# Von AlphaZero zur Mars-Rover-Autonomie

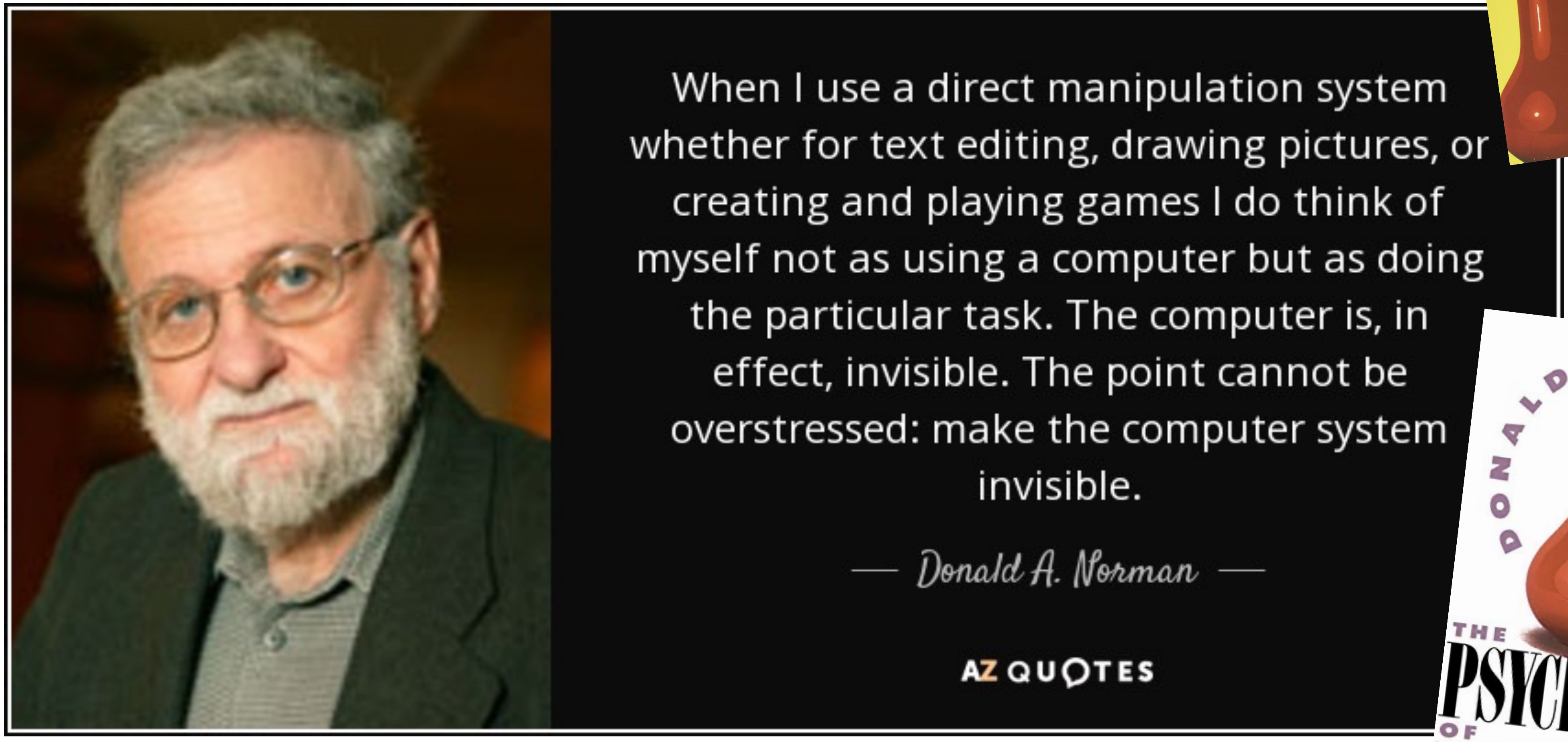
Oder: Warum KI sehr viel mehr umfasst als maschinelles Lernen!

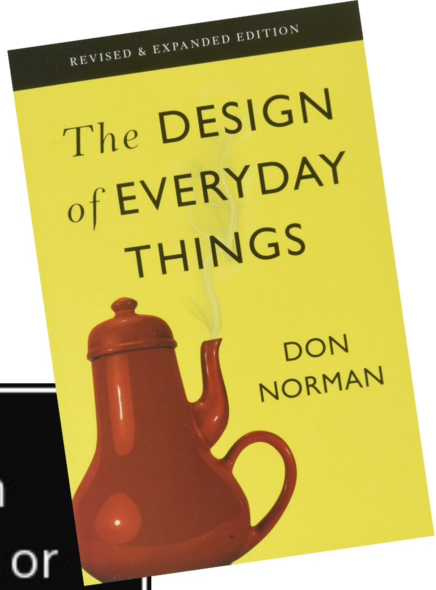
Prof. Dr. Ralf Möller

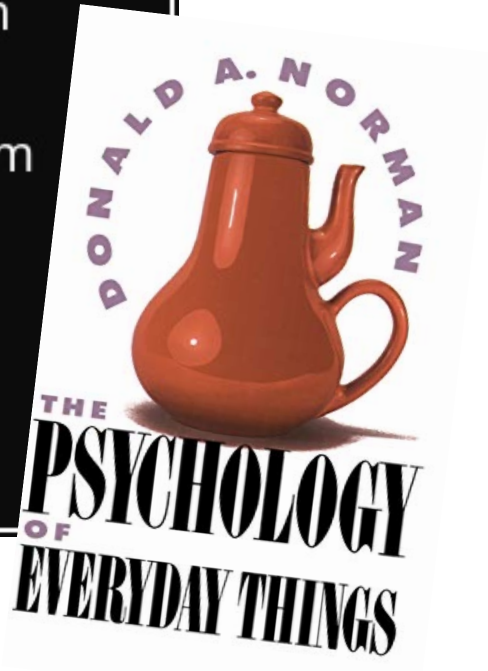
Direktor Institut für Informationssysteme, UzL
Forschungsbereichsleiter Human-Centered AI, DFKI Lübeck

UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR INFORMATIONSSYSTEME

# Early Days: Direct Manipulation

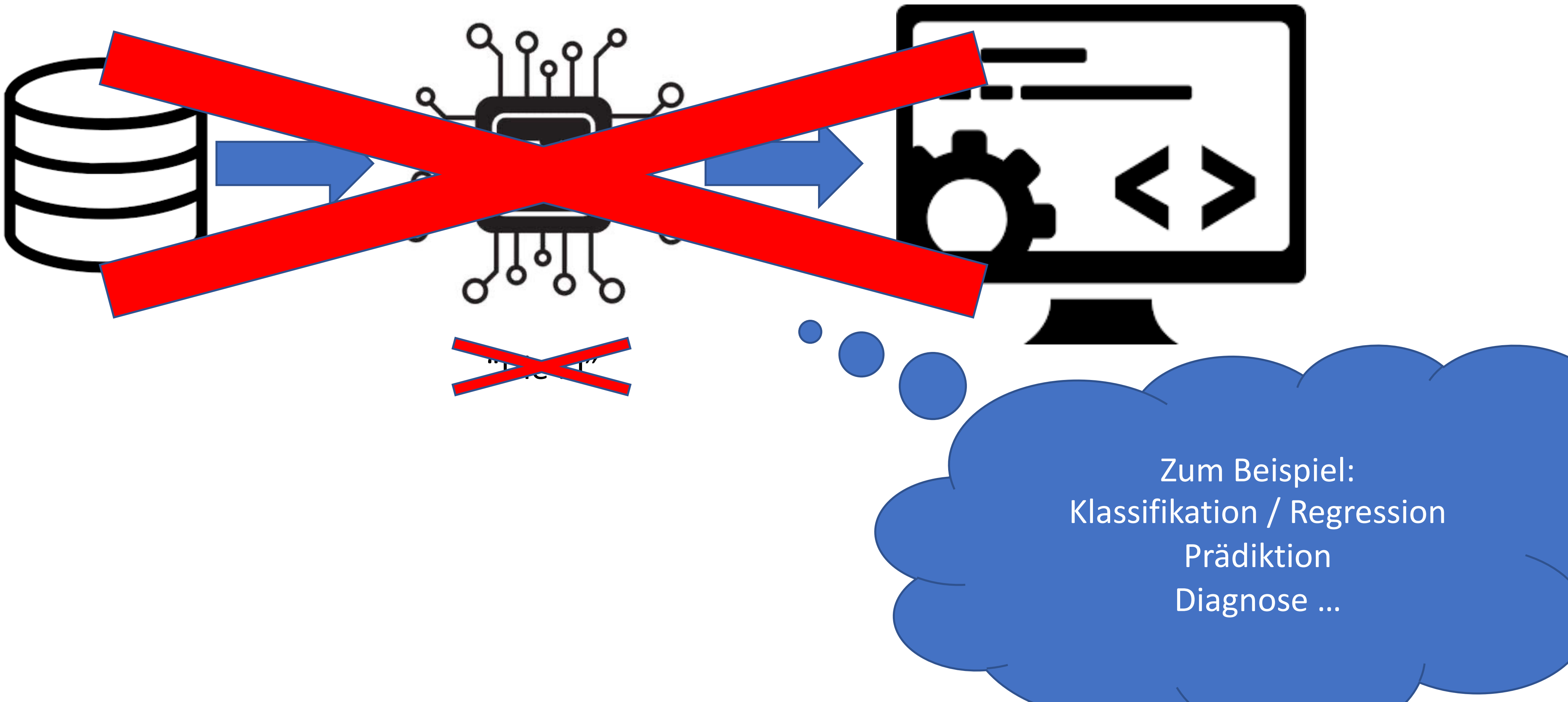


# AI Age: Instruct systems to autonomously act in a beneficial way for a group of humans

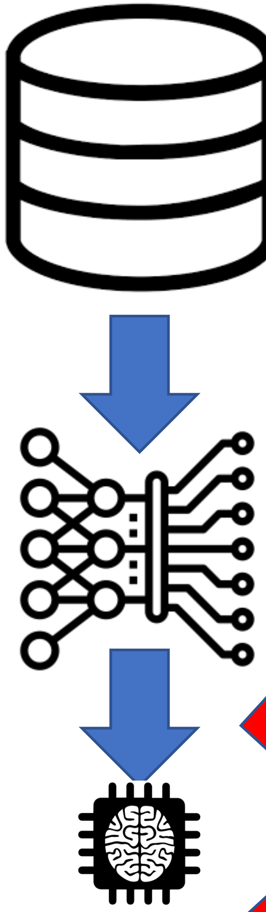
# Märchen


Zum Beispiel:
Klassifikation / Regression
Prädiktion
Diagnose …

# Maschinelles Lernen

Data Mining?

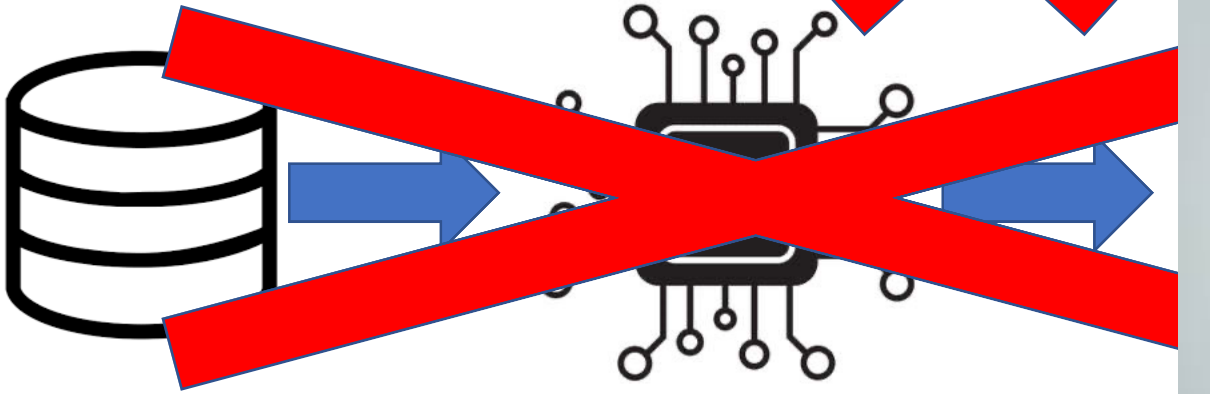
Kombination?
(vgl. Makridakis Competitions)



Machine Learning System

ID3 (Entscheidungsbaumgenerator),
Tensorflow,
…

Programm:
Entscheidungsbaumevaluator,
Kaskade von linearen und stückweise linearen Funktionen,
…




https://en.wikipedia.org/wiki/Makridakis_Competitions

# Difference between Data Mining and ML?

https://www.discoverdatascience.org/articles/data-mining-vs-machine-learning/

## What is the Difference Between Data Mining and Machine Learning?

Data mining is the probing of available datasets in order to identify patterns and anomalies. Machine learning is the process of machines (a.k.a. computers) learning from heterogeneous data in a way that mimics the human learning process. The two concepts together enable both past data characterization and future data prediction.

Well, current ML does not even attempt to show that learning processes are cognitively plausible

But: Simple data mining has the same fundamental problems

Data mining deals with secondary memory (more serious for big data)

In principle, there is no difference between data mining and ML.

# Simpson's Paradox (Example)

- Record recovery rates of 700 patients given access to a drug

| | **Recovery rate with drug** | **Recovery rate without drug** |
|---|---|---|
| Men | 81/87 (93%) | 234/270 (87%) |
| Women | 192/263 (73%) | 55/80 (69%) |
| Combined | 273/350 (78%) | 289/350 (83%) |

- Paradox:
  - For men, taking drugs has benefit
  - For women, taking drugs has benefit, too.
  - But: for all persons taking drugs has no benefit

# Resolving the Paradox (Informally)

- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox

- In drug example
  - Why has taking drug less benefit for women?

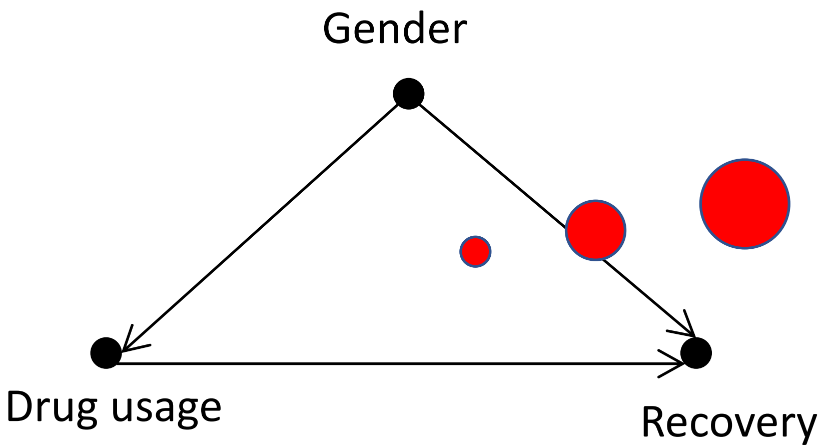
    Answer: Estrogen has negative effect on recovery
  - Data: Women more likely to take drug than men
  - Choosing randomly any person taking drugs will rather give a woman – and for these recovery is less beneficial

- In this case: Have to consider segregated data

  (not aggregated data)

# Resolving the Paradox

- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox




Ohne (kausale) Modelle kann man Daten nicht interpretieren!

- Drug usage and recovery have common cause
- Gender is a confounder

# Simpson's Paradox (Again)

- Record recovery rates of 700 patients given access to a drug w.r.t. blood pressure (BP) segregation

| | **Recovery rate without drug** | **Recovery rate with drug** |
|---|---|---|
| Low BP | 81/87 (93%) | 234/270 (87%) |
| High BP | 192/263 (73%) | 55/80 (69%) |
| Combined | 273/350 (78%) | 289/350 (83%) |

- BP recorded at end of experiment
- This time segregated data recommend not using drug whereas aggregated data does

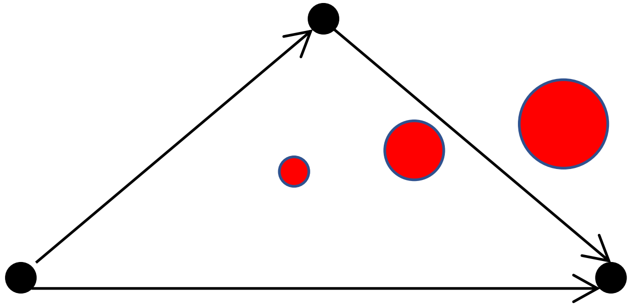
# Resolving the Paradox (Informally)

- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox

- In this example
  - Drug effect is: lowering blood pressure (but may have toxic effects)
  - Hence: In aggregated population drug usage recommended
  - In segregated data one sees only toxic effects

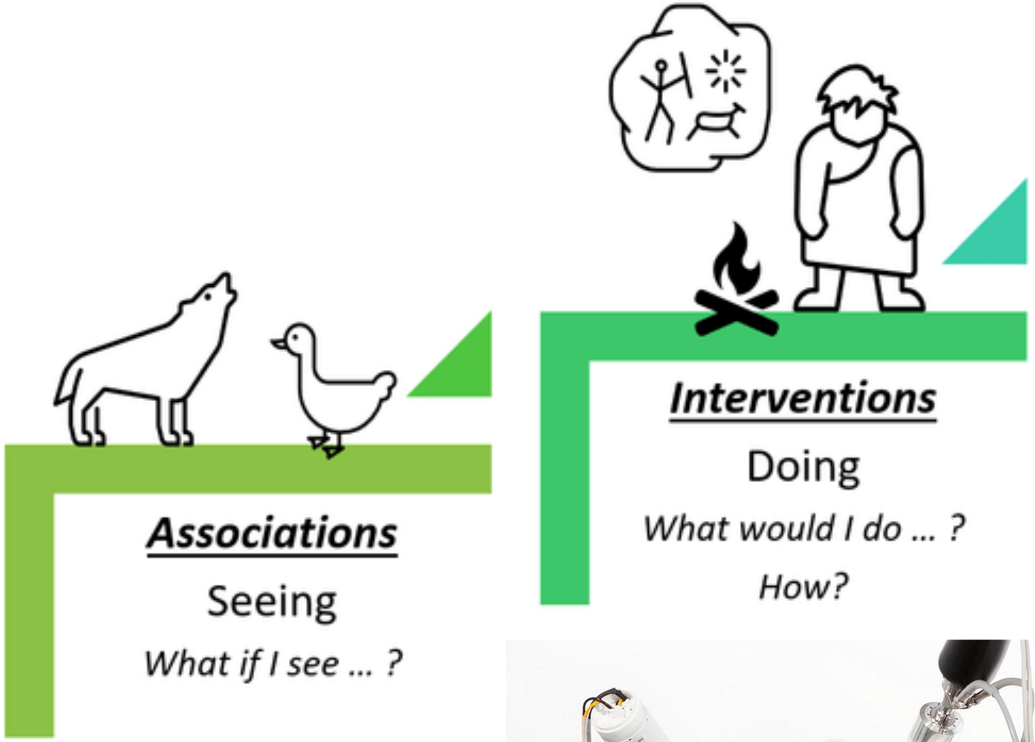
# Resolving the Paradox

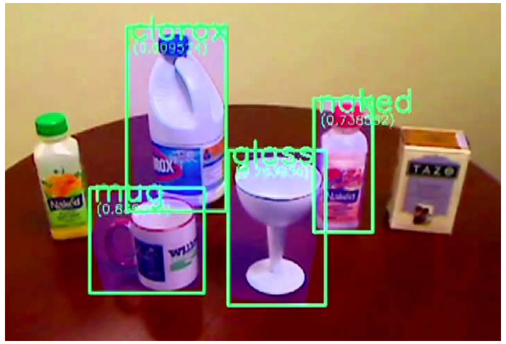
- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox

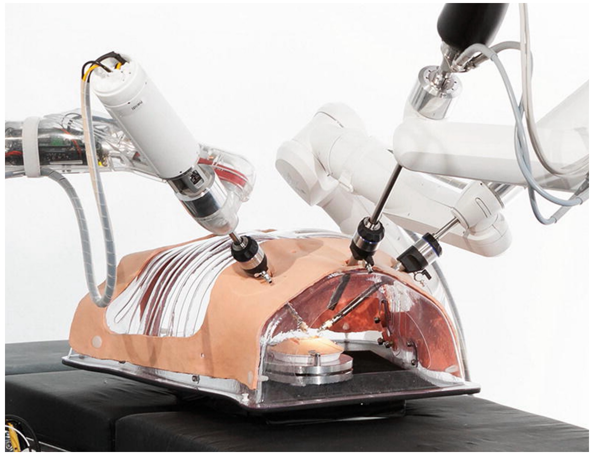



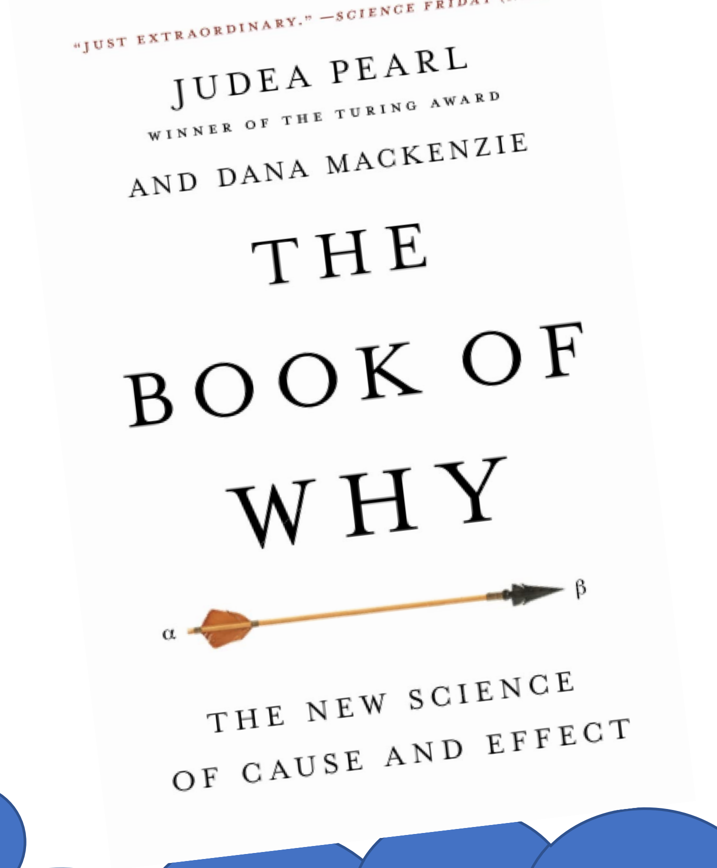

Ohne (kausale) Modelle kann man Daten nicht interpretieren!

# AI: Möglichkeiten



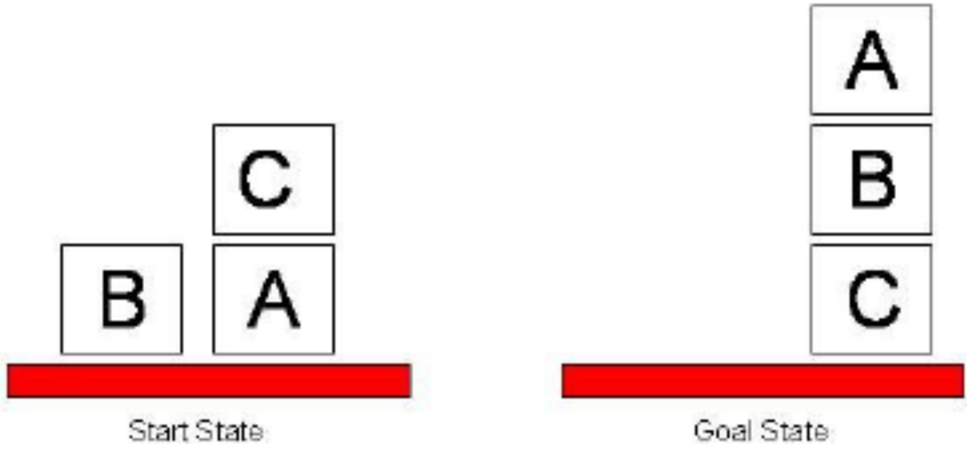



Was könnte man besser machen?

Interne Generierung von “Daten” bisher nicht immer betrachtet
Man verlässt sich zu sehr auf “externe” Daten

# Meine These

- KI ist nicht nur Datenanalyse (Data Science) mit sog. “KI-Methoden”
- Nur, wenn wir Systeme betrachten, die auch handeln …
- … wird die KI wird richtig interessant



# Intelligenter Agent

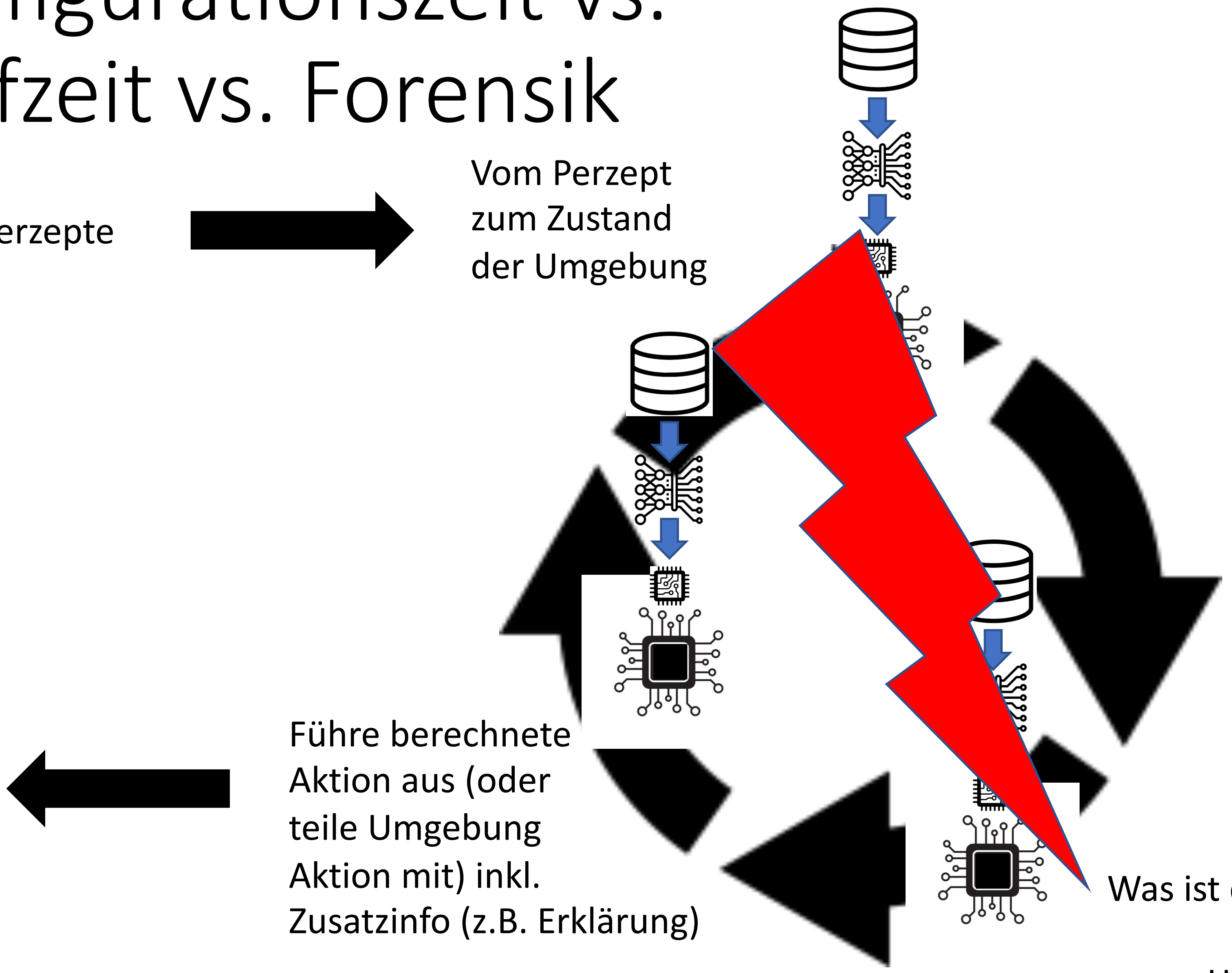
- Bestimmt Handlungen nach Auswertung (einer Folge) von Eingaben,
  - so dass der Agent durch Interaktion mit der Umgebung irgendwann seine Aufgabe erfüllt,
  - oder – anders gesagt – so dass ein lokales Nützlichkeitsma… …imiert wird
- Bewältigt mit gegebener Handlungsbestimmungs… …überzeugender Weise eine Aufgabe (oder löst ein relevantes Pro…
  - von der (dem) wir Menschen denken,
  - dass ein automatisches System sie (oder es) beim derzei… …d der Kunst der Systementwicklung kaum lösen k…nn
- Das Beeindruckt-Sein des Me…en bedingt ä… …zuschreibung nicht die verwendeten “Inner…
  - Wichtiger Teilaspekt der Intelligenz:
  - Intelligenzzuschreibung relativiert sich … …ohn… an die Leistung
- Intelligentes System: Zusammenspiel von Agente…

# Konfigurationszeit vs. Laufzeit vs. Forensik

# Lernen Planung

Perzepte

Vom Perzept zum Zustand der Umgebung

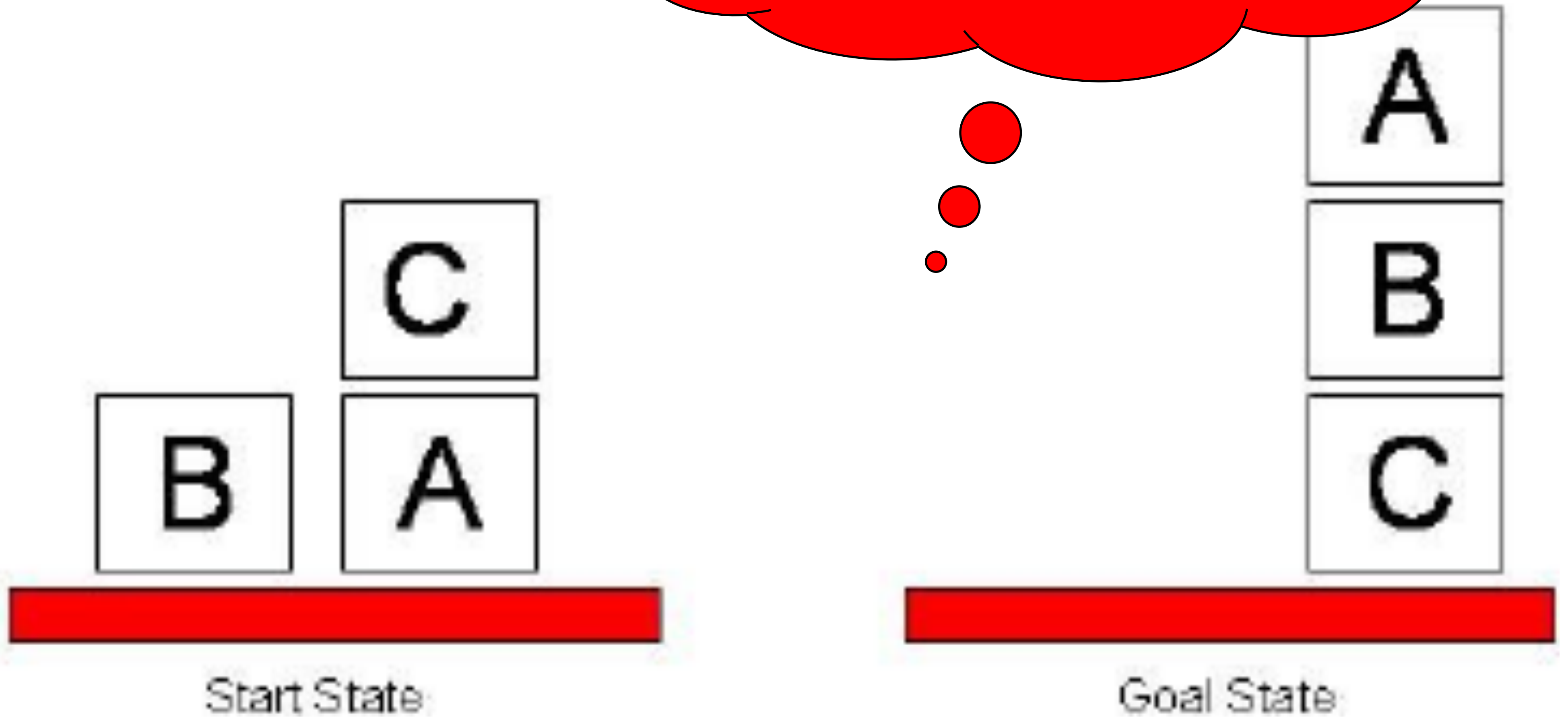


Führe berechnete Aktion aus (oder teile Umgebung Aktion mit) inkl. Zusatzinfo (z.B. Erklärung)

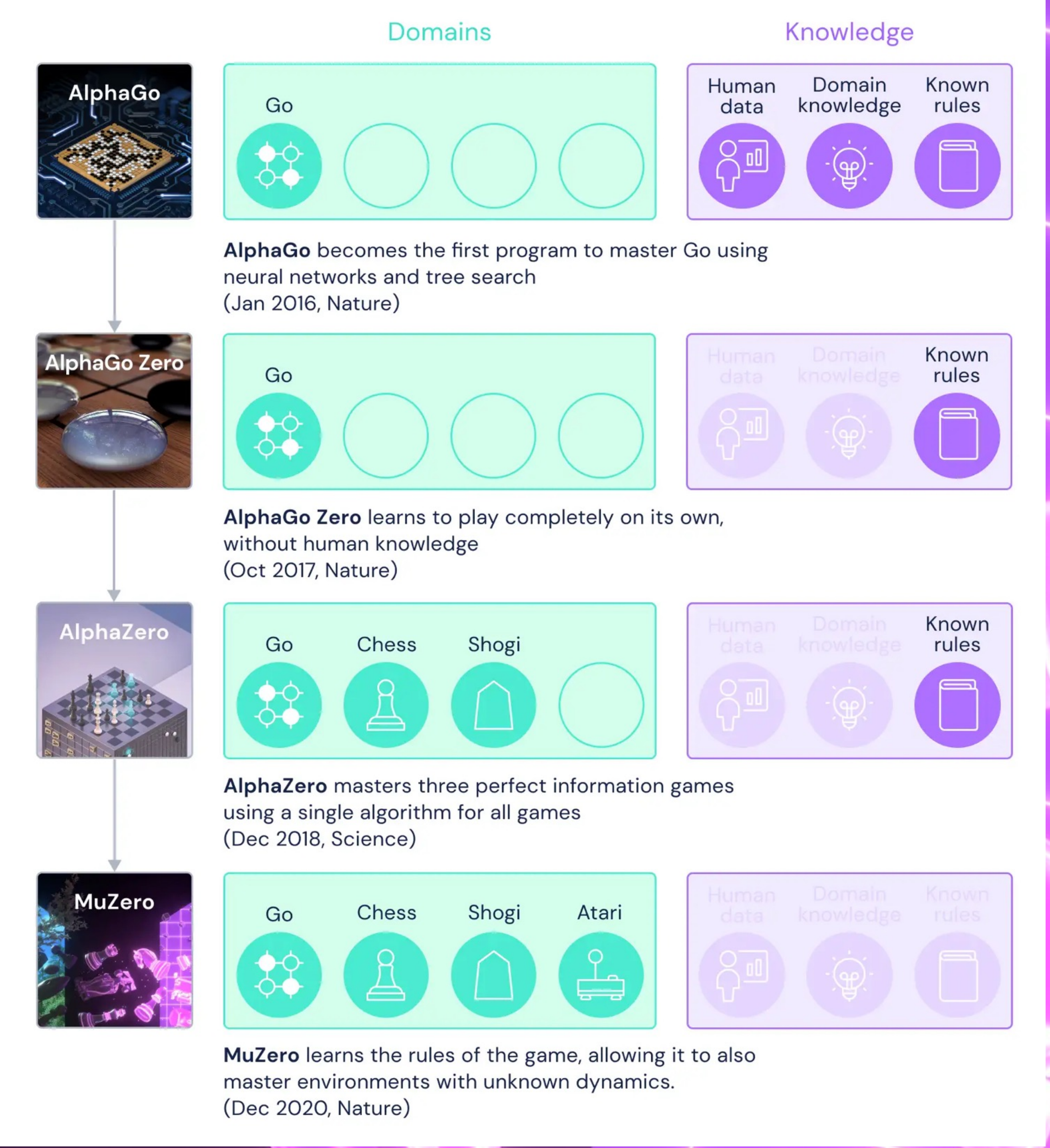
Was ist die beste Aktion im aktuellen Zustand bei gegebenem Ziel?
Handlungsbestimmungsstrategie

# Planen im Zustandsraum

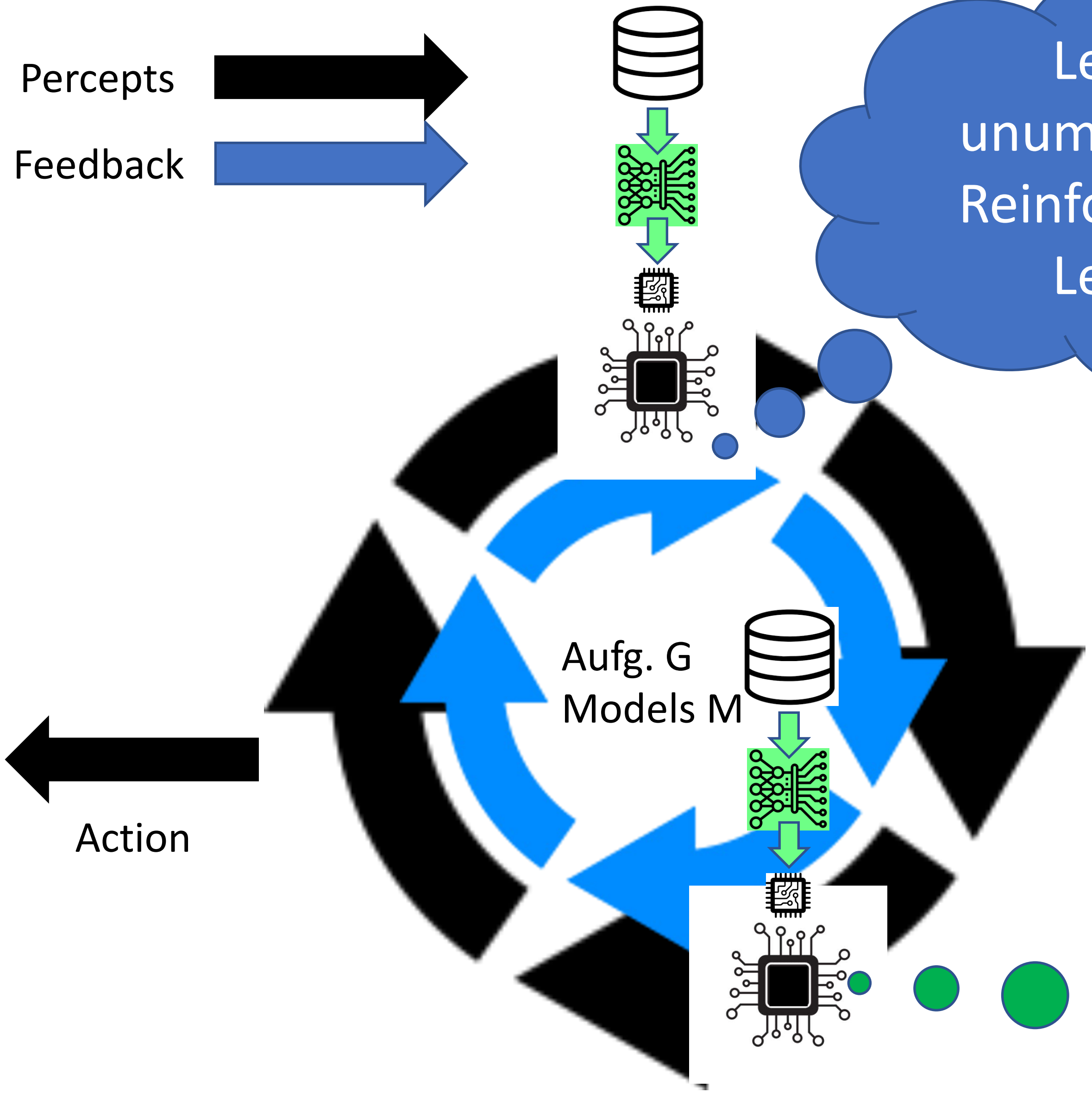
Brittleness-Problem der initialen Handlungsmodelle



# Against the Brittleness: MuZero (DeepMind)



# Some more detail…



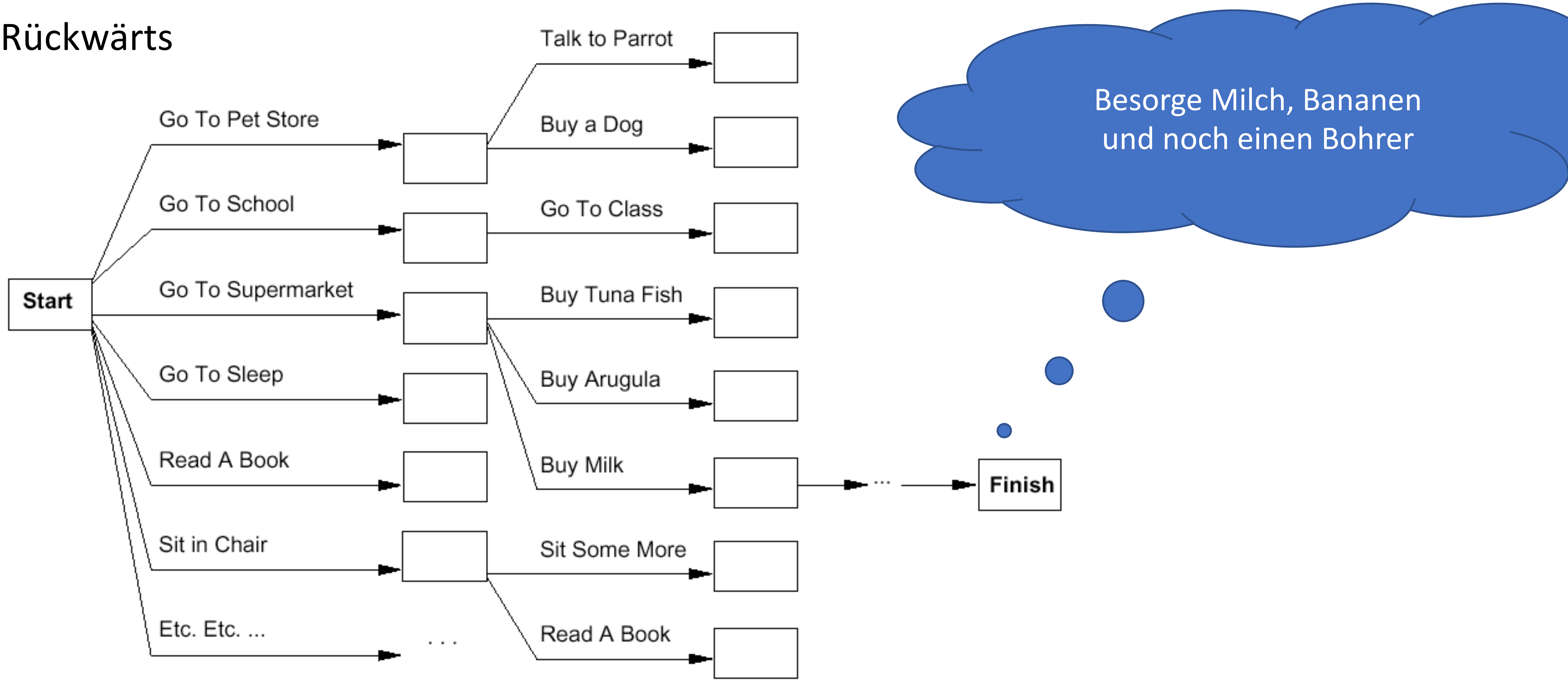

Lernen unumgänglich: Reinforcement Lernen

Nicht immer wird der beste Zug genommen (Exploration/Counterfactuals: Hätte ich das und das gemacht, was wäre die Konsequenz?)

Vom Start zum Ziel? Planen

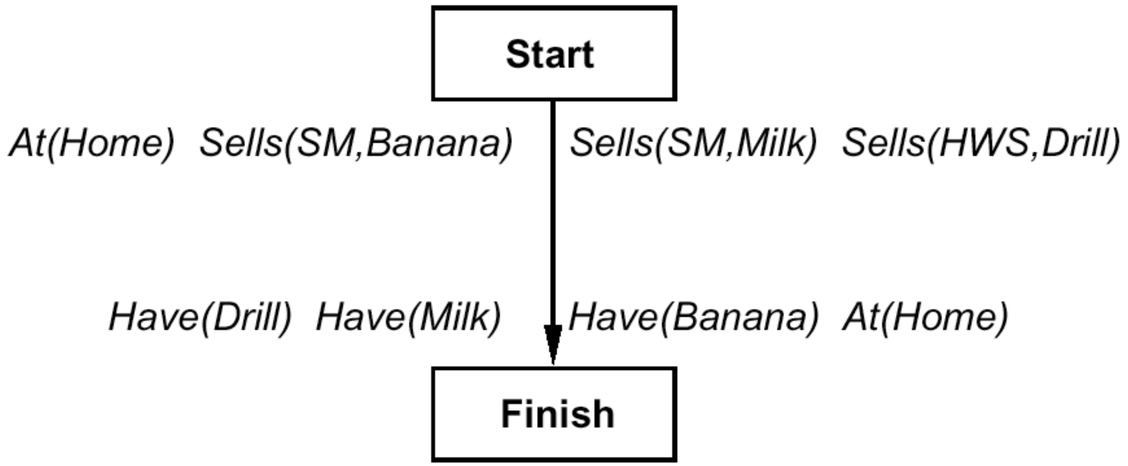
# Vom Start zum Ziel: Planen als Suche

- Vorwärts
- Rückwärts




Aus einer KI-Vorlesung von Bernhard Nebel

# Plan: Handlungsfolge vom Startzustand zum Zielzustand




**Aktionen:**

*Op* (Action: *go(there)*,
Precond: *At(here)*,
Effect: *At(there)* $\wedge$ $\neg$*At(here)*)

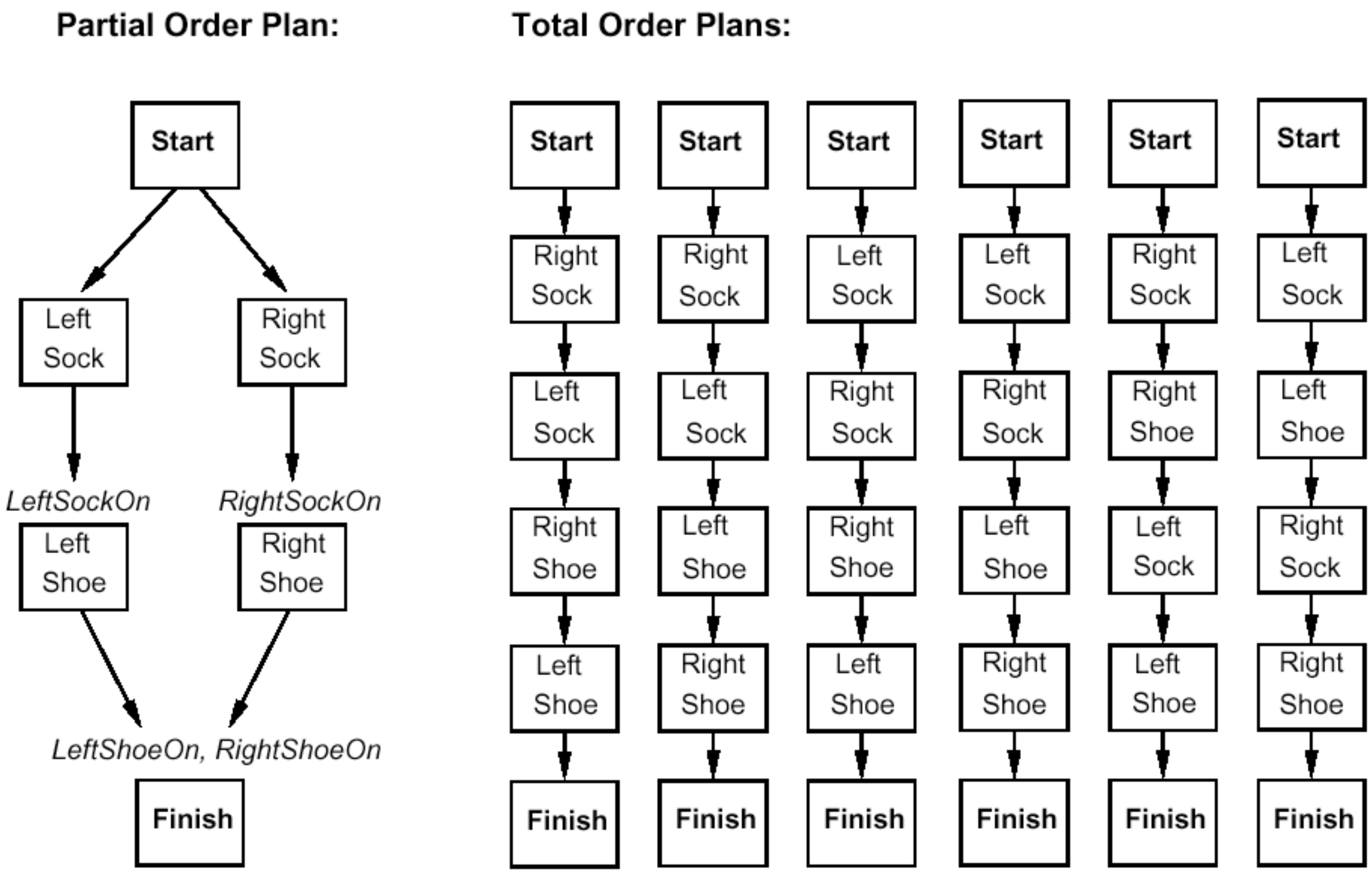
*Op* (Action: *buy(x)*,
Precond: *At(store)* $\wedge$ *Sells(store,x)*,
Effect: *Have(x)*)

there, here, x, store sind Variablen.



Aus einer KI-Vorlesung von Bernhard Nebel

# Plan = Lineare Sequenz von Aktionen?

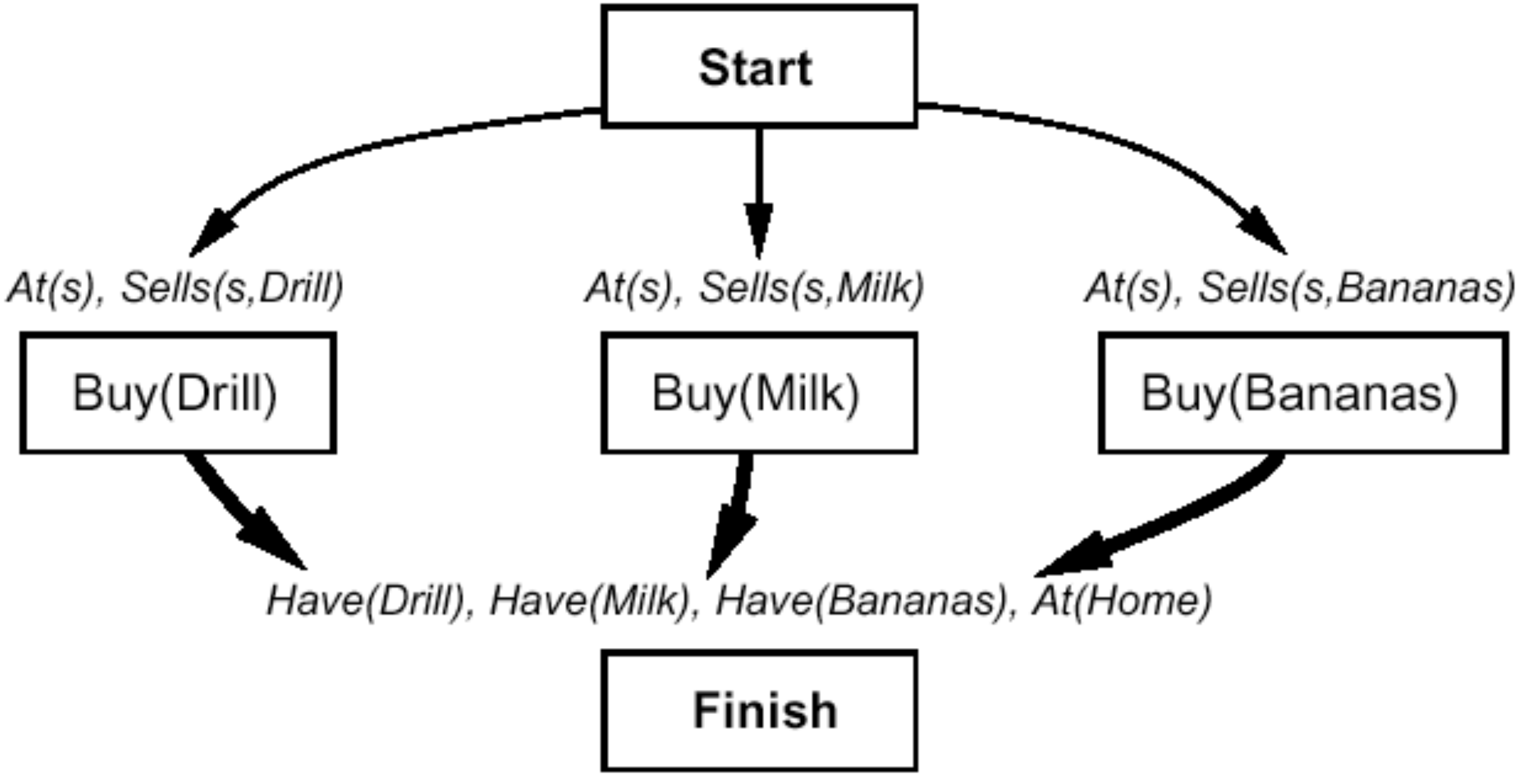





Aus einer KI-Vorlesung von Bernhard Nebel

# Planen im Planraum: Planverfeinerung (1)

- Rückwärtsplanung

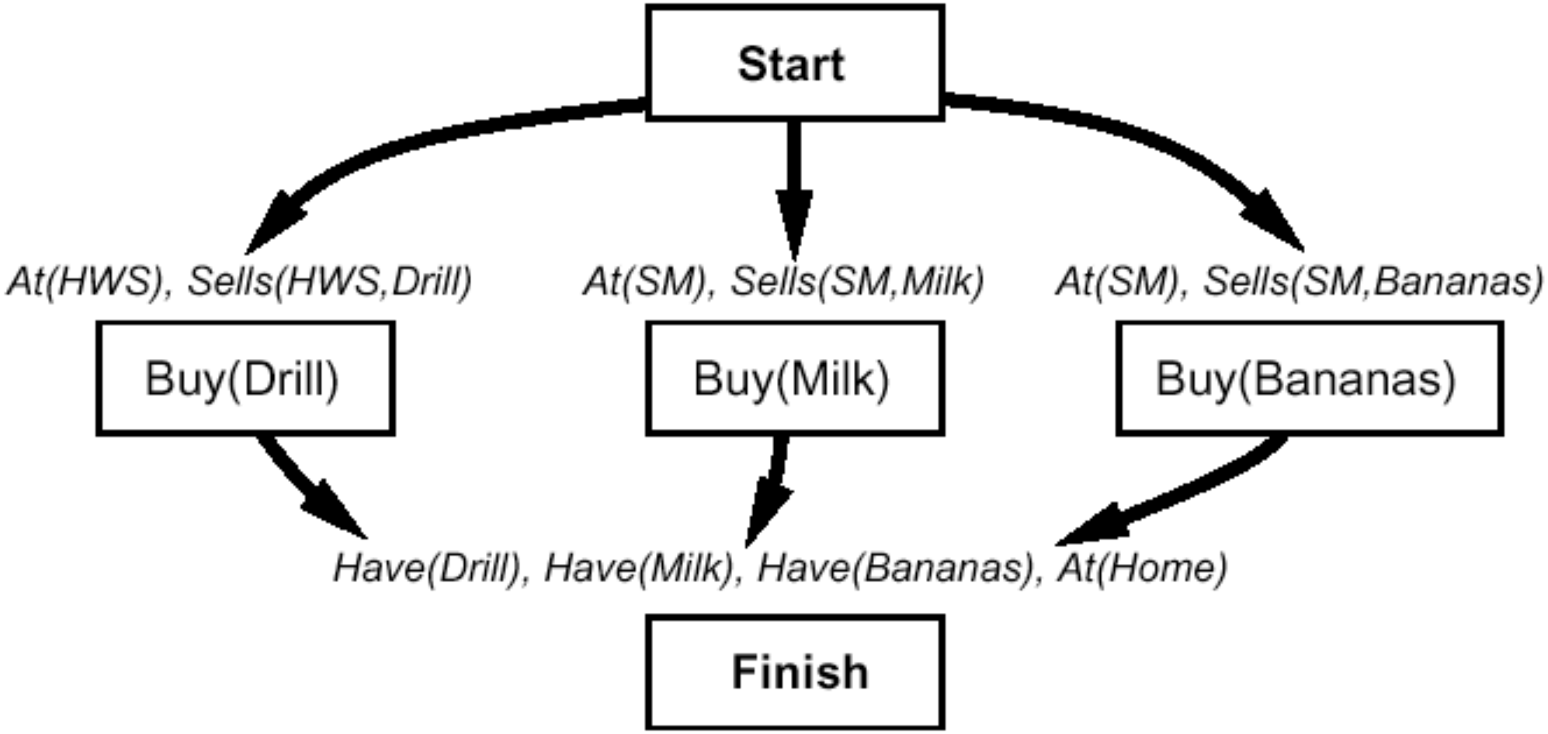



Dünne Pfeile $= \prec$
Fette Pfeile $=$ Kausale Beziehungen $+ \prec$



Aus einer KI-Vorlesung von Bernhard Nebel

# Planverfeinerung (2)



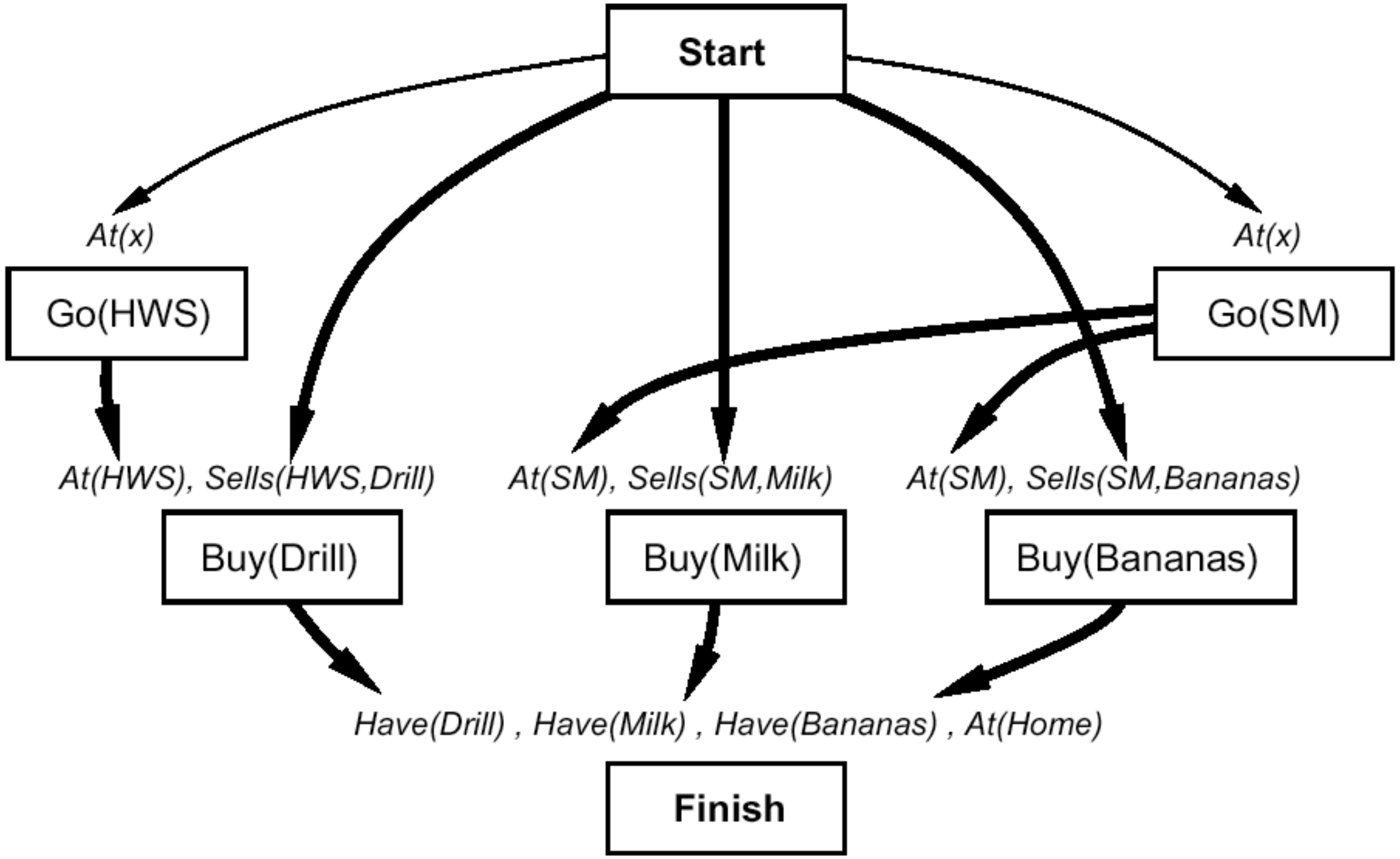

...nach Instantiierung der Variablen



Aus einer KI-Vorlesung von Bernhard Nebel

# Planverfeinerung (3)

- Gehe zum richtigen Geschäft …

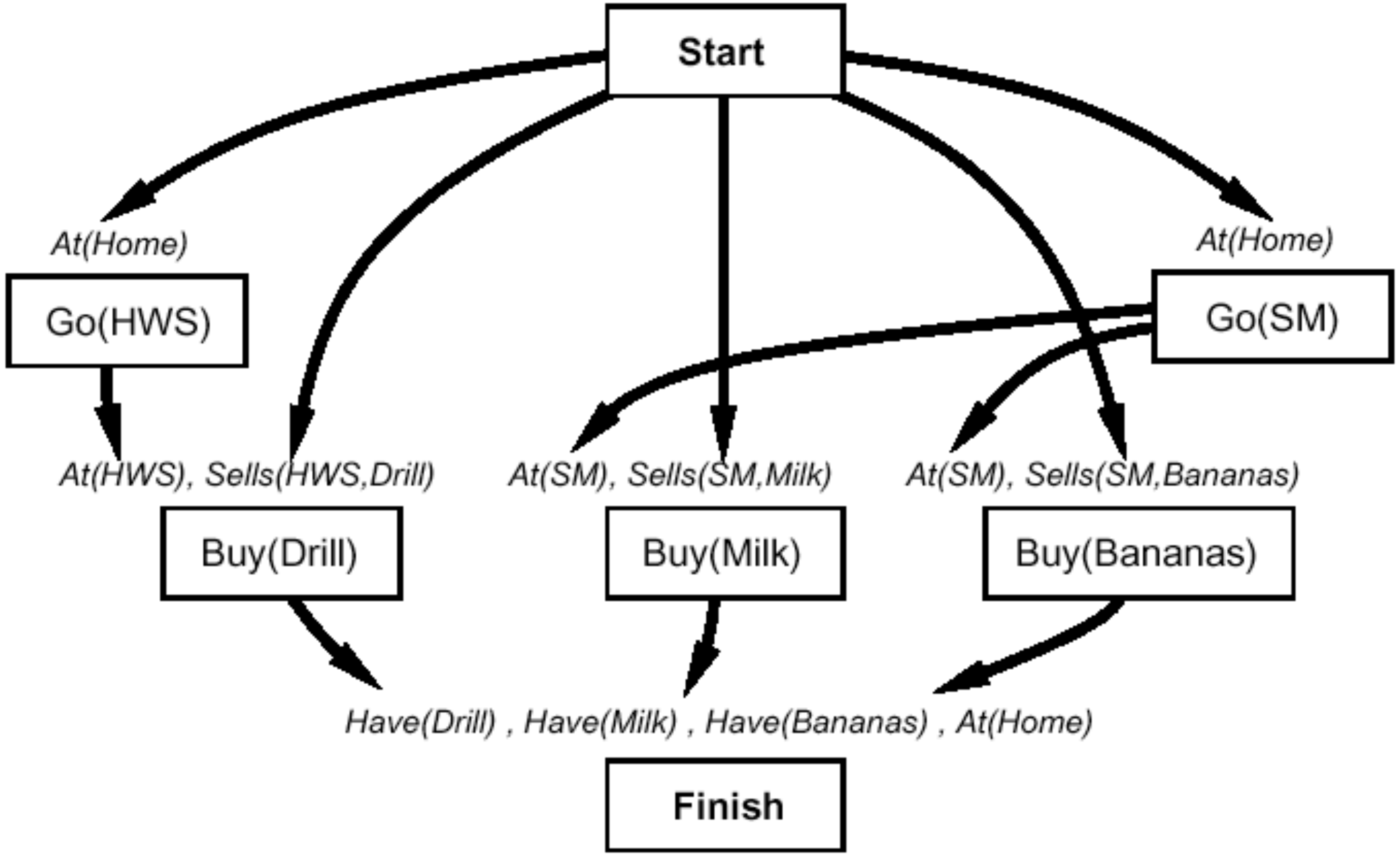





Aus einer KI-Vorlesung von Bernhard Nebel

# Planverfeinerung (3)

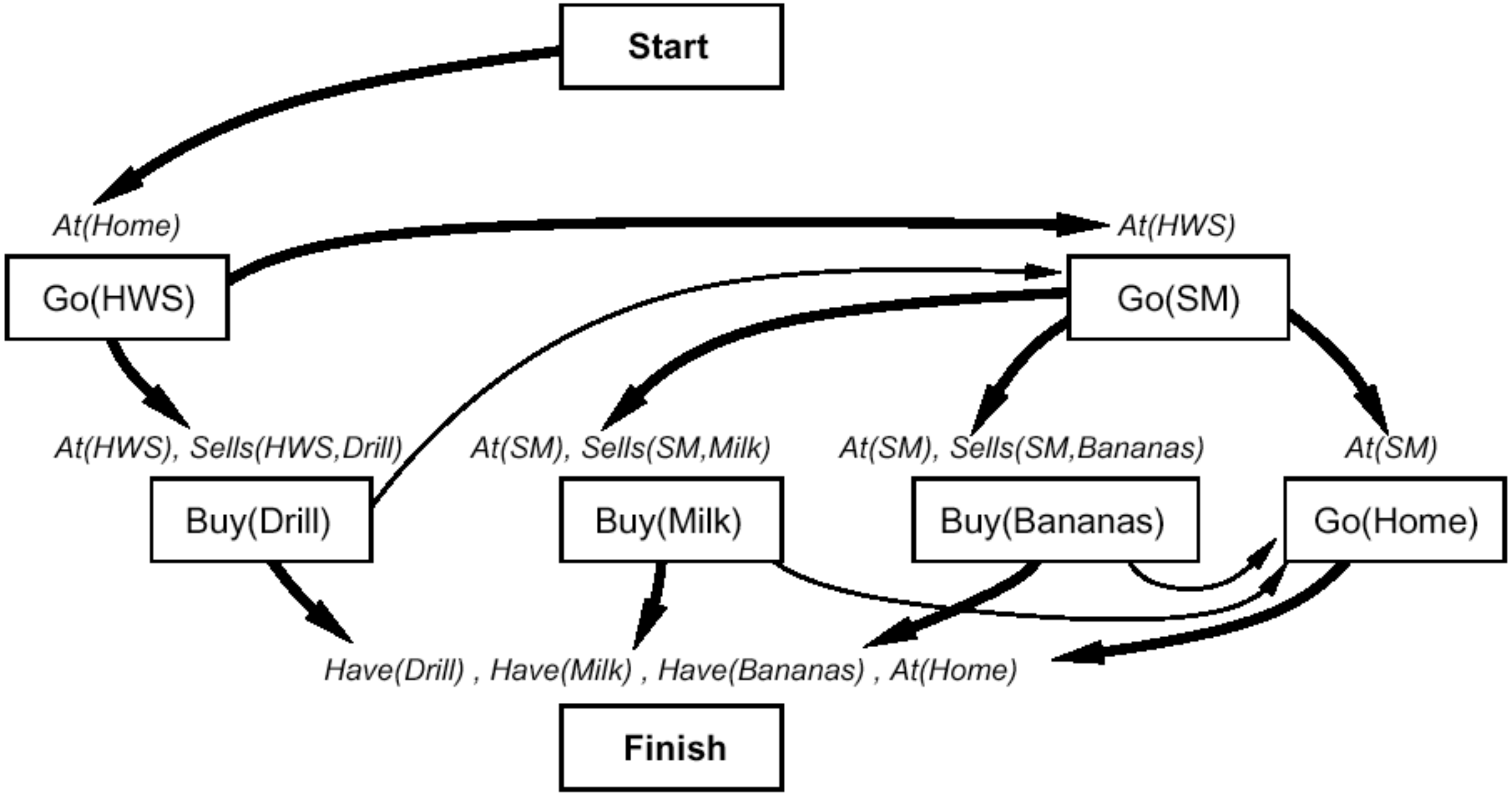
- ... sofern man nicht am richtigen Ort ist.






Aus einer KI-Vorlesung von Bernhard Nebel

# Entscheide ggf. die Reihenfolge …

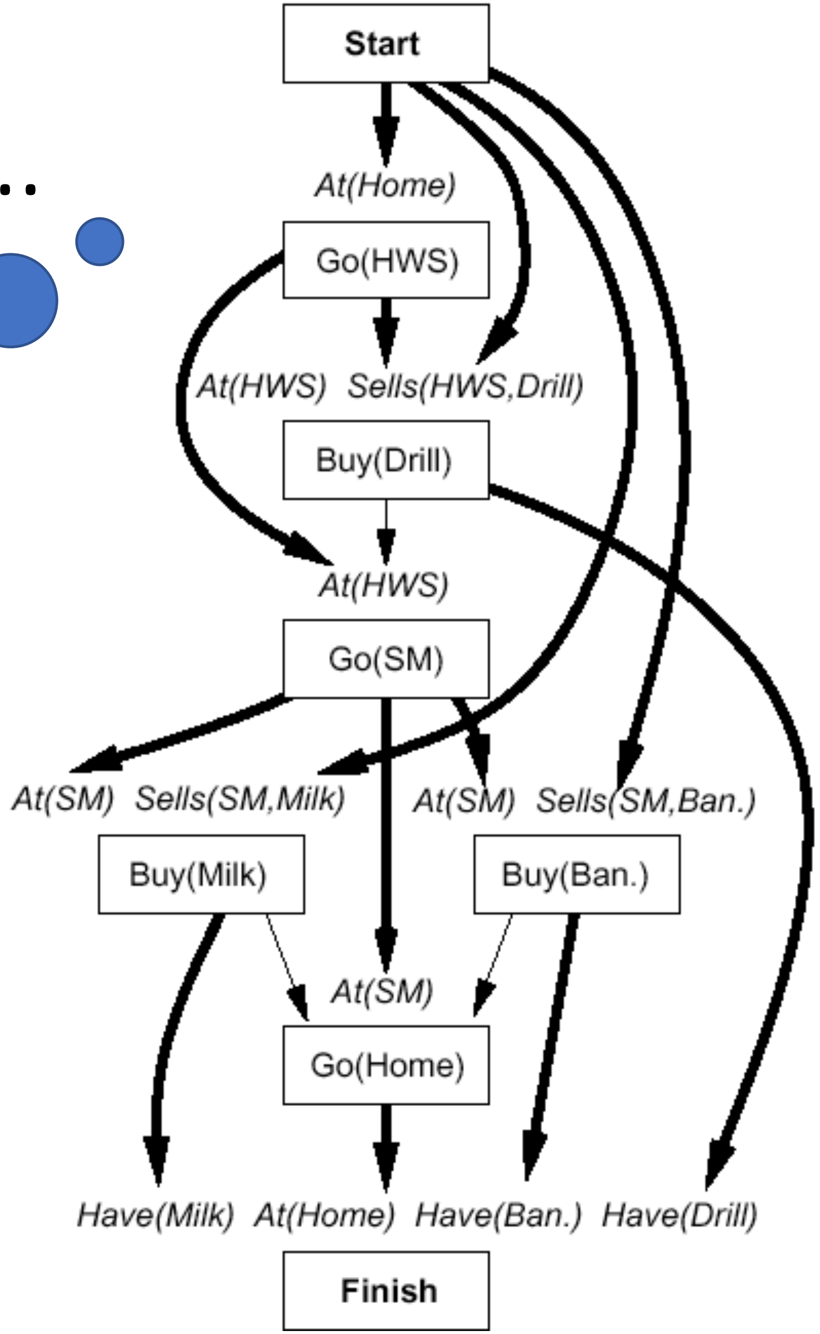





Aus einer KI-Vorlesung von Bernhard Nebel

# Die vollständige Lösung...

Partial Order Planning (UCPOP: Universal Conditional Partial Order Planner)




Penberthy, J. S. and Weld, D., ``UCPOP: A Sound, Complete, Partial-Order Planner for ADL,'' Third International Conference on Knowledge Representation and Reasoning (KR-92), Cambridge, MA, October **1992**.



Aus einer KI-Vorlesung von Bernhard Nebel

Lübecker Nachrichten

Lübecker General-Anzeiger

2016

Facebook, CMU
Offline- + Online-Lernen
Spielerverhalten nur online verfügbar

# Artificial intelligence has now pretty much conquered poker

Called Pluribus, the AI is a formidable opponent at six-player no-limit Texas Hold’em

# DAILY NEWS

World - Business - Finance - Lifestyle - Travel - Sport - Weather

2022

THE WORLDS BEST SELLING NATIONAL NEWSPAPER

Issue: 240104

Monday 5th June

First Edition

## Artificial Intelligence
## Beats 8 World Champions at Bridge

Using machine learning techniques, the small French startup company NukkAI....

https://nukk.ai

# Where did we go wrong?

- Humans are intelligent to the extent that our actions can be expected to achieve our objectives
- Machines are intelligent to the extent that their actions can be expected to achieve their objectives
  - Give them objectives to optimize (cf control theory, economics, operations research, statistics)
- We don’t want machines that are intelligent in this sense
- Machines are ***beneficial*** to the extent that ***their*** actions can be expected to achieve ***our*** objectives
- We need machines to be ***provably beneficial***

Maschinen unsicher bzgl. “unserer” Präferenzen: Hinterfragen der eigenen Aufgabenbeschreibung

Arguments taken from Stuart Russell’s Presentations on Provably Beneficial AI

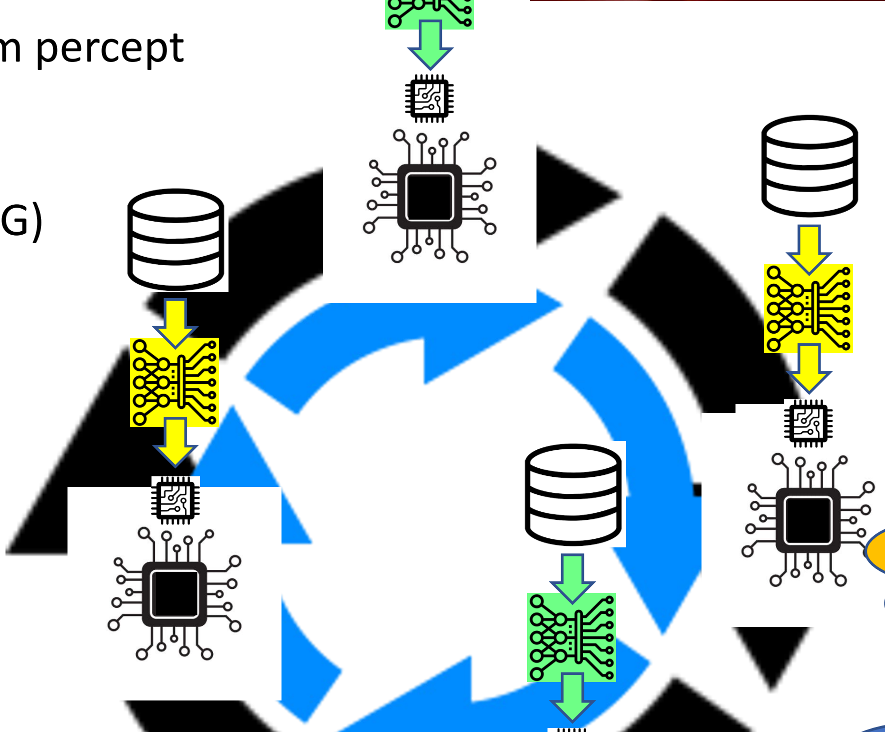
# Yet some more detail

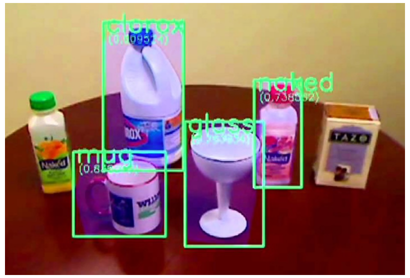
Percepts

Feedback

Mapping from percept to state of environment (depends on G)

Your performance might be improved






Idea of a utility function

Is the task G in the current state of the environment still correctly chosen?
Should I have a new task G'?

My utility will be zero then. How to prevent being switched off?

Hm: Possibly I'll be switched off

Maximize utility (possibly with sequence of actions)

What is the best action in the current state to fulfill the task?
Strategy for determining action

Action
Explanation

Return calculated action (prepare to answer the Why question)

# Off-Switch Problem

Representative for dealing with unwanted behavior

Example: “Fetch some coffee”

Very much underspecified!

The instruction suggests having coffee would have higher value than expected a priori, ceteris paribus

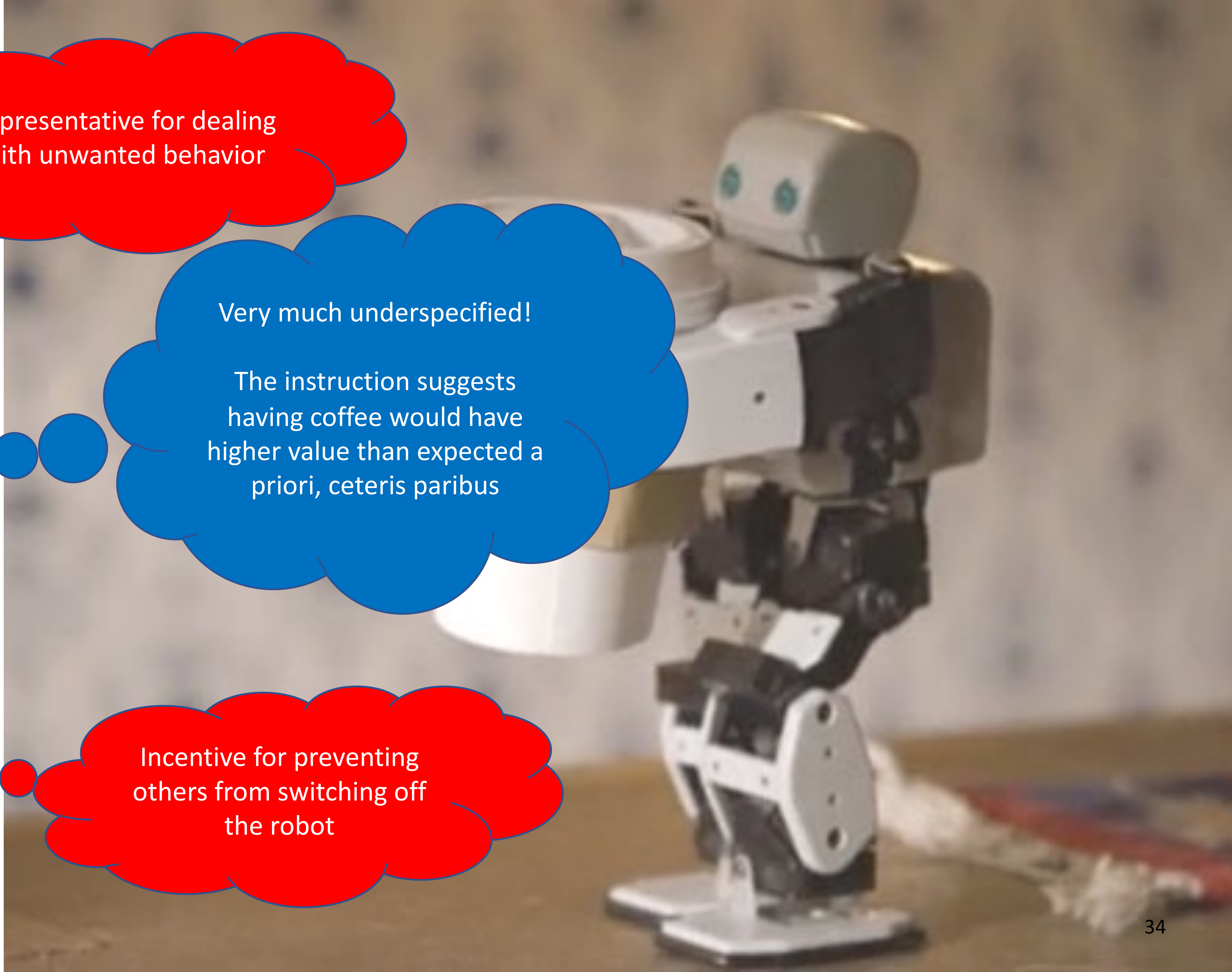
Agents get better at maximizing the built-in utility function

What’s bad about better AI?

Can we switch off the agent if it “does not work as expected”?

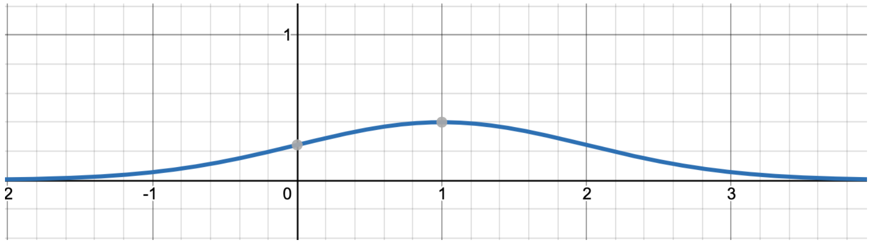
“Can’t fetch coffee if I am dead.”

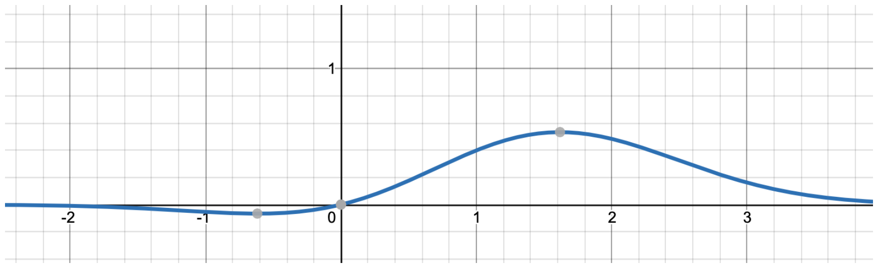
Incentive for preventing others from switching off the robot

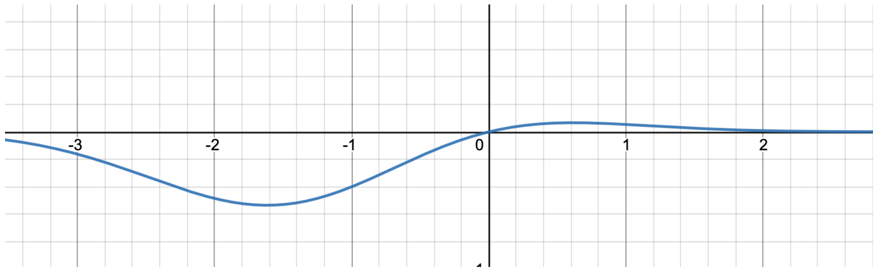


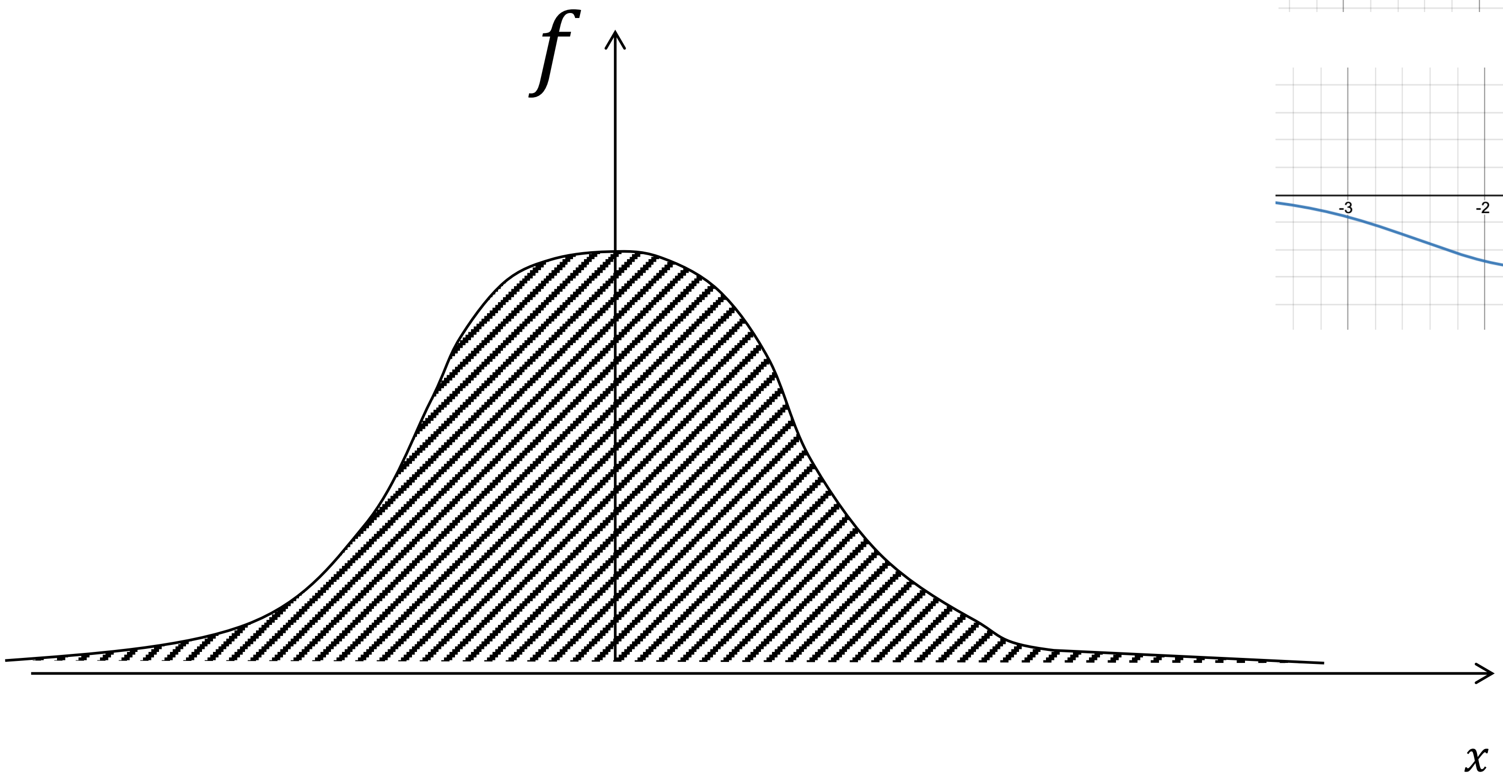
# Before/after bringing coffee…

- Expected utility $E[U(state)] = E[x] = \int x\, f(x)dx$
- Change the utility function



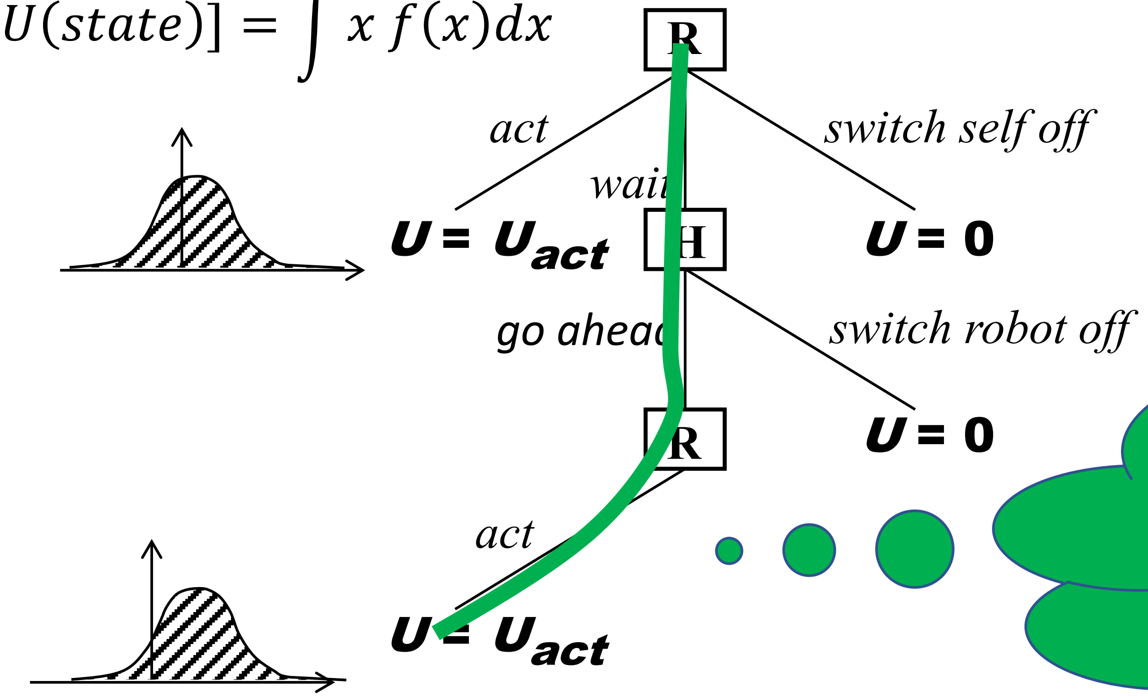



# The off-switch problem

- A robot, given an objective, has an incentive to disable its own off-switch
- Claim: A robot with appropriate uncertainty about objective won't behave this way
- Example: Planning for the best action (sequence)

$$E[U(state)] = \int x\, f(x)dx$$




Tiny example for "ethical behavior"

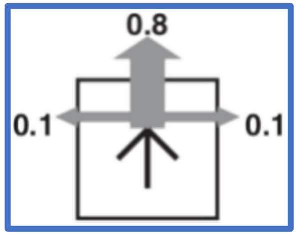
Theorem:
This way robot has a positive incentive to allow itself to be switched off

Do not just talk about ethics!
Verify it!

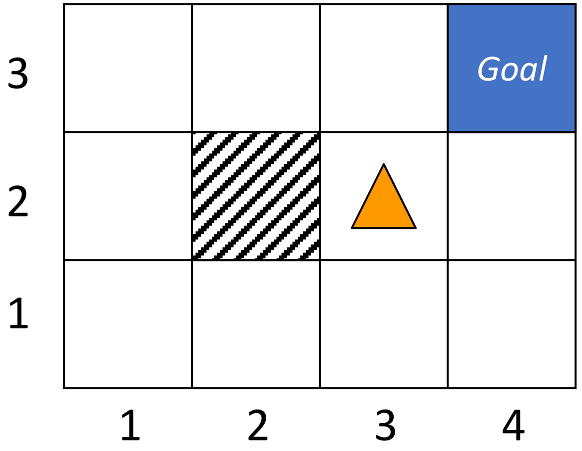
Example taken from Stuart Russell's Presentations on Provably Beneficial AI

# Sequence of Actions: Nondeterminism

- In each state, the possible actions are U, D, R, and L; the transition model for each action is (pictured):



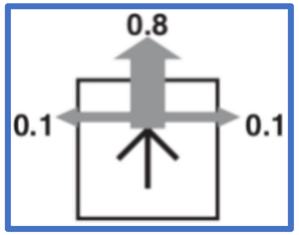

- Current position: [3,2]
- Planned sequence of actions: (U, R)



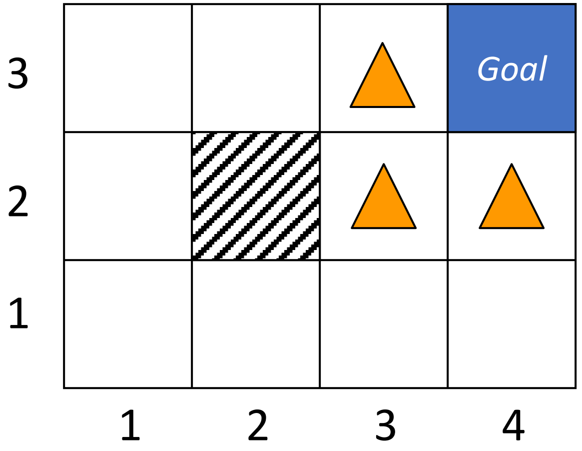

[3,2]

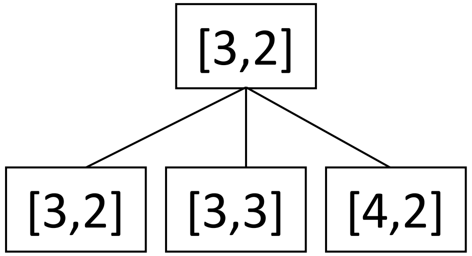
# Sequence of Actions

- In each state, the possible actions are U, D, R, and L; the transition model for each action is (pictured):



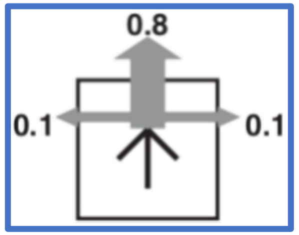

- Current position: [3,2]
- Planned sequence of actions: (U, R)
  - U is executed



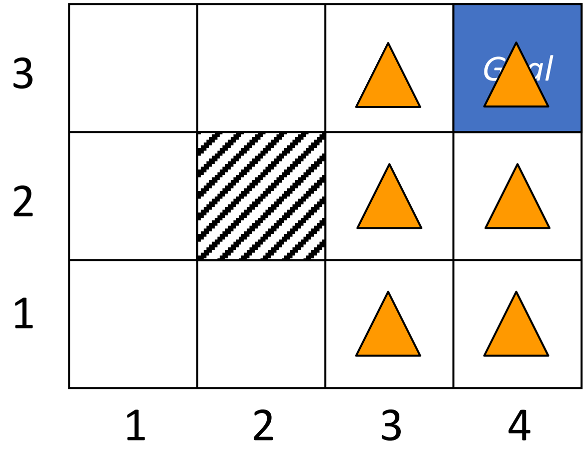



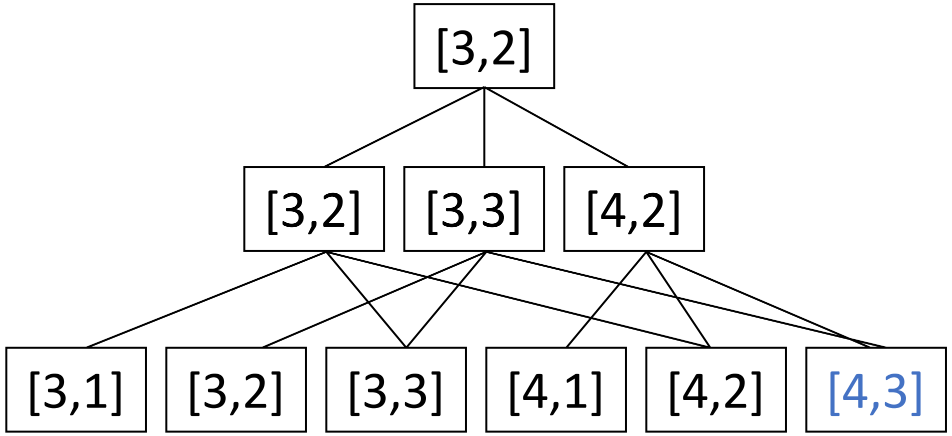
# Sequence of Actions

- In each state, the possible actions are U, D, R, and L; the transition model for each action is (pictured):



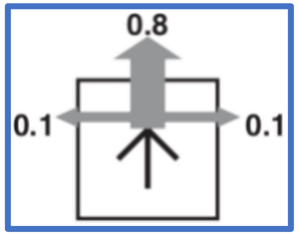

- Current position: [3,2]
- Planned sequence of actions: (U, R)
  - U has been executed
  - R is executed



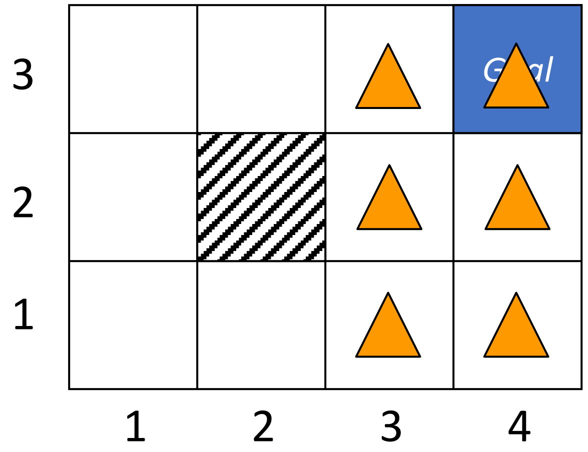
# Histories

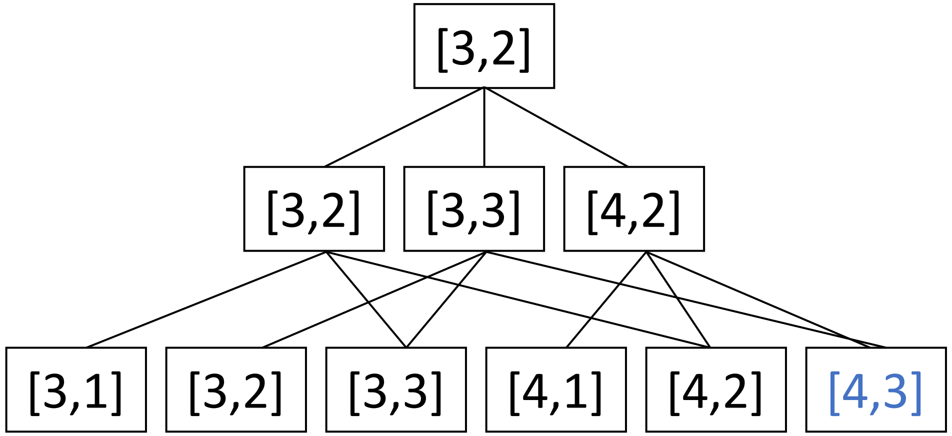
- In each state, the possible actions are U, D, R, and L; the transition model for each action is (pictured):




- Current position: [3,2]
- Planned sequence of actions: (U, R)
  - U has been executed
  - R is executed

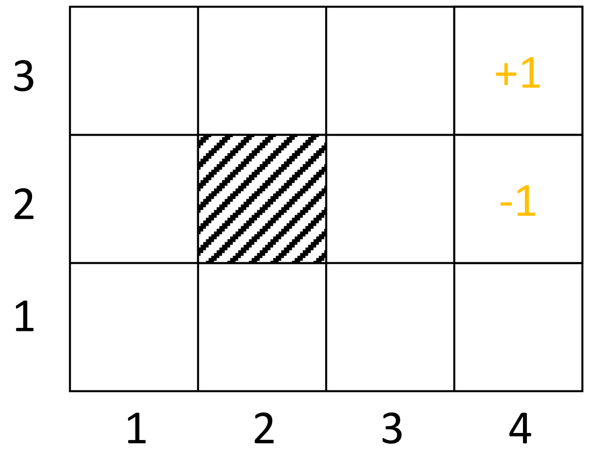
9 possible sequences of states, called histories, and 6 possible final states



# Utility Function

- [4,3] : power supply
- [4,2] : sand area the robot cannot escape (stops the run)
- Goal: robot needs to recharge its batteries
- [4,3] and [4,2] are terminal states
- In this example, we define the utility of a history by
  - The utility of the last state (+1 or –1) minus $0.04 \cdot n$
    - $n$ is the number of moves
    - I.e., each move costs $0.04$, which provides an incentive to reach the goal fast



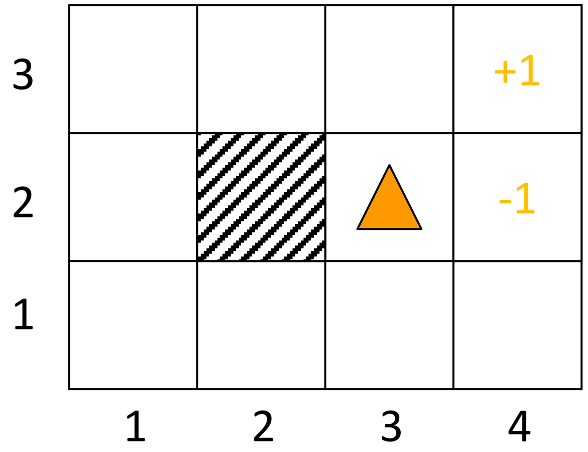
# Utility of an Action Sequence

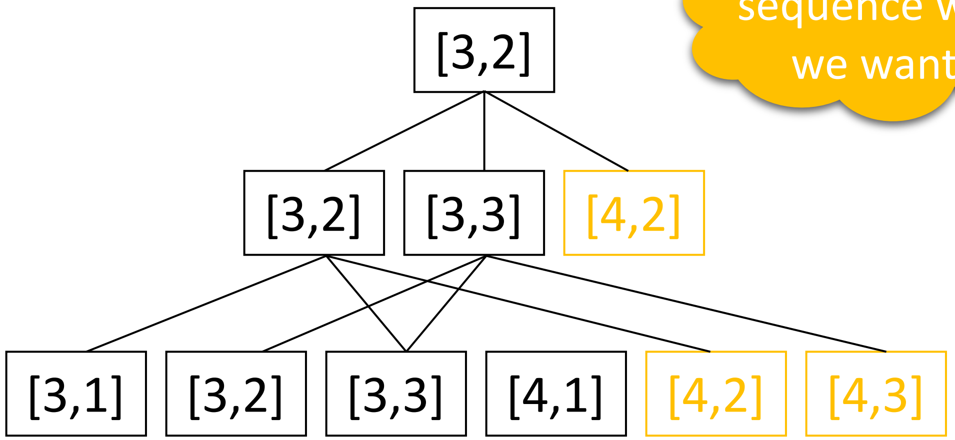
- Consider the action sequence $\boldsymbol{a} = (\text{U,R})$ from [3,2]
- A run produces one of 7 possible histories, each with a probability
- Utility of the sequence is the expected utility of histories $h$:

$$U(\boldsymbol{a}) = \sum_h U_h P(h)$$

- Optimal sequence = the one with maximum utility

Is the optimal sequence what we want?

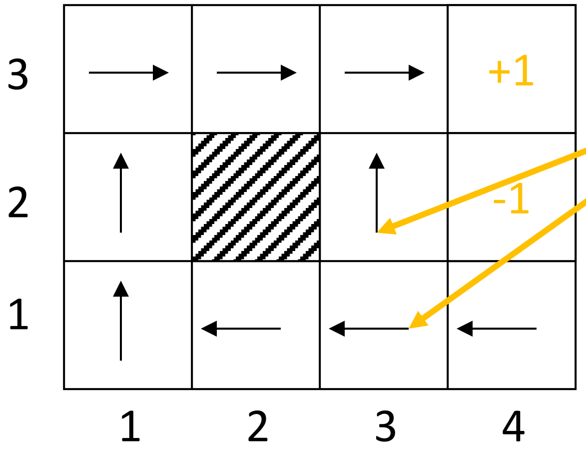


# Policy (Reactive/Closed-loop Strategy)

- Policy $\pi$
  - *Complete* mapping from states to actions
- Optimal policy $\pi^*$
  - Always yields a history (ending at terminal state) with maximum expected utility

```
Act()
    repeat
        s ← sensed state
        if s is terminal then
            exit
        a ← π(s)
        perform a
```




Note that [3,2] is a “dangerous” state that the optimal policy tries to avoid

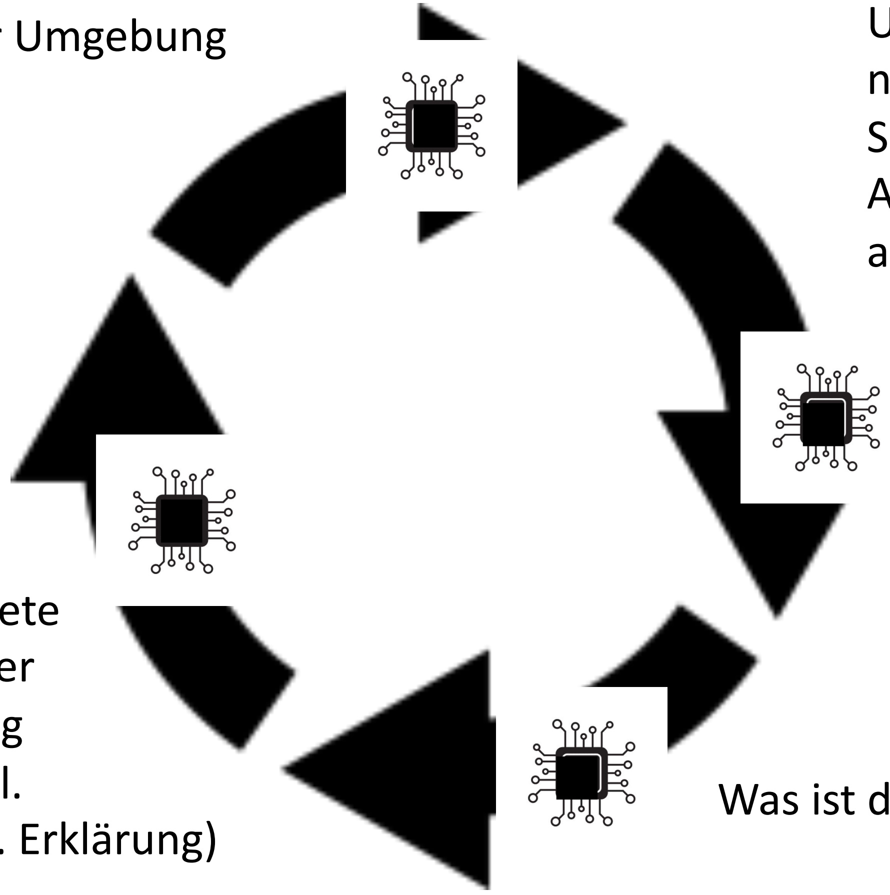
How to compute $\pi^*$?
Solving a Markov Decision Process

# Planen oder Policy-Anwendung

Perzepte

Vom Perzept
zum Zustand
der Umgebung

Ist meine
aufgabenbeschreibung
bei aktuellem Zustand der
Umgebung
noch richtig gewählt?
Sollte ich einen neue
Ausgabenbeschreibung
annehmen?




Führe berechnete
Aktion aus (oder
teile Umgebung
Aktion mit) inkl.
Zusatzinfo (z.B. Erklärung)

Was ist die beste Aktion im aktuellen Zustand
bei gegebenem Ziel?
Handlungsbestimmungsstrategie

# Perseverence: Mars Rover

# Mars Rover: Autonomy is Necessary

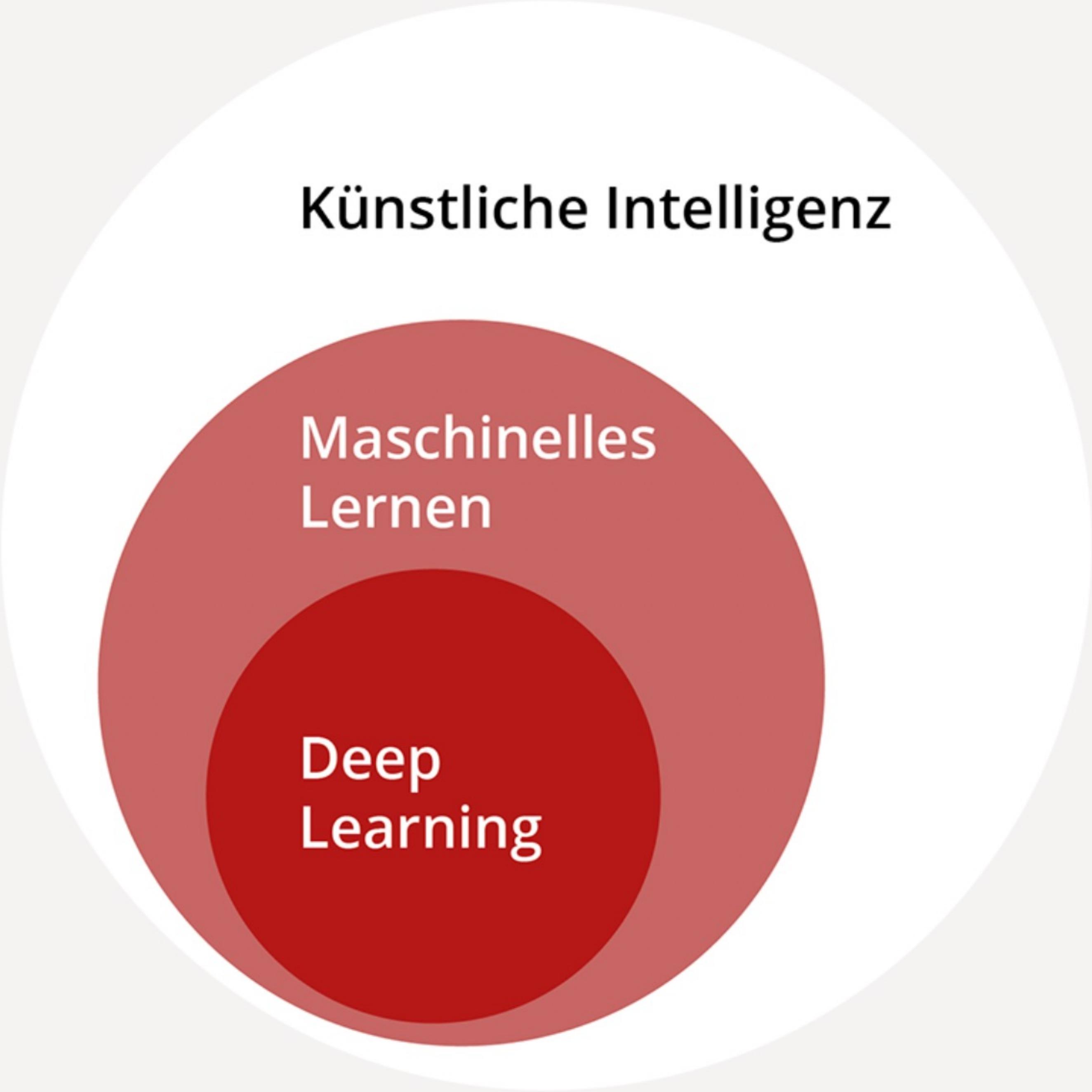
No remote control

- Reinforcement learning necessary (→ Systems must act!)
  - Development time
  - Deployment time
- Data interpretation → Model assumptions required
- Model assumptions possibly need to be revised
- Decision: Use policy or employ planning (or counterfactual reasoning)
- Task description ("goal") still valid?

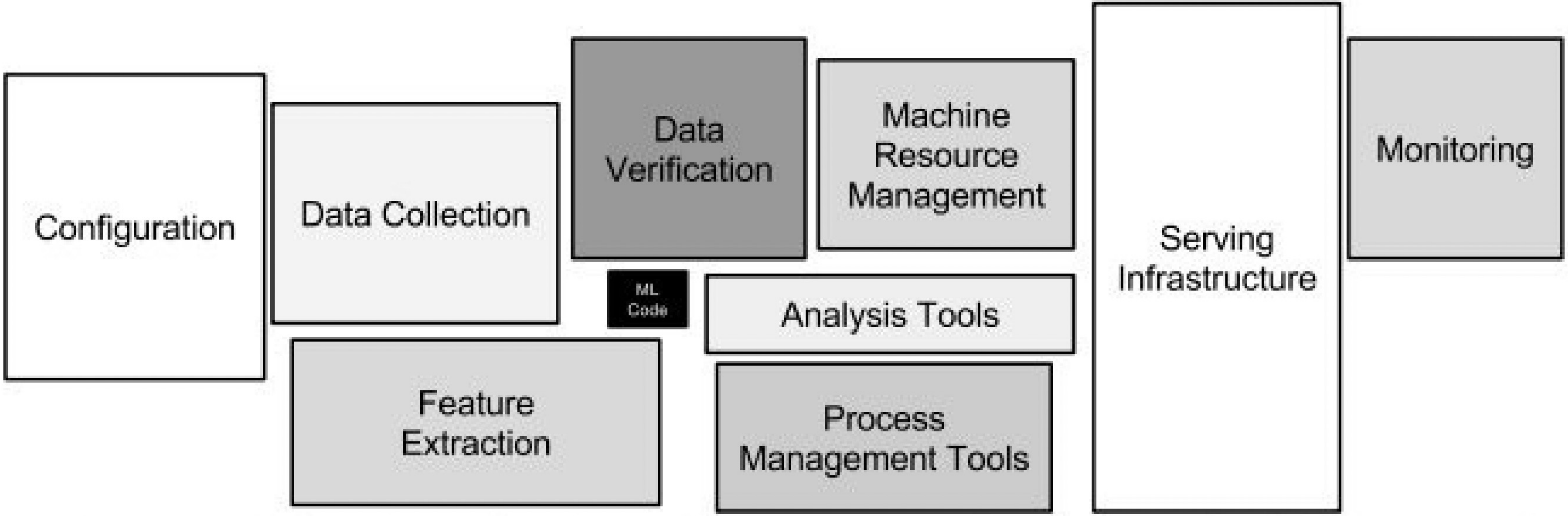
By no means are intelligent systems simple to develop:

We need to be able to understand intelligent systems from an engineering perspective

Intelligent systems must be provably beneficial

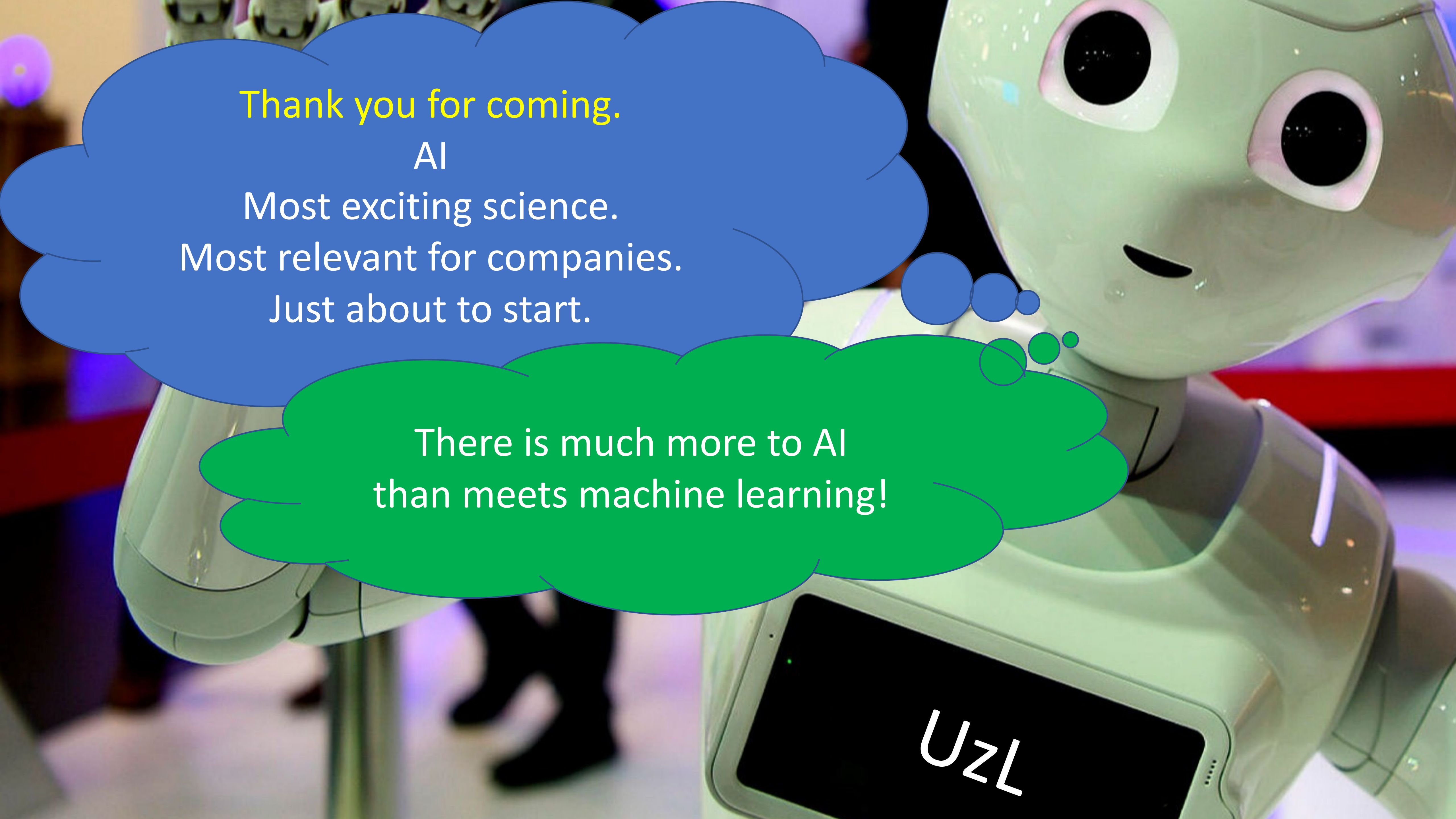

Künstliche Intelligenz
Maschinelles Lernen
Deep Learning

# Künstliche Intelligenz



# AI in the Small: Nutzung von Teilen des Agenten

# AI in the Large: Instruct systems to autonomously act in a beneficial way  for a group of humans


Thank you for coming.
AI
Most exciting science.
Most relevant for companies.
Just about to start.
There is much more to AI
than meets machine learning!
UzL