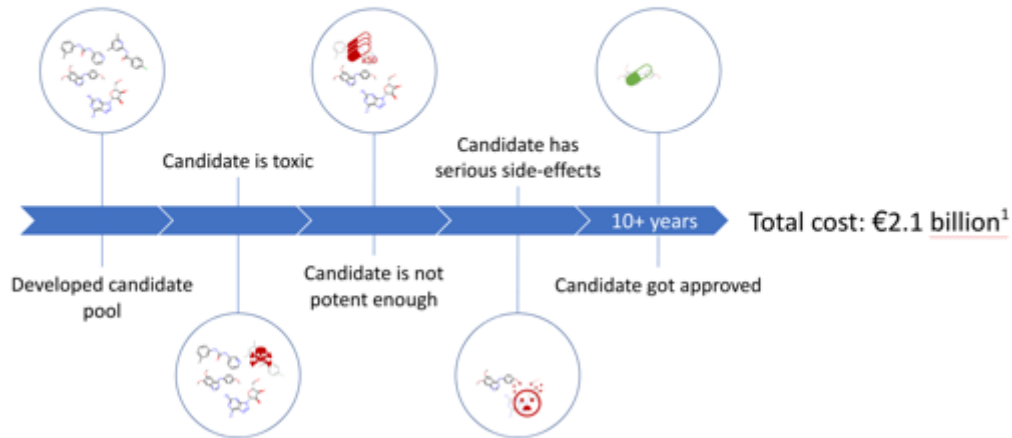


# Multiple agents

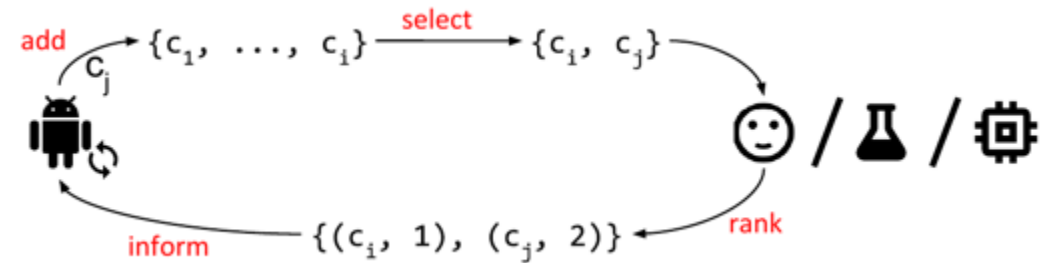


# Multi-Objective RL

Developing a drug is slow and expensive



<sup>1</sup> \$2.8 billion, exchange rate of 2013 (DiMasi et al., 2016)

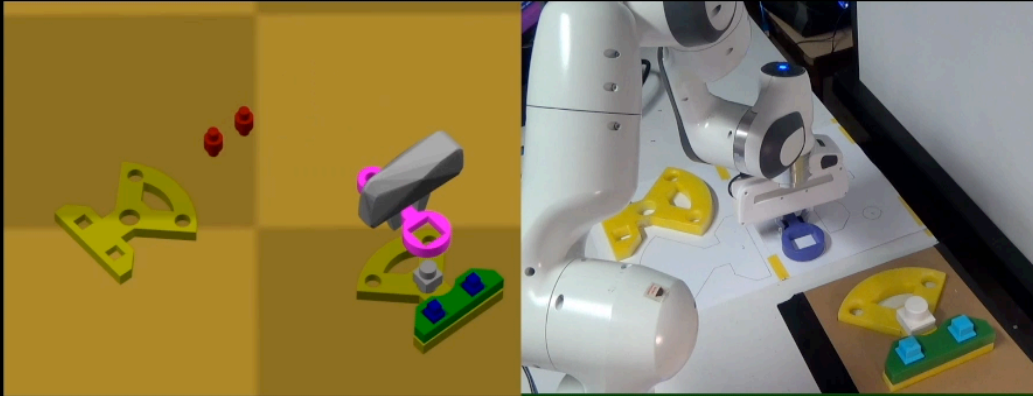


- Drug design, scheduling, manufacturing, robots, ...



# RL and digital twins

Expected skill execution



Learning  
Quadrupedal  
Locomotion  
over  
Challenging  
Terrain

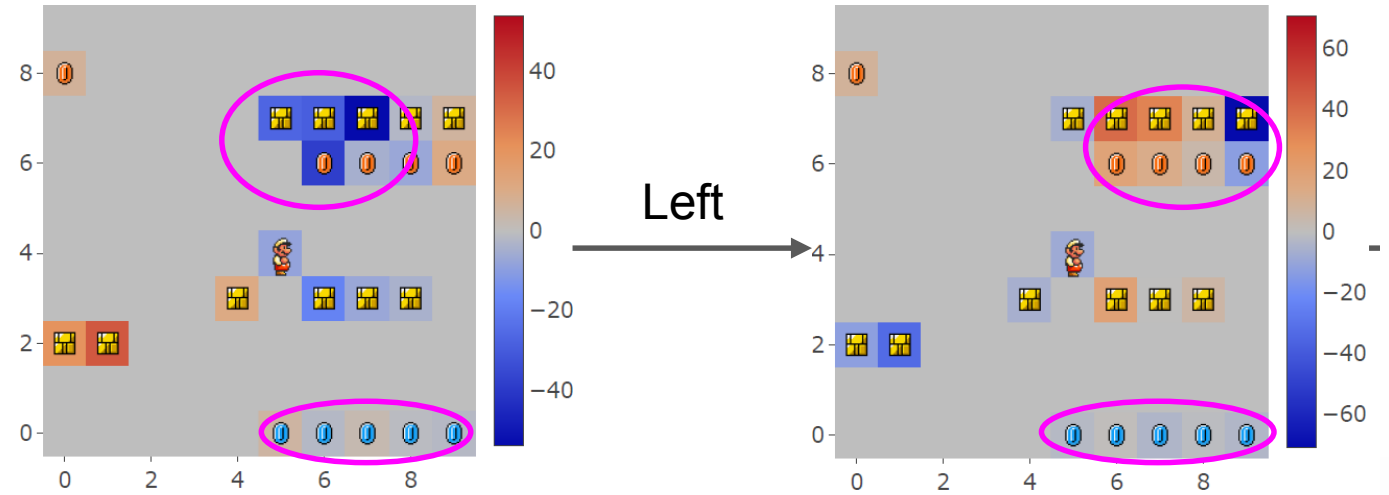
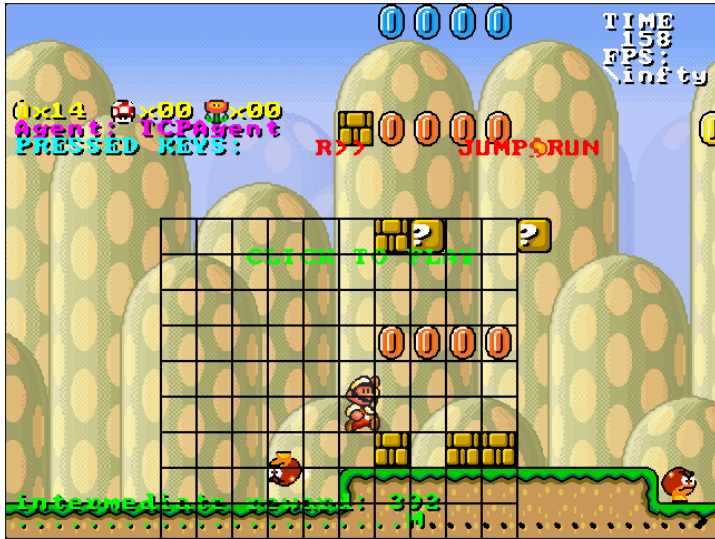


PhD Joris De Winter  
Ann Nowé (VUB AI Lab), Bram Vanderborcht (VUB R&MM )

# If an RL agent told you to sell all your shares, would you do it?

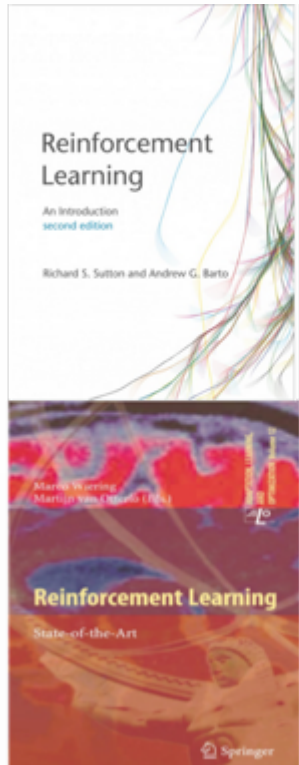


# Explainable RL



- Crucial for GDPR compliance

# More on RL?



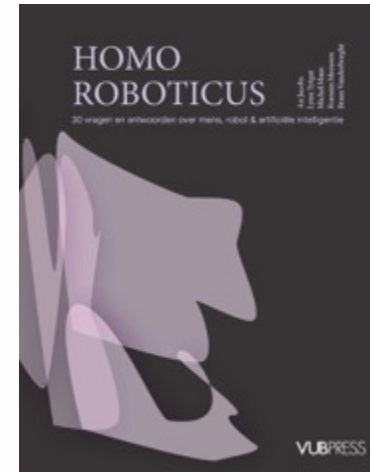
Reinforcement Learning: an introduction

R.S. Sutton and A.G. Barto

Available for free online

Reinforcement Learning: State-of-the-Art,  
M.Wiering and M. van Otterlo

Covers more advanced topics



[steckdenis.be/phd\\_thesis.pdf](https://steckdenis.be/phd_thesis.pdf)

A Gentle introduction to Reinforcement Learning  
Ann Nowé & Tim Brys

# Concluding remarks

What can RL mean for your company?

# Take-home messages

- RL goes beyond traditional machine learning and data mining techniques
  - Taking action rather than revealing patterns
  - Prescriptive, not descriptive, and takes into account long-term effects
  - Can learn in the real world, or a model, any combination of these is possible
  - Domain knowledge can be incorporated (to reduce the cost of learning)

# Take-home messages

- **RL can solve a wide variety of tasks**
  - Optimal control of machines and robots (e.g., increased productivity, safely interacting with humans)
  - Advanced planning and resource allocation (e.g., multi-objective, dynamic, personalized)
  - Fraud detection and prevention (e.g., selection which transactions to check, dynamic authentication)
  - Recommender systems (e.g., context-aware, personalized)
  - Advance price setting strategies (e.g., incentivize citizens to consume less or more energy at specific times)
  - What-if scenarios (e.g., uplift modelling, minimize churn, actionable digital twins)
  - And many more...

# Take-home messages

- RL is not a magical bullet
  - Great potential, but requires expertise
    - Defining the environment, actions and allowed transitions, objective (reward scheme), etc.
    - Incorporating domain knowledge
    - Requires learning in real life, or representative simulator, or sufficiently annotated data (states, actions, next state)
- Current challenges
  - The “human-in-the loop”: how should RL interact with humans
    - To express the objective can be challenging in multi-criteria settings
    - Trust is important
    - GDPR regulations: not only privacy-related, but also to ensure explainability and accountability
    - Further improvements on more efficient learning, e.g., coupling with digital twin



# Next steps

- **Feedback from industry**
  - On which topics would you like more information? Interest in tutorial-based seminars?
  - Which problems in your company could be solved by RL? What do you expect from RL solutions? What is in your opinion still missing? Do you have other specific concerns?
  - Survey will be sent afterwards to collect answers
- **Organize follow-up sessions (2021)**
  - Based on collected feedback
  - On more advanced technical topics or on specific applications
  - Open discussion
- **Other questions? Want to collaborate?**
  - R&D projects, applied PhDs, training, consultancy, internships, bachelor and master theses, etc.
  - More info:
    - <https://ai.vub.ac.be/services/>
    - <https://vubtechtransfer.be>

# VUB AI Experience Centre (since Nov 2019)

- “One-stop shop” for everyone interested in AI
  - Meeting place for researchers, companies, policymakers and the general public
  - Offers demonstrators/prototypes and test infrastructure
  - Used for seminars, company visits, events
- Partnership between 10+ VUB research groups



More info:  
[experience-centre.ai](https://experience-centre.ai)

Hans De Canck  
[hans.de.canck@vub.be](mailto:hans.de.canck@vub.be)



@aibrussels



## Wegwijs in Artificiële Intelligentie: van idee tot realisatie

Artificiële intelligentie (AI), wat is het en hoe gaat u ermee aan de slag? Voor welke valkuilen moet u opletten? Laat u inspireren en begeleiden door Agoria's AI-experten, praktijkgerichte tools en ons netwerk van specialisten om uw bedrijf de veerkracht te geven die nodig is om te blijven groeien en innoveren.



### Wat is AI en wat zijn de business drivers?

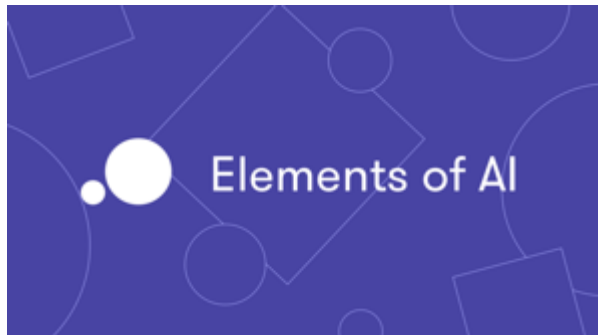
Artificiële intelligentie is niet de zoveelste hype. De technologie wordt al ingezet in tal van sectoren: gezondheidszorg, energie, maakindustrie en nog veel meer. Leer uit de inzichten en ervaringen van meer dan dertig Belgische



### Hoe vermijdt u de typische AI-valkuilen?

Uit ervaring weten we dat men vóór, tijdens of na de ontwikkeling van een AI-toepassing vaak belangrijke elementen over het hoofd ziet. Die elementen noemen we de AI Blindspots. Om te ontdekken welke AI Blindspots uw project

# Other AI courses



<https://www.elementsofai.com>

## Lifelong Learning Program

<https://ai.vub.ac.be/lifelong-learning-program/>

### WE ARE AN EDUCATION PROGRAM FOR PROFESSIONALS

We provide training and coaching to non-academic professionals who need to work with AI in their daily lives



#### For Policymakers & Journalists

What **impact** does AI have on society and ethics?

We can help you formulate your own opinion by explaining to you what AI is and sharing with you in-depth information on applications of today. We can also show you the projected role of humans in the future development of AI technology.



#### For CEOs & Tech Investors

What **value** can AI bring to a business?

We can help you establish your niche by drafting a roadmap specific to your operations and anticipative of tech trends. We have the tools to assess the AI maturity of your company objectively, trace out its market potential and identify obstacles in your future business development.



#### For Data Scientists, Architects & Developers

How do I **implement** AI into my project concretely?

If you need in-depth technical help, we can help you implement algorithms, perform data analysis and verify the scientific rigour of your AI architecture.

# Interested in collaborating? Contact us!



**Ann Nowé**

Director AI Lab

Ringleader Flanders AI  
Research Program

[ann.nowe@vub.be](mailto:ann.nowe@vub.be)



**Denis Steckelmacher**

Senior researcher  
Reinforcement Learning

[denis.steckelmacher@vub.be](mailto:denis.steckelmacher@vub.be)



**Leander Schietgat**

Challenge manager  
Flanders AI Research  
Program

[leander.schietgat@vub.be](mailto:leander.schietgat@vub.be)



**Gill Balcaen**

Business developer

[gill.balcaen@vub.be](mailto:gill.balcaen@vub.be)

Thanks for  
your attention

