Reasoning and Decision-Making under Uncertainty

1. Session: Introduction

Prof. Dr.-Ing. Stefan Kopp
Center of Excellence „Cognitive Interaction Technology“
AG Sociable Agents

CITEC

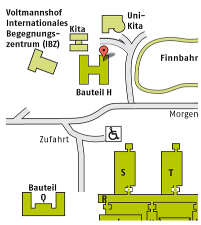
Sociable Agents

Organisatorisches

Sociable Agents
http://www.techfak.uni-bielefeld.de/ags/soa/

Prof. Dr.-Ing. Stefan Kopp

- skopp@techfak.uni-bielefeld.de
- Sprechstunde: Fr 13-14, H1-115
- Tel: (106-)12144



Semesterapparat: Universitätsbibliothek, FB Informatik

Web: www.techfak.uni-bielefeld.de/~skopp/Lehre/Uncertain_SS13

Organisatorisches

Voraussetzungen:

- Ansätze und Methoden der Künstlichen Intelligenz
- Mathematische Grundlagen der Wahrscheinlichkeitstheorie
- Algorithmen & Datenstrukturen

Leistungspunkte: 6 LPs für Vorlesung und Übung

- Teilnahme an der VL
- erfolgreiches Bearbeiten der Übungsaufgaben
- Bestehen der Abschlussprüfung/Klausur (→ benotete EL)

Modul „Vertiefung Künstliche Intelligenz“ = 10 LP

- 4 LP aus weiterem Seminar

Übungen

- Sebastian Ptock (sptock@techfak.uni-bielefeld.de), Raum H1-115a
- Belegnummer 392102 (bitte alle in den eKVV-Verteiler eintragen!)
- Web: http://www.techfak.uni-bielefeld.de/~sptock/tutki/index.html
- Termin: Mi, 16-18, in H1-111a (nicht C6-200!)
- Start am 17. April, ab 24. April zweiwöchentlich

Übungen

Praktische Programmier-Übungen (in Python) zu ausgewählten Modellen und Algorithmen aus der Vorlesung

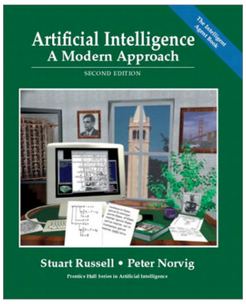
Inhalte:

- Einführung in Python und Numpy (Termin 1 & 2)
- Implementierung eines Reasoning-Systems mittels Bayes-Netzen und Inferenzalgorithmen
- Implementierung eines Decision-Making-Systems mittels Markov-Entscheidungsprozessen

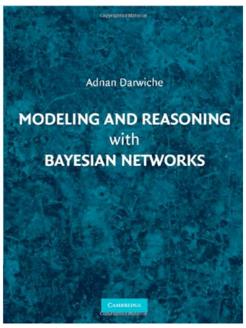
Leistungsanforderung:

- Bearbeitung und fristgerechte Abgabe der Übungsaufgaben

Literatur



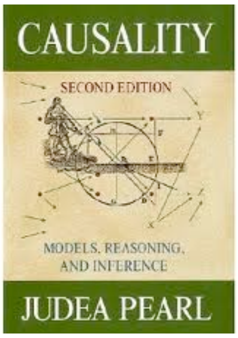
Russell & Norvig: Artificial Intelligence: A Modern Approach. Prentice Hall, 2nd Edition, 2003



Darwiche: Modeling and Reasoning with Bayesian Networks. Cambridge Univ. Press, 2009

Literatur

Judea Pearl, Probabilistic reasoning in intelligent systems, Morgan Kaufmann, 1989



J. Pearl: Causality - Models, Reasoning and Inference (2nd edition). Cambridge Univ. Press, 2009

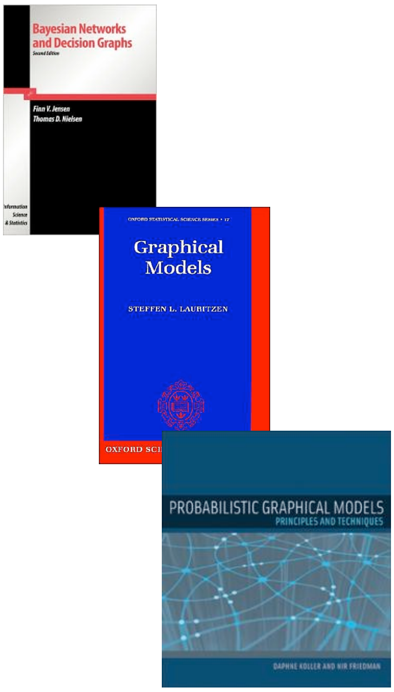
Literatur

Daniel Barber, Bayesian Reasoning and Machine Learning, Cambridge Univ. Press, 2012.

Kostenfreie Online-Version!

http://www.cs.ucl.ac.uk/staff/d.barber/brml/

Weiterführende Literatur



Finn V. Jensen, Bayesian networks and decision graphs, , Springer, 2001

Steffen L. Lauritzen, Graphical models, Oxford, 2002

Daphne Koller & Nir Friedman, Probabilistic Graphical Models, MIT Press, 2009

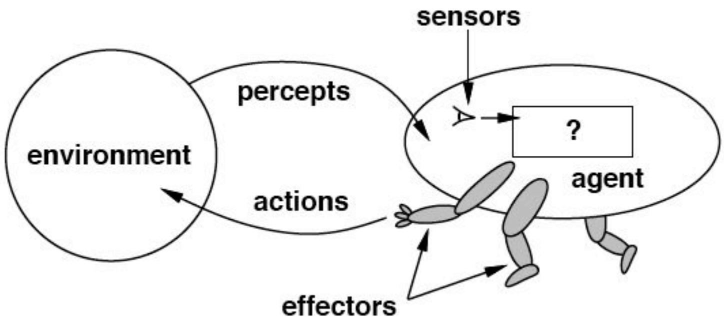
Introduction

Warming-up exercise

<u>Answer these questions:</u>

- How does a classical A.I. system work (in principle)?
- What kinds of uncertainties might it face?
- What may they arise from?

Introduction



Basic idea:

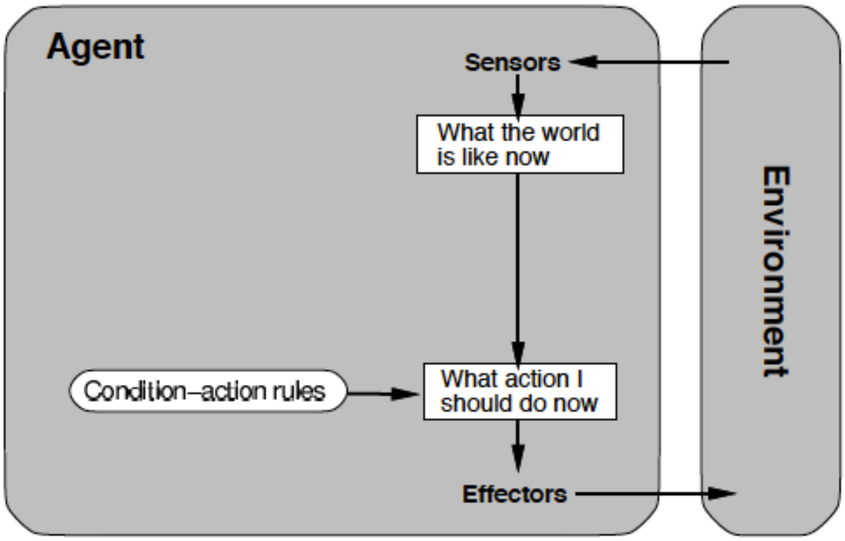
Agents interacting autonomously with an environment through sensors and effectors (e.g., Russell & Norvig 1995)

Intelligent agent

From the outside:

- for each possible percept sequence, does whatever action it expects to maximize its performance measure (rational agent)
- on the basis of the evidence provided by its percept sequence and whatever built-in knowledge and preferences it has
- based on some form of reasoning or planning that involves possible outcomes of actions or action sequences

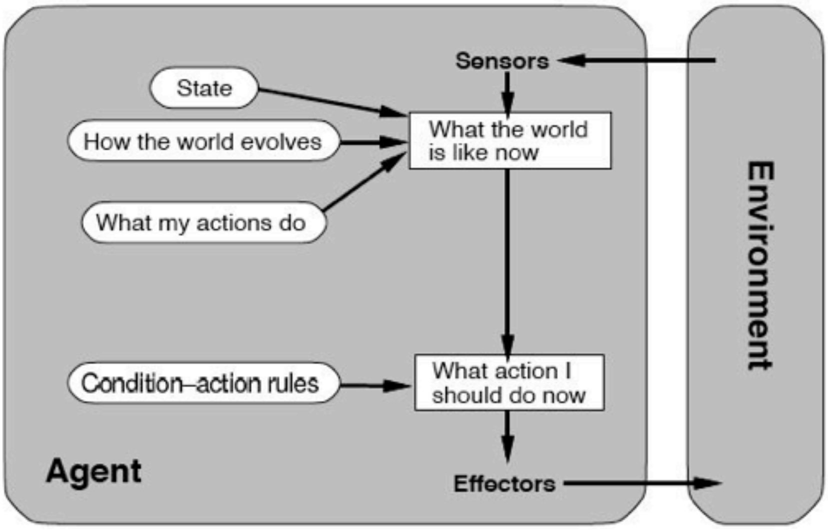
Intelligent agent



Simple reflex agent

(Russell & Norvig)

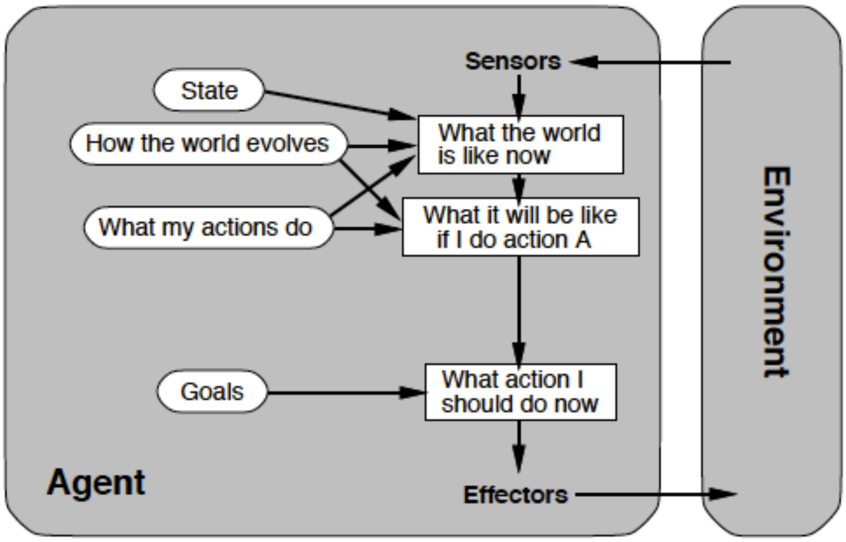
Intelligent agent



Reflex agent with internal state

(Russell & Norvig)

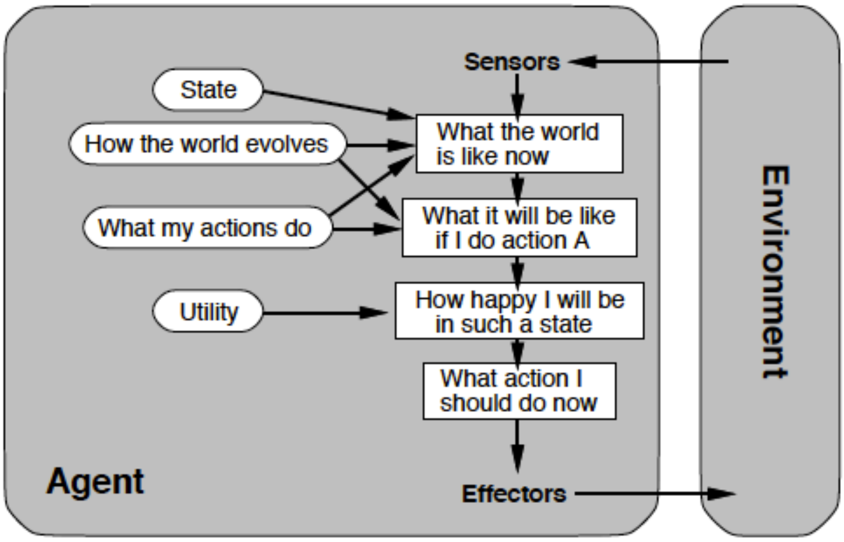
Intelligent agent



Goal-based agent

(Russell & Norvig)

Intelligent agent



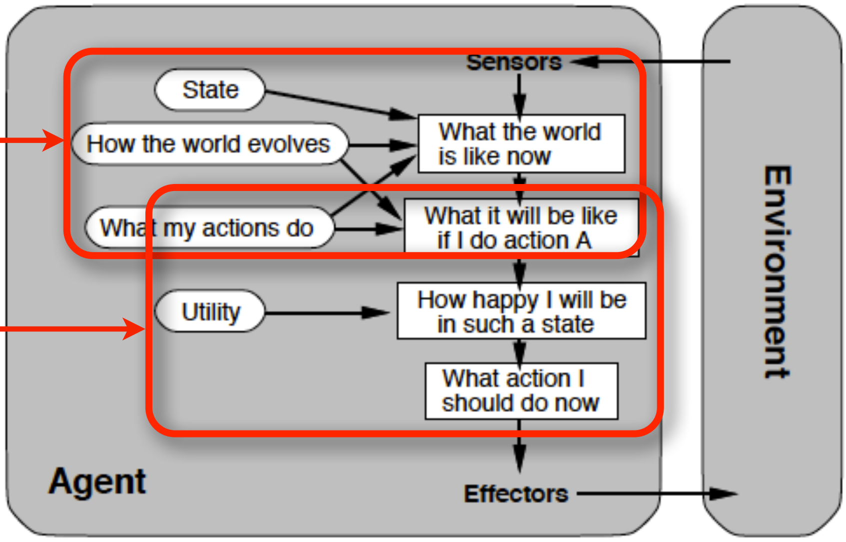
Utility-based agent

(Russell & Norvig)

Intelligent agent

Reasoning, inference

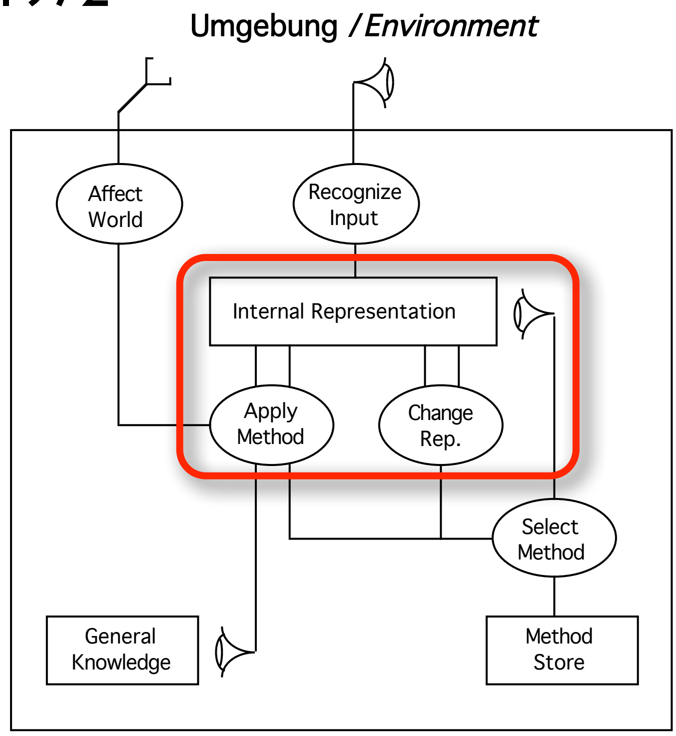
Decision-making, action selection



Intelligent agent - 1972

General Intelligent Agent

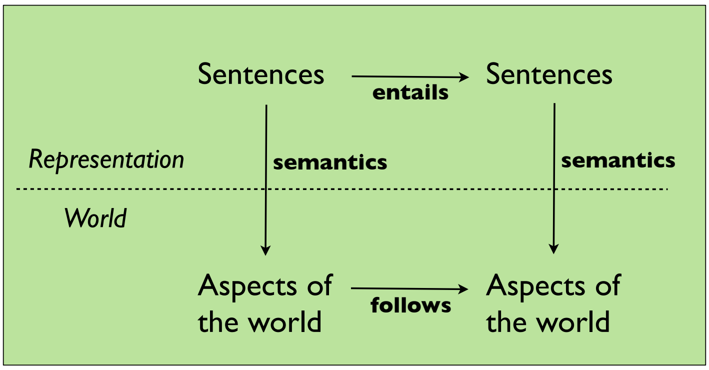
(Newell & Simon 1972)



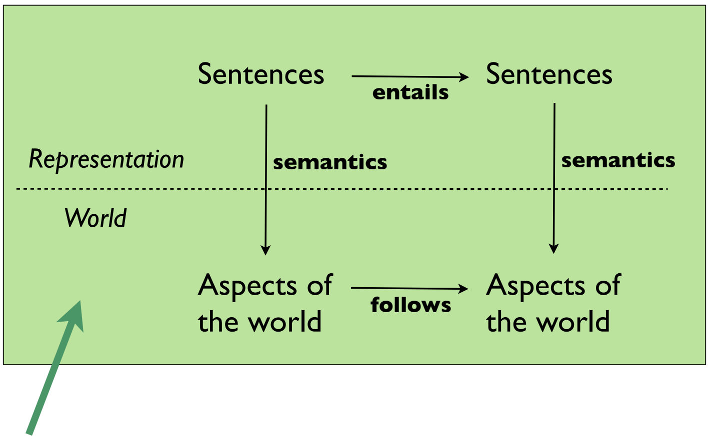
Intelligent agent

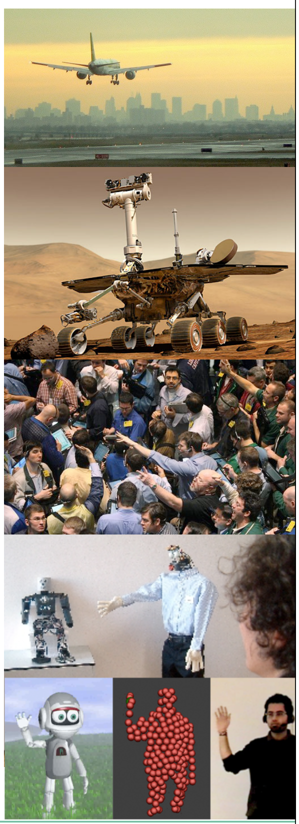
<u>Key principle</u>:

- internal representation of (parts of) the environment
- reasoning using an inference calculus
- decision-making based on preferences (goals) and search



Real-life domains





What makes many domains notoriously hard?

size, large or unknown complexity, highly dynamic, weakly predictable, limited observability, ...

Sources of uncertainty in classical reasoning

Incomplete knowlegde

- no knowledge of all causal relations, their antecedents or consequents
- precise information would be too complex
- need to make default assumptions or approximations

Conflicting information

- local conclusions may become invalid later, need to be retracted

Cumulation of uncertainties

- uncertainty about antecedents increases uncertainty of conclusion:
  Sunny [0.7] and (Sunny → Warm) [0.8] => Warm [?]
- uncertainty accumulates when chaining rules/inferences

Example: The doorbell problem

Logics-based formulation:

1. AtDoor(x) → Doorbell
2. Short-Circuit → Doorbell
3. Doorbell → Wake(John)
4. Light-Bedroom → Wake(John)
5. Extremely-Tired(John) → NOT Wake(John)

Given: Doorbell rang at 12 o‘clock midnight

- Can we say Wake(John) given Doorbell?
- Can we say AtDoor(X) given Doorbell?

Example: The doorbell problem

Can we say Wake(John) given Doorbell?

- Deductive reasoning:
  Doorbell → Wake(John), Doorbell ⇒ Wake(John)
- Locality:
  - ignores exceptions, e.g., Wake(John) less likely if he is so tired
  - ignores other reasons, e.g., Wake(John) more likely if also Light-Bedroom
- Detachement:
  - ignores validity of antecedent Doorbell, e.g., Wake(John) less likely when finding out that no one was at the door, or invalid when NOT Doorbell
  - ignores other possible reasons, e.g. Wake(John) more likely when finding out that both Doorbell AND Light-Bedroom, but *not* when both have the same underlying cause

Example: The doorbell problem

Can we say AtDoor(X) given Doorbell?

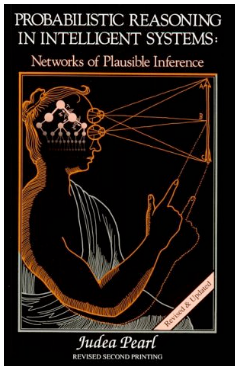
- Abductive reasoning:
  AtDoor(x) → Doorbell, Doorbell ⇒ AtDoor(x)
- Locality:
  - ignores other explanations in KB, e.g., Short-Circuit may also be true
  - ignores human-like causal reasoning, e.g., support for Short-Circuit reduces belief in AtDoor (one reason is sufficient, „explaining away“)

Human-like plausible reasoning requires bi-directional reasoning combining uncertain diagnostic and predictive inferences

Limits of classical logics-based reasoning

Modularity, i.e. locality and detachment of logics-based inference creates semantic deficiencies when trying to incorporate uncertainties

- improper handling of bi-directional inference
- difficulties in retracting conclusions
- improper treatment of correlated sources of evidence



More uncertainty in decision-making

Most domains are not observable, not static, and non-deterministic -- when taking decisions an agent can rarely...

- know the state of the world exactly and completely
- be sure that the state has not changed in the meantime
- be sure that its actions will bring about the desired effects

Different kinds of indeterminacy of an environment

- Bounded: actions can have unpredictable effects, but these can be enumerated in action description axioms
- Unbounded: preconditions and effects are too large to enumerate

Different kinds of decision problems

Single-state problem

- Environment is static, deterministic, and fully observable, i.e. can be encoded in one single state
- Agent knows exactly which state it is now in and will be in
- *Solution*: (sequence of) action that can be executed (open-loop)

Sensorless (conformant) problem

- Partial knowledge of states, but known actions
- Agent may have no idea which state it is in
- Each action may lead to one of several possible states
- *Solution* (if any): (sequence of) action that will do the job in any case

Different kinds of decision problems

Contingency problem

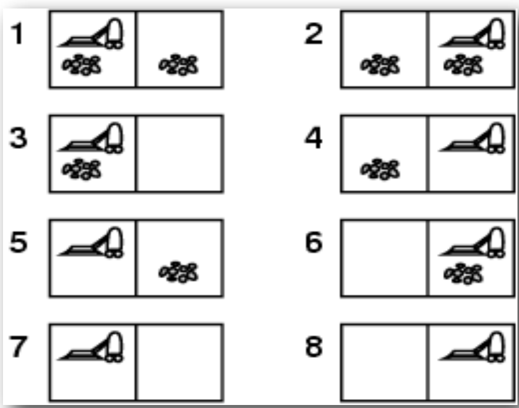
- Environment is non-deterministic, i.e. effects of actions are uncertain, or only partially observable
- Each percept provides new, but partial information after each action
- *Solution*: no fixed action sequence, plan for contingency, interleave monitoring, decision-making and execution (closed-loop)

Exploration problem

- Extreme case of contingency problem: environment and actions are fully unknown up-front
- *Solution*: unclear, agent must act to discover states and actions

Example: vacuum world

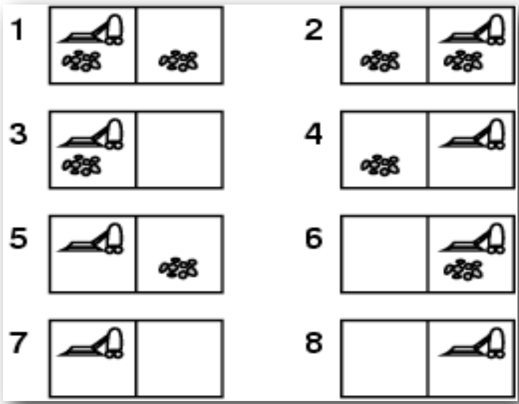
- Single-state, start in #5.
  Solution?



Task: Clean the room (#7 or #8)

Example: vacuum world

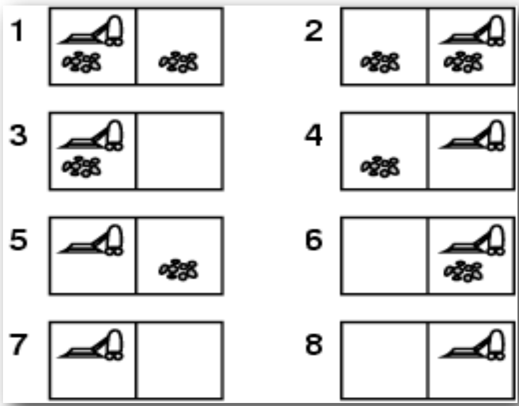
- Single-state, start in #5.
  Solution? *[Right, Suck]*



Task: Clean the room (#7 or #8)

Example: vacuum world

- Single-state, start in #5.
  Solution? *[Right, Suck]*
- Sensorless, start in one of {*1,2,3,4,5,6,7,8*}, e.g. *Right* goes to {*2,4,6,8*} and *[Right, Suck]* to {*4,8*}
  Solution?

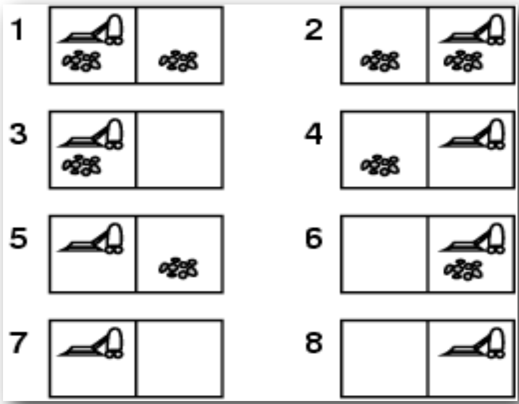


Task: Clean the room (#7 or #8)

Example: vacuum world

- Single-state, start in #5.
  Solution? *[Right, Suck]*

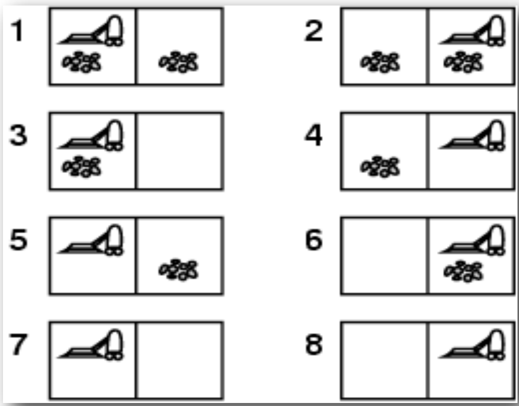
- Sensorless, start in one of {*1,2,3,4,5,6,7,8*}, e.g. *Right* goes to {*2,4,6,8*} and *[Right, Suck]* to {*4,8*}
  Solution?
  [Right,Suck,Left,Suck]
  Search in *sets* of states



Task: Clean the room (#7 or #8)

Example: vacuum world

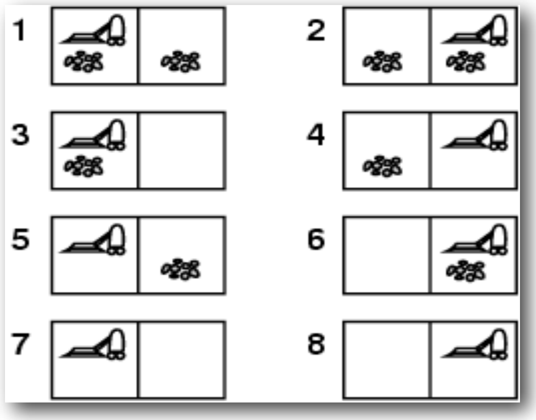
- Contingency problem
  - Non-deterministic: *Suck* may dirty a clean carpet
  - Partially observable: location? dirt at current location?
  - Percept: *[Left, Clean]*, i.e., start in #5 or #7 or ??
    Solution?



Task: Clean the room (#7 or #8)

Example: vacuum world

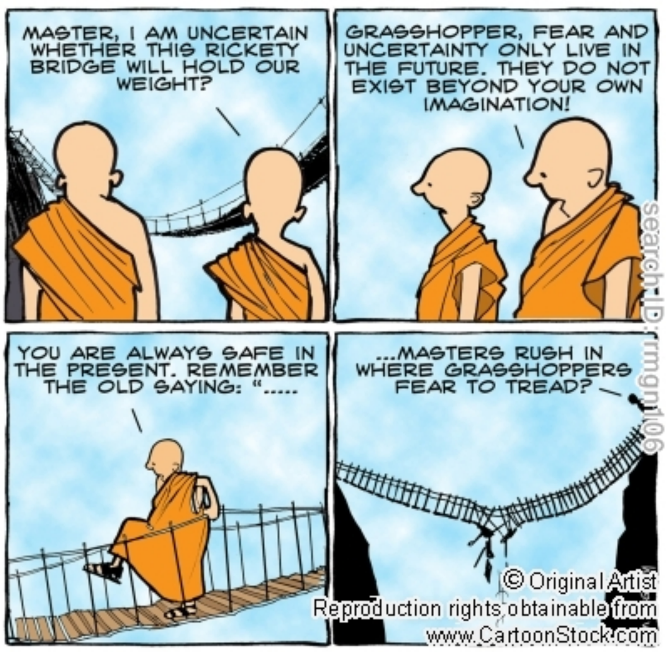
- Contingency problem
  - Non-deterministic: *Suck* may dirty a clean carpet
  - Partially observable: location? dirt at current location?
  - Percept: *[Left, Clean],* i.e., start in #5 or #7 or ??
    Solution?
    *[Right, **if** dirt **then** Suck, Left, **if** dirt **then** Suck] + goto 1 until clean*
    need to take actions based on contingencies arising during execution



Task: Clean the room (#7 or #8)

Uncertainty remains!

Question: How to deal properly with uncertainty in autonomous intelligent agents?



Outline of this lecture

- Robust planning
  - ways to cope with complex, uncertain problems classically
- The probabilistic turn
  - uncertainty, probability theory & degrees of belief
- Bayesian Networks
  - inferences, interventions & causal effects
  - actions, utilities & decisions (DBN, BDN)
- Markov Decision Problems
  - complex decisions in complex situations
- Current trends
  - Relational probabilistic models
  - Markov/Bayesian Logic Networks