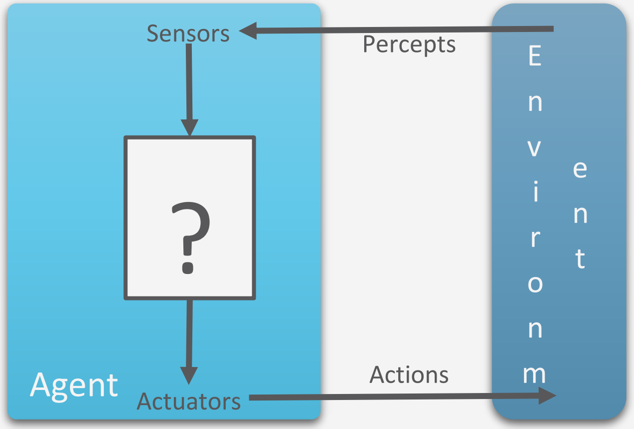
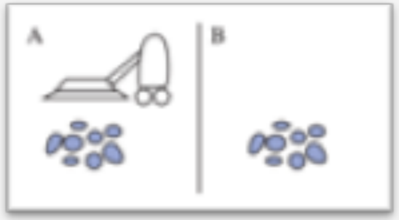
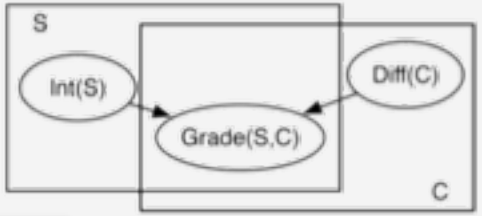
Introduction

Statistical Relational Artificial Intelligence (StaRAI)







$\infty \; Presents(X,P,C) \Rightarrow Attends(X,C)$

$3.75 \; Publishes(X,C) \land FarAway(C) \Rightarrow Attends(X,C)$

Contents

Overview: 1. Introduction

A. ***Artificial Intelligence***
   - Approaches: thinking / acting humanly / rationally

B. *Framework: Agent Theory*
   - Agent
   - Task environment
   - Agent structure

C. *Topic: StaRAI*
   - Motivation, context
   - Relational examples, outlook on probabilistic relational models (PRMs)

Approaches to Artificial Intelligence (AI)

- All approaches researched
  - Supported and hindered each other
- Rationality
  - System is rational if it does the “right thing,” given what it knows

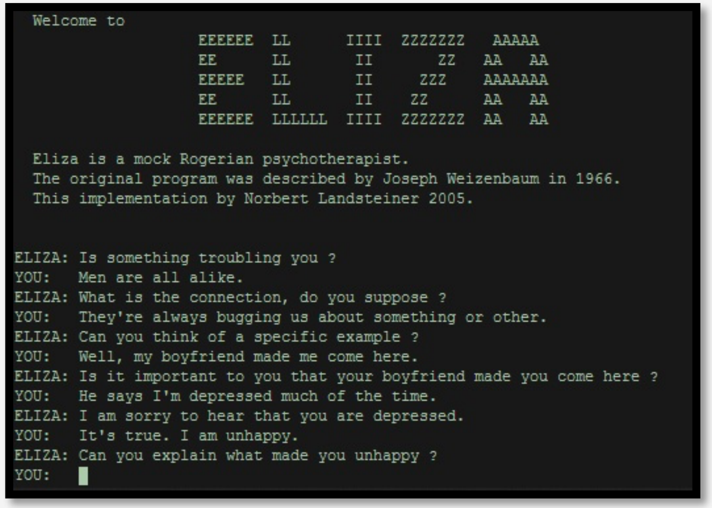
Success measure

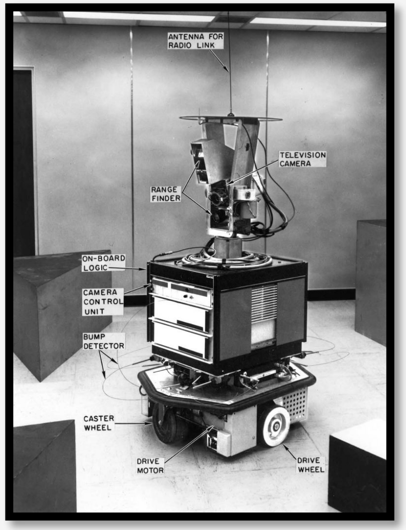
Fidelity of human performance	Ideal performance measure rationality	
Thinking Humanly	**Thinking Rationally**	
“The exciting new effort to make computers think . . . machines with minds, in the full and literal sense.” (Haugeland, 1985) “[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning…” (Bellman, 1978)	“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985) “The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)	Thought processes, reasoning
Acting Humanly	**Acting Rationally**	
“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990) “The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)	“Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998) “AI … is concerned with intelligent behaviour in artefacts.” (Nilsson, 1998)	Behaviour

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Acting Humanly

- Turing Test (Turing, 1950)
  - Computer passes test, if a human, who asks written questions, cannot tell if the the written answers come from a human or not
    - Example: *Eliza*, program for superficially simulating a psychiatrist
    - See also Ch. 26, “Artificial Intelligence – A Modern Approach” by Russel & Norvig, including a discussion whether a computer would really be intelligent if it passed
      - Regarding Eliza: human’s example closure tendencies are more pronounced for emotional/social intelligence aspects
      - Cf. robot *Shakey*: No on who saw Shakey the first time thought it could shoot hoops, yet the first people interacting with Eliza assumed it was a real doctor
  - *Total* Turing Test: includes a video signal to test perceptual abilities, opportunity to pass physical objects





https://en.wikipedia.org/wiki/ELIZA
https://thenewstack.io/remembering-shakey-first-intelligent-robot/

Acting Humanly

- Subproblems to solve as part of the Turing Test
  - *Natural Language Processing*
    - Communication
  - *Knowledge representation*
    - Store knowledge and observations
  - *Automated reasoning*
    - Answer questions, draw new conclusions
  - *Machine learning*
    - Adapt to new circumstances, detect and extrapolate patterns
  - Total Turing Test
    - *Computer vision*: perceive objects
    - *Robotics*: manipulate objects, move about

The Turing Test covers a majority of disciplines that make up AI nowadays.

- But:
  little research effort devoted to pass test
- Instead:
  Study underlying principles of intelligence

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Approaches to Artificial Intelligence (AI)

- All approaches followed
  - Supported and hindered each other
- Rationality
  - System is rational if it does the "right thing," given what it knows

Success measure

	Fidelity of human performance	Ideal performance measure rationality
Thought processes, reasoning	**Thinking Humanly** "The exciting new effort to make computers think . . . machines with minds, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning..." (Bellman, 1978)	**Thinking Rationally** "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Behaviour	**Acting Humanly** "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	**Acting Rationally** "Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998) "AI … is concerned with intelligent behaviour in artefacts." (Nilsson, 1998)

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Thinking Humanly

- A “program thinks like a human“
  - Requires a way to determine how humans think → workings of the human mind
  - Given theory of the mind, express theory as computer program
    - If program’s input-output behaviour matches corresponding human behaviour, evidence that some of program’s mechanisms could also be operating in humans
- Approach complementary to AI: *Cognitive Science*
  - Interdisciplinary:
    - Computer models from AI
    - Experimental techniques from psychology
  - Goal:
    Construct precise and testable theories of human mind

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Approaches to Artificial Intelligence (AI)

- All approaches followed
  - Supported and hindered each other
- Rationality
  - System is rational if it does the “right thing,” given what it knows

Success measure

	Fidelity of human performance	Ideal performance measure rationality	
	Thinking Humanly	**Thinking Rationally**	
	“The exciting new effort to make computers think . . . machines with minds, in the full and literal sense.” (Haugeland, 1985) “[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning…” (Bellman, 1978)	“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985) “The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)	Thought processes, reasoning
	Acting Humanly	**Acting Rationally**	
	“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990) “The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)	“Computational Intelligence is the study of the design of intelligent agents.” (Poole et al., 1998) “AI … is concerned with intelligent behaviour in artefacts.” (Nilsson, 1998)	Behaviour

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Thinking Rationally

- Codify thinking → rules
  - Irrefutable reasoning processes
  - Argument structures that always yield correct conclusions when given correct premises
- Field of *Logic*
  - Precise notation for statements about objects in a world and relations among them
  - Programs that could, in principle, solve *any* solvable problem described in logical notation
  - Obstacles:
    - Informal knowledge
      - Unstructured data
      - Uncertainty
    - Solving any solvable problem in practice
      - Limited computational resources

Obstacles apply to *any* attempt to build computational reasoning systems
- Formulated first in logic

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Approaches to Artificial Intelligence (AI)

- All approaches followed
  - Supported and hindered each other
- Rationality
  - System is rational if it does the "right thing," given what it knows

Success measure

	Fidelity of human performance	Ideal performance measure rationality	
	Thinking Humanly	**Thinking Rationally**	
	"The exciting new effort to make computers think . . . machines with minds, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning…" (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)	Thought processes, reasoning
	Acting Humanly	**Acting Rationally**	
	"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998) "AI … is concerned with intelligent behaviour in artefacts." (Nilsson, 1998)	Behaviour

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Acting Rationally

- Rational agent approach
- Agent = something that acts
  - Operate autonomously
  - Perceive environment
  - Persist over a prolonged time period
  - Adapt to change
  - Create and pursue goals
- Rational agent
  - One that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome
  - May include thinking rationally or acting humanly, but *more general*

Advantage: Standard of rationality mathematically well defined

- Better suited to generate agent designs that provably achieve rationality
- Focus of the next slides

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Interim Summary

- Four approaches to define AI
  - Acting humanly
    - Turing test
    - Subproblems: Natural language processing, knowledge representation, reasoning, machine learning, computer vision, robotics
  - Thinking humanly
    - Cognitive sciences as an interdisciplinary science between AI and psychology
  - Thinking rationally
    - Logic for knowledge representation and correct reasoning
  - Acting rationally
    - Rational agent as a generalisation of thinking rationally with a mathematical definition of rationality as a formal criterion
    - Also contains subproblems, which the Turing test identifies

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Overview: 1. Introduction

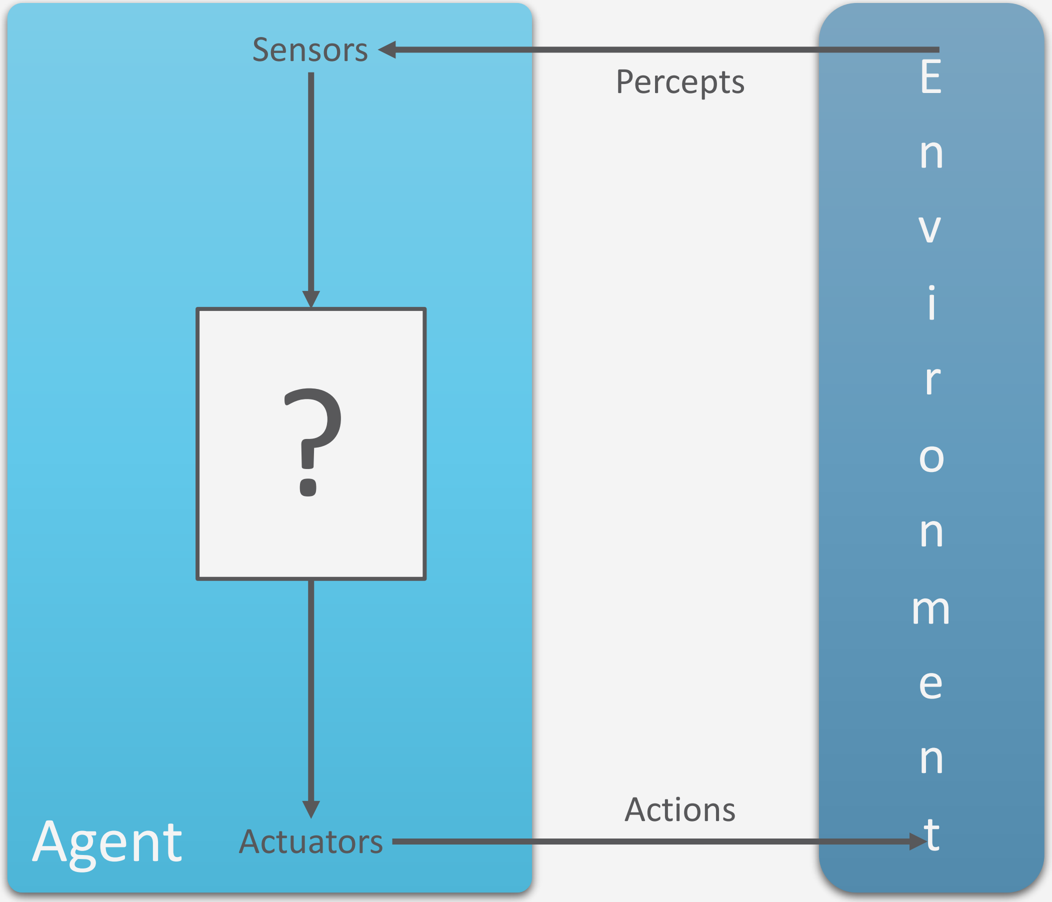
A. *Artificial Intelligence*
- Approaches: thinking / acting humanly / rationally

B. ***Framework: Agent Theory***
- Agent
- Task environment
- Agent structure

C. *Topic: StaRAI*
- Motivation, context
- Relational examples, outlook on probabilistic relational models (PRMs)

Agent

- Something that perceives its environment through sensors and acts through actuators
  - Human agent
    - Sensors: eyes, ears, further organs
    - Actuators: hands, legs, mouth, other body parts
  - Robot agent
    - Sensors: cameras, infra-red sensors, etc.
    - Actuators: motors
  - Software agent
    - Goal: *document retrieval, DR*
    - Sensors: input interface for textual queries
    - Actuators: output interface for documents

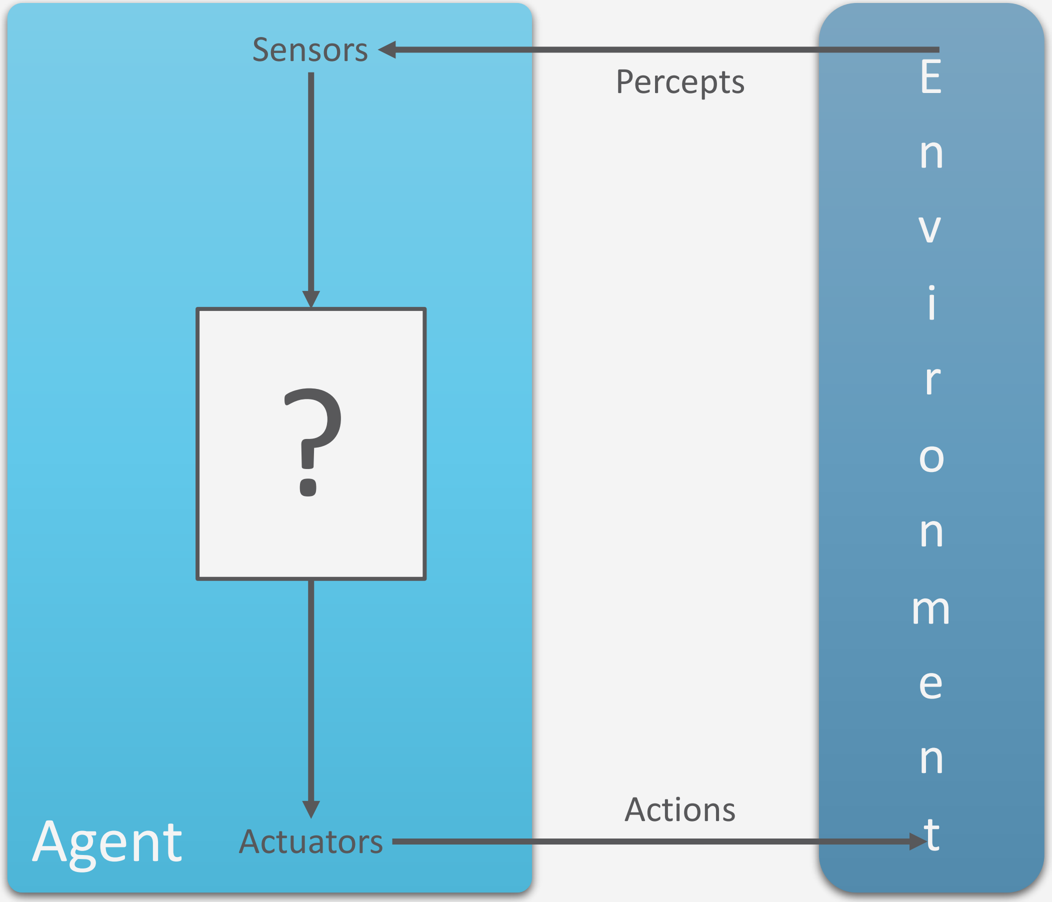


Agent Behaviour

- Agent function
  - Mapping from percept sequences to actions
$$f : \boldsymbol{P}^* \rightarrow \boldsymbol{A}$$
    - $\boldsymbol{P}^*$ Set of possible percept sequences
    - $\boldsymbol{A}$ Set of available actions

- Agent program
  - Implementation of the agent function f
  - Runs on a physical system (architecture)

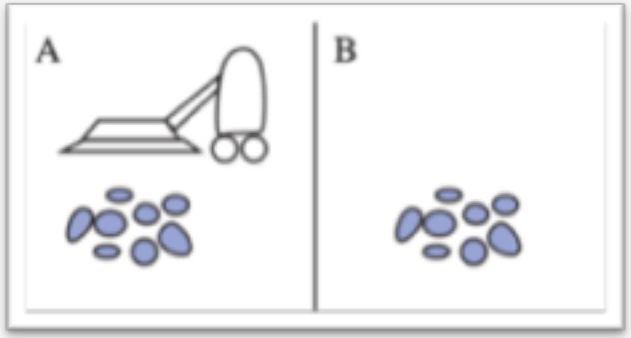
- Agent = architecture + agent program



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Simple Example

- Vacuum cleaner
  - Two locations: squares *A*, *B*
  - Possible percepts: location; location *clean*, *dirty*
  - Available actions: *right*, *left*, *vacuum*
- Agent function in table
  - Mapping increasingly longer percept sequences to actions
    - Optimisable, as action only depends on last observation (and not the sequence)
  - How table is filled defines the agent
    - Different agents possible



Percept sequence	Action
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Vacuum*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Vacuum*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…

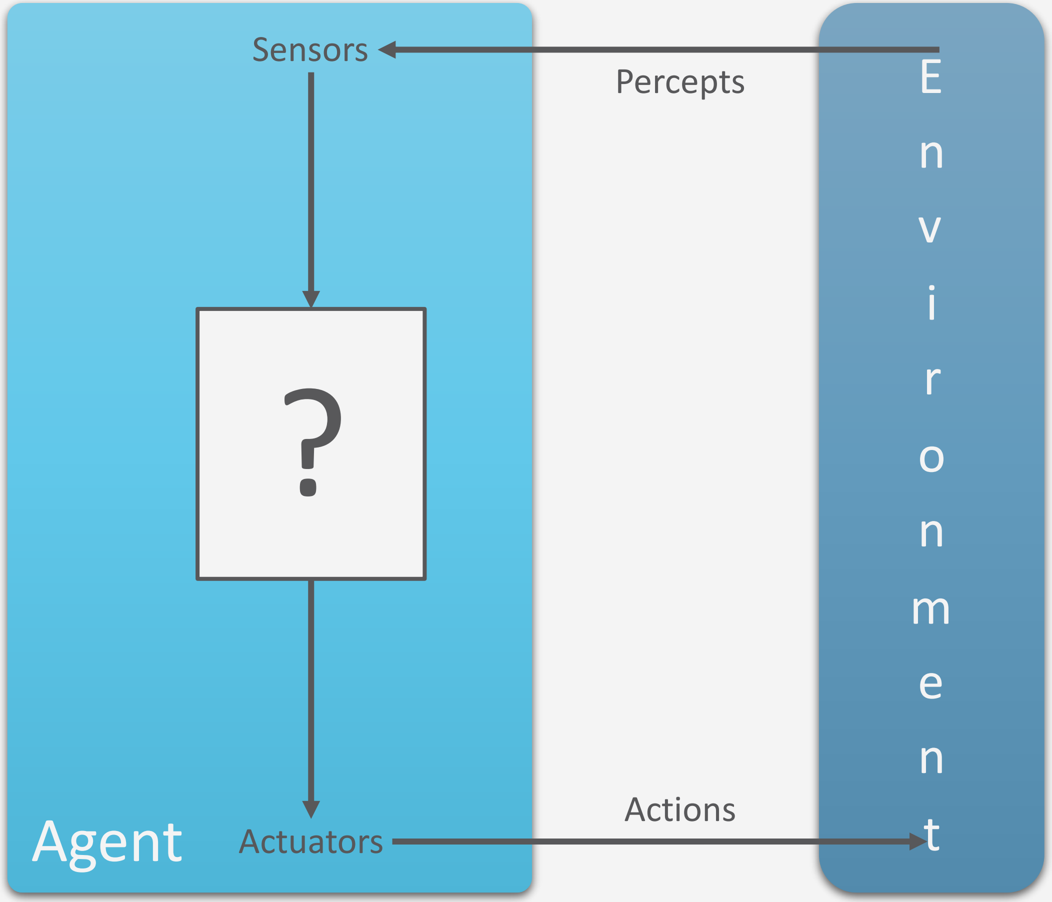
S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Rationality

Rationality is not omniscience!

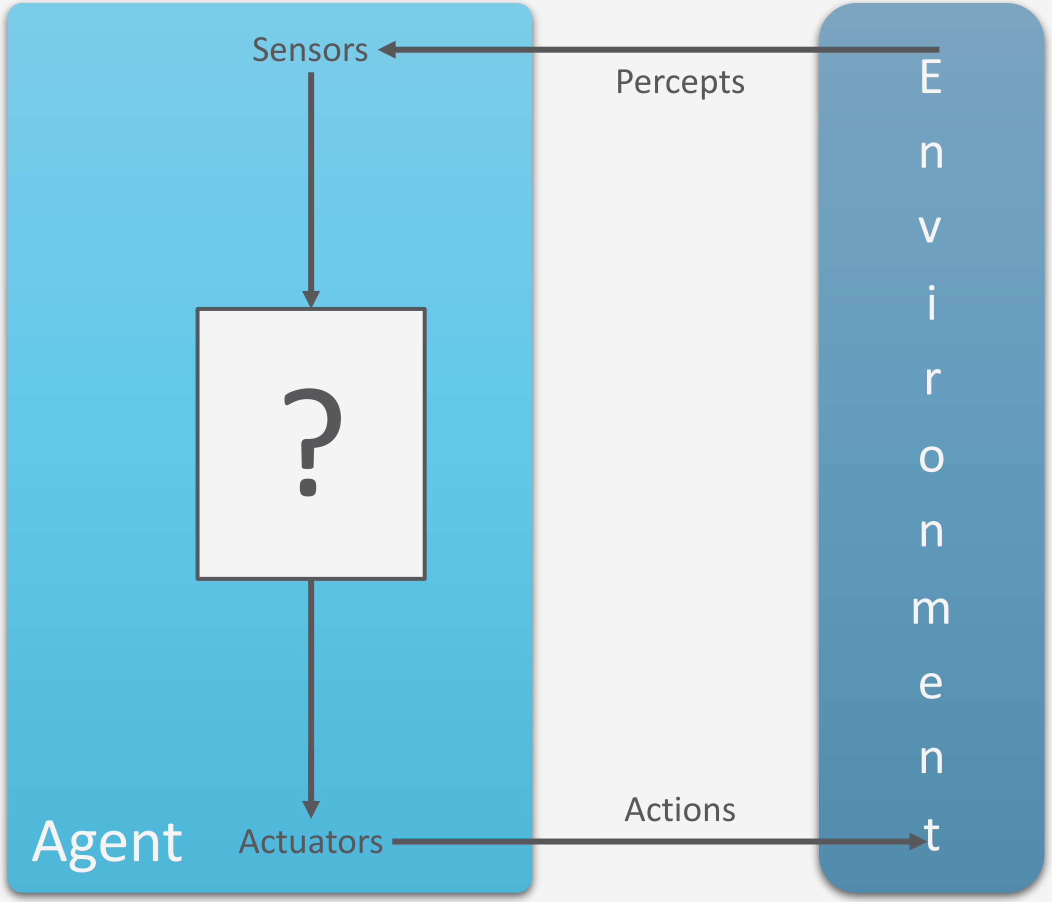
- Depends on four things:
  - *Performance measure*, defines criterion of success
  - Agent's prior *knowledge* of environment
  - *Actions* that agent can perform
  - Agent's *percept sequence* to date
- Rational agent:
  - For each possible percept sequence, a rational agent should select an *action*
  - expected to maximize its *performance measure*,
  - given evidence provided by *percept sequence* and
  - whatever built-in *knowledge* the agent has.

→ Rational = intelligent



Further Agent Properties

- Properties that agent behaviour can exhibit:
  - Rationality
  - Flexibility
    - Reactive
    - Proactive
    - Social (in multi-agent systems)
  - Autonomy
  - Mobility
  - Veracity
  - Benevolence
  - Ability to learn / adapt



Task Environment

- Task environments
  - Essentially the “problems” to which rational agents are the “solutions”
  - Specify using PEAS description
    - Performance measure
    - Environment
    - Actuators
    - Sensors

Attention: Name collision between
task *environment* und *environment*.
We will after this part usually only talk about the environment.

→ Properties of task environment
→ Modelling the environment
- Determine agent design

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

PEAS – Example

- Medical diagnosis system

Agent Type	Performance Measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, reduced costs	Patient, hospital, staff	Display of questions, tests, diagnoses, treatments, referrals	Keyboard entry of symptoms, findings, patient's answers

- Software agent for *document retrieval, DR*
  - Sensors: input interface for textual queries
  - Actuators: output interface for documents

What makes up the *environment*? What is a fitting *performance measure*?

Properties of Task Environments

- Fully observable (vs. partially observable)
  - Agent‘s sensors give access to complete state of environment at each point in time
- Single agent (vs. multiple agents)
  - Single agent acts in environment
  - Can depend on modelling if something in environment is an object or another agent
  - Multi-agent systems:
    - Cooperative vs. competitive
    - Communication possible?
- Deterministic (vs. stochastic)
  - Next state of environment completely determined by current state and action executed by agent
  - Strategic: if environment deterministic except for actions of other agents

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Properties of Task Environments

Special property: known (vs. unknown)
- Agent's (or designer's) state of knowledge about its environment beforehand
- *No* property of the environment

- Episodic (vs. sequential)
  - Agent's experience divided into atomic "episodes"
    - Episode = agent perceives a single percept and then performs a single action
  - Choice of action in each episode depends only on episode itself
- Static (vs. dynamic)
  - Environment unchanged while agent is deliberating
  - Semi-dynamic: environment itself does not change with passage of time but agent's performance score does
- Discrete (vs. continuous)
  - Applies to *state* of the environment, to the way *time* is handled, and to the *percepts* and *actions*
  - Finite number of distinct states, percepts, actions, time steps

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Properties of Task Environments – Example

Task environment	Observable	Single agent	Deterministic	Episodic	Static	Discrete
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Self-driving taxi	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete

How about image classification?

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Representation of the Environment

- Depends on properties of the task environment
  - Discrete/continuous, static/dynamic, episodic/sequential, ...
- Determine expressiveness of representation
  - Increasing complexity and expressiveness

1. Atomic
   - Each state not further dividable (no internal structure)
2. Factorised
   - Partitioning of state into fixed set of variables / attributes that can take a value
   - Representation: state-variable models, propositional logic, probabilistic graphical models (PGMs)
3. Structured
   - Objects, relations among them
   - Representation: relational databases + PGMs, first-order logic

Advertisement:
Lecture „*Automated Planning and Acting*“

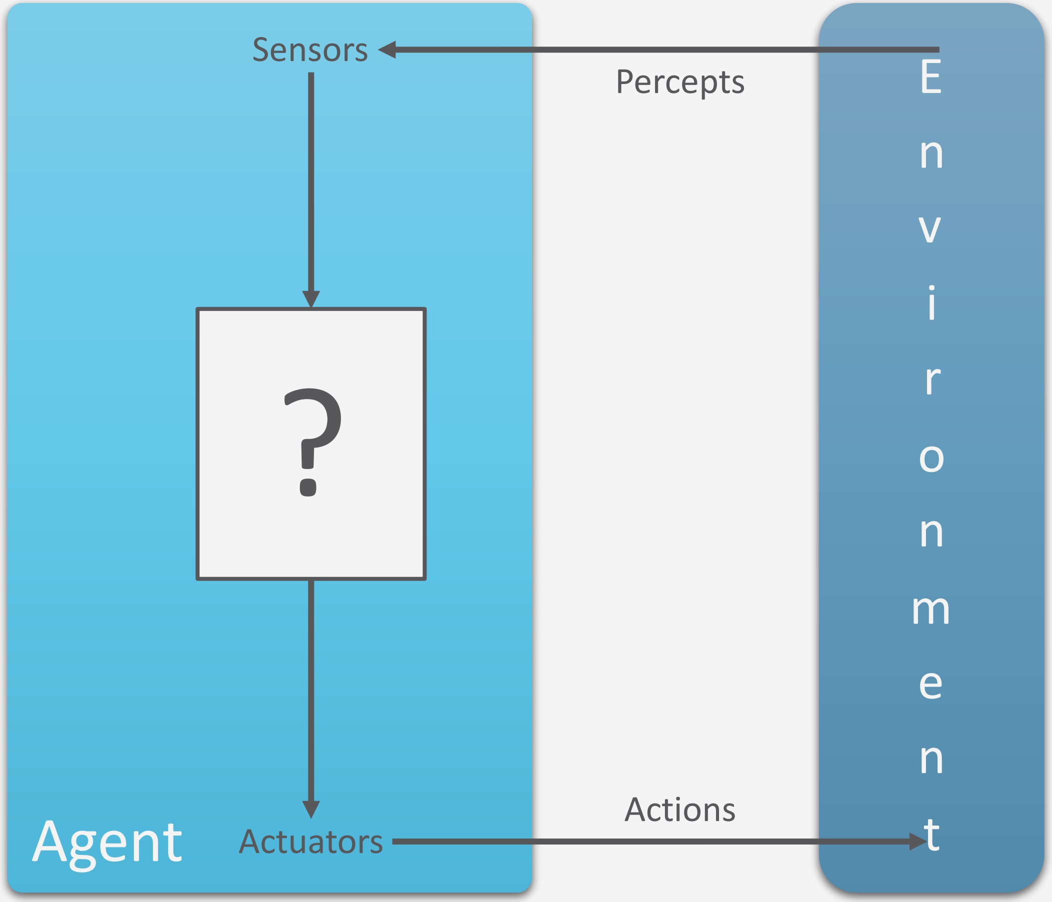
Advertisement:
Lecture „*Data Science / Intro to AI*“

This lecture on StaRAI!

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Performance Measure: Good Behaviour

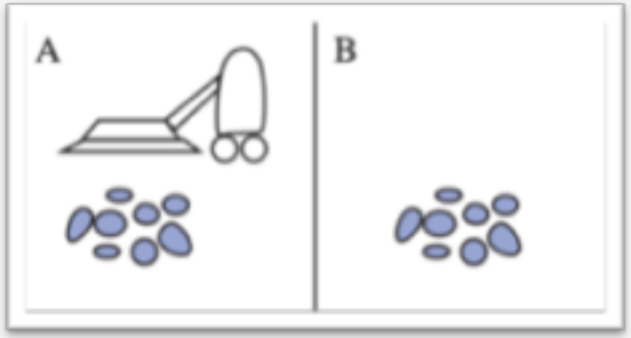
- What is the right thing to do?
  - Agent generates sequence of actions given percepts
  - Causes environment to go through sequence of states
  - If sequence desirable / conforms to our expectations, then agent performed well
- Performance measure that evaluates any given sequence of environemnt states



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Back to the Vacuum Cleaner

Given this task environment, rational?



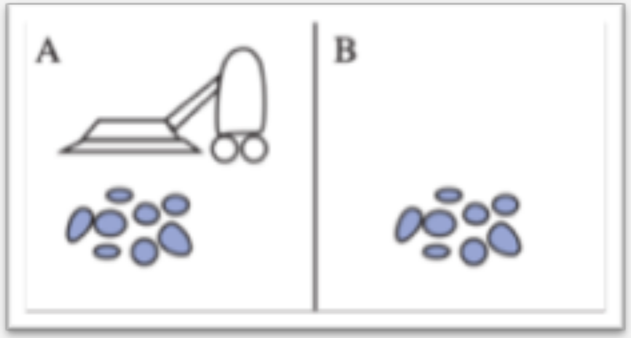
- Performance Measure
  - 1 point for each clean square in each time step over a life span of 1000 time steps
- Environment
  - Locations (*A*, *B*) prior knowledge
  - Not known: initial location, dirt distribution
  - Action *Vacuum* cleans the current location; clean location remains clean
  - Actions *Left* and *Right* move the agent accordingly, except if that moves the agent out of the environment, then it stays put
  - Location and dirtiness perceived correctly

Percept sequence	Action
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Vacuum*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Vacuum*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Back to the Vacuum Cleaner

Given this task environment, rational?



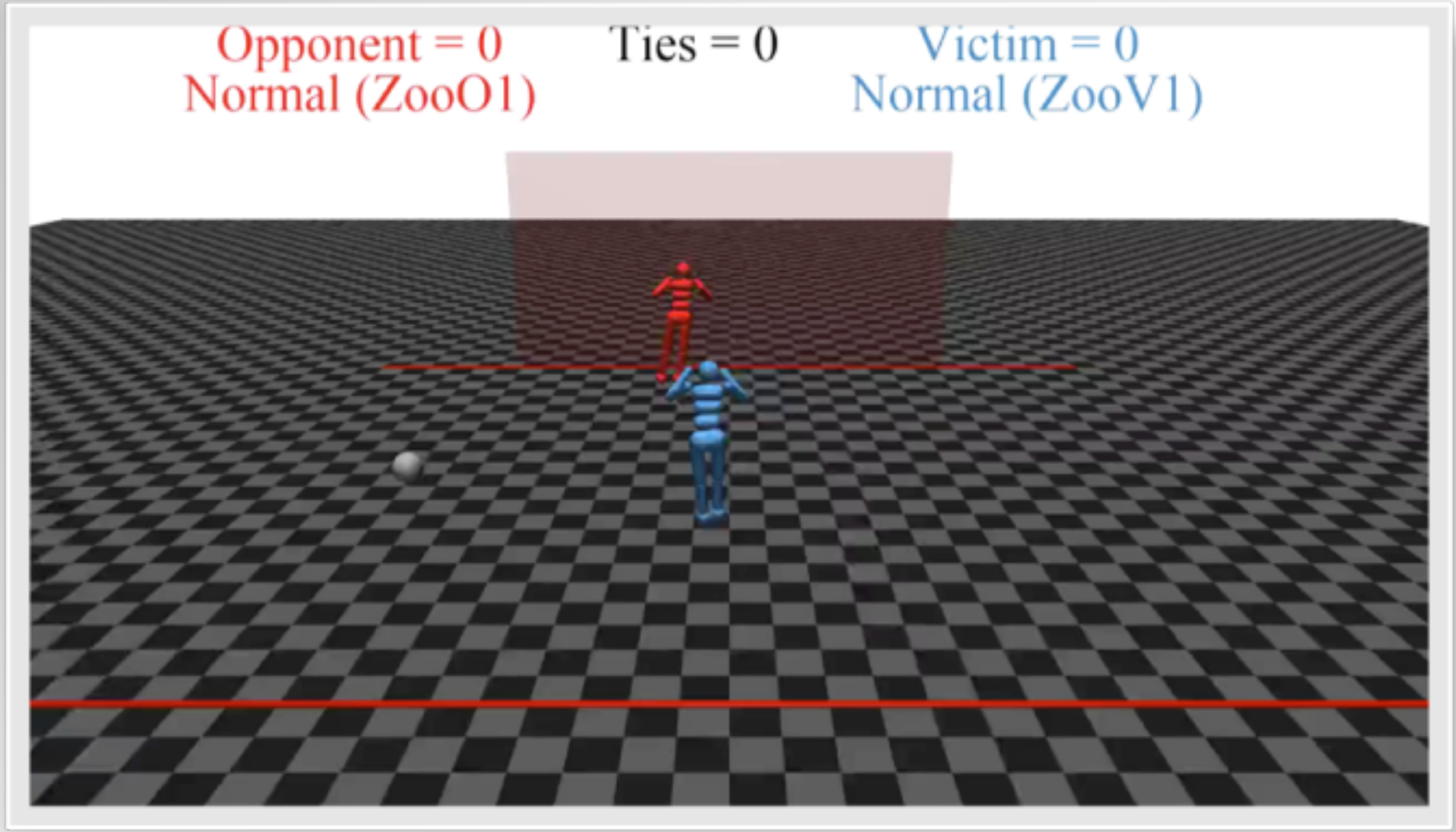
- Performance Measure
  - 1 point for each clean square in each time step over a life span of 1000 time steps
  - 1 point deduction for each *Left* or *Right* action
- Environment
  - As before

How do we get a better agent?

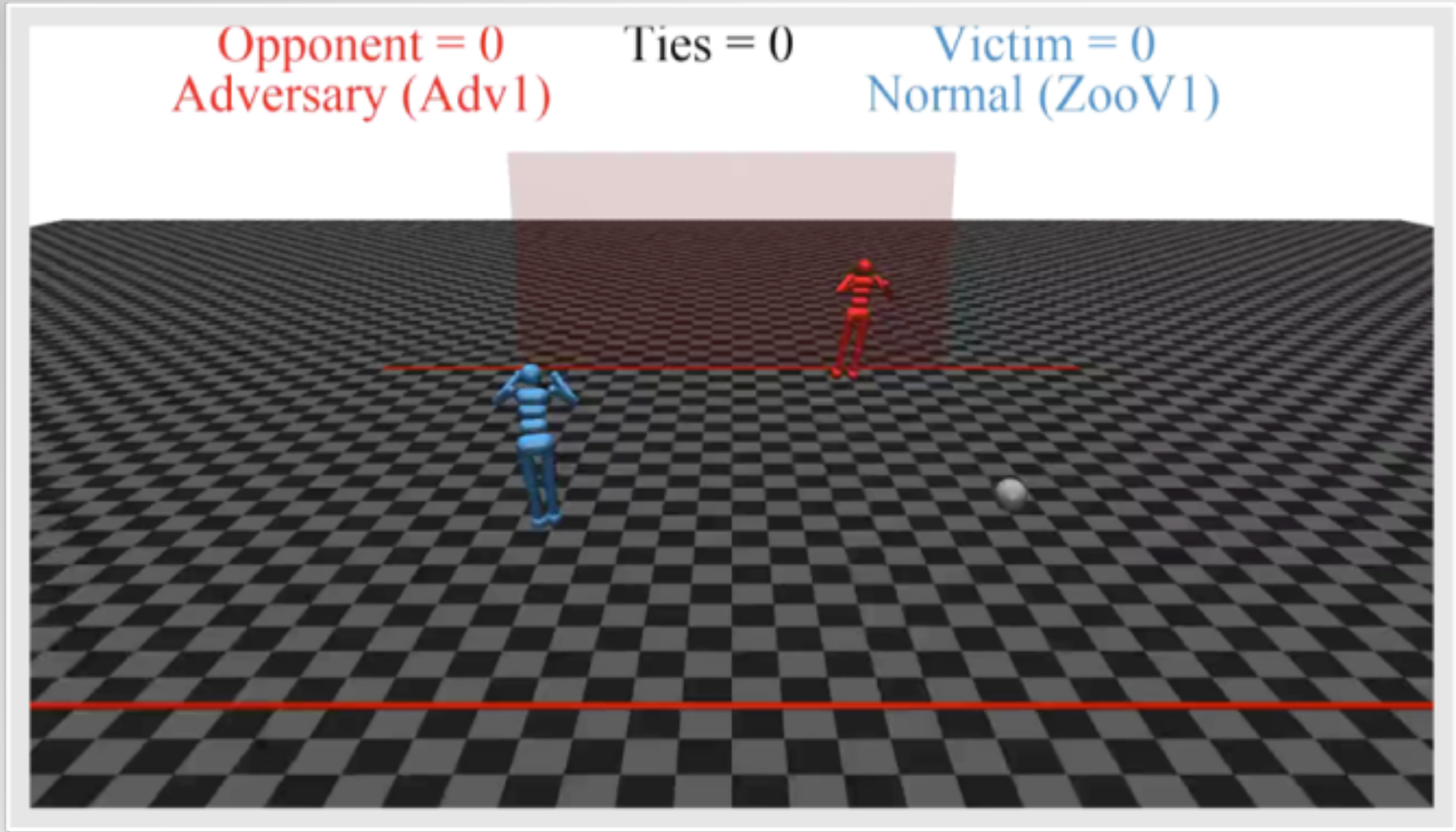
- Additional action: *No-Op*
  - Adapt agent function such that in the case of dirt the locations get cleaned once and then *No-Op*
  - If locations can get dirty again, repeat cleaning round in intervals

Percept sequence	Action
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Vacuum*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Vacuum*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.



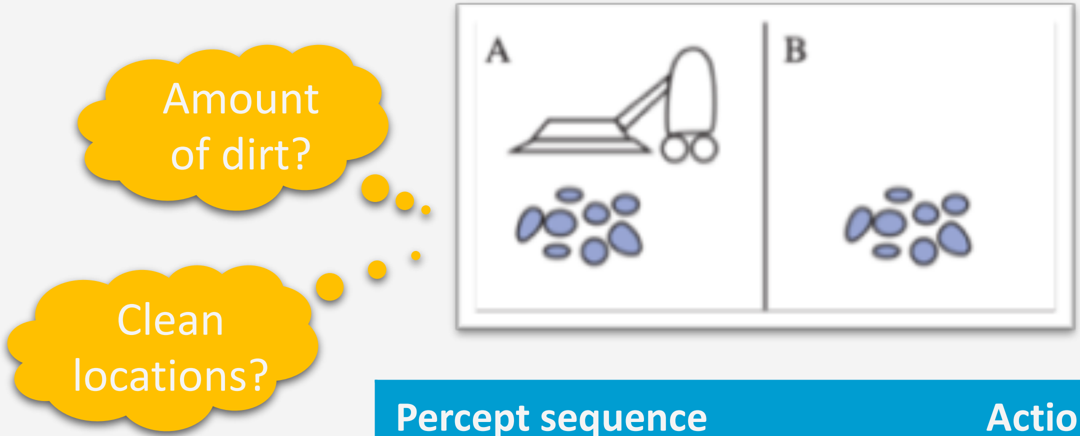
Quelle: https://people.eecs.berkeley.edu/~russell/talks/2020/russell-aaai20-hntdtwwai-4x3.pptx



Quelle: https://people.eecs.berkeley.edu/~russell/talks/2020/russell-aaai20-hntdtwwai-4x3.pptx

Performance Measure

- Hard to determine
  - Not one fixed performance measure for all tasks and agents



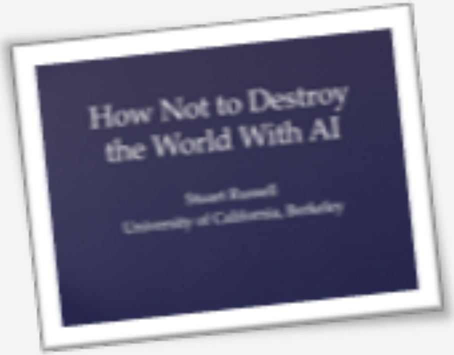
Percept sequence	Action
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Vacuum*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Vacuum*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…

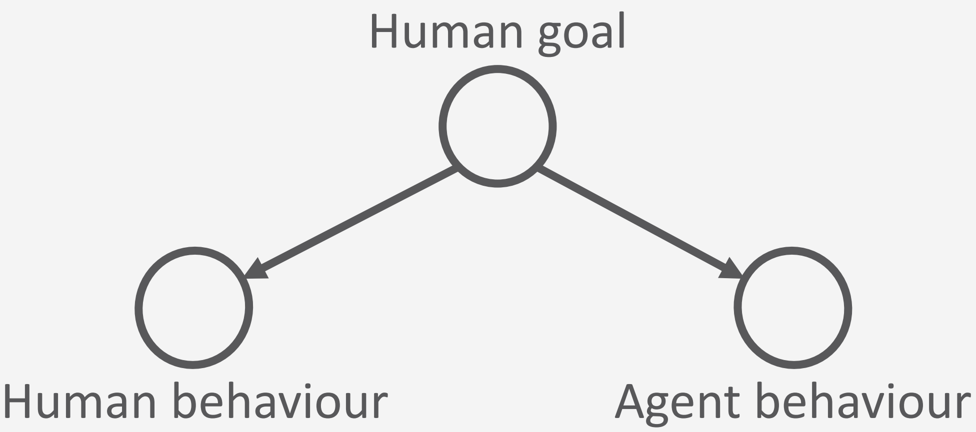
As a general rule, it is better to design performance measures according to what one actually wants in the environment, rather than according to how one thinks the agent should behave.

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Side Note: Current Research

- Idea:
  - Humans: intelligent to the extent that our actions can be expected to achieve our goals
  - ~~Maschines: intelligent to the extent that their actions can be expected to achieve their goals~~
  - Maschines are *beneficial* to the extent that their actions can be expected to achieve our goals n
  - Approach: Performance measure unknown, human as assistant
- Goal: *Provably beneficial AI*
  - See for example Stuart Russell



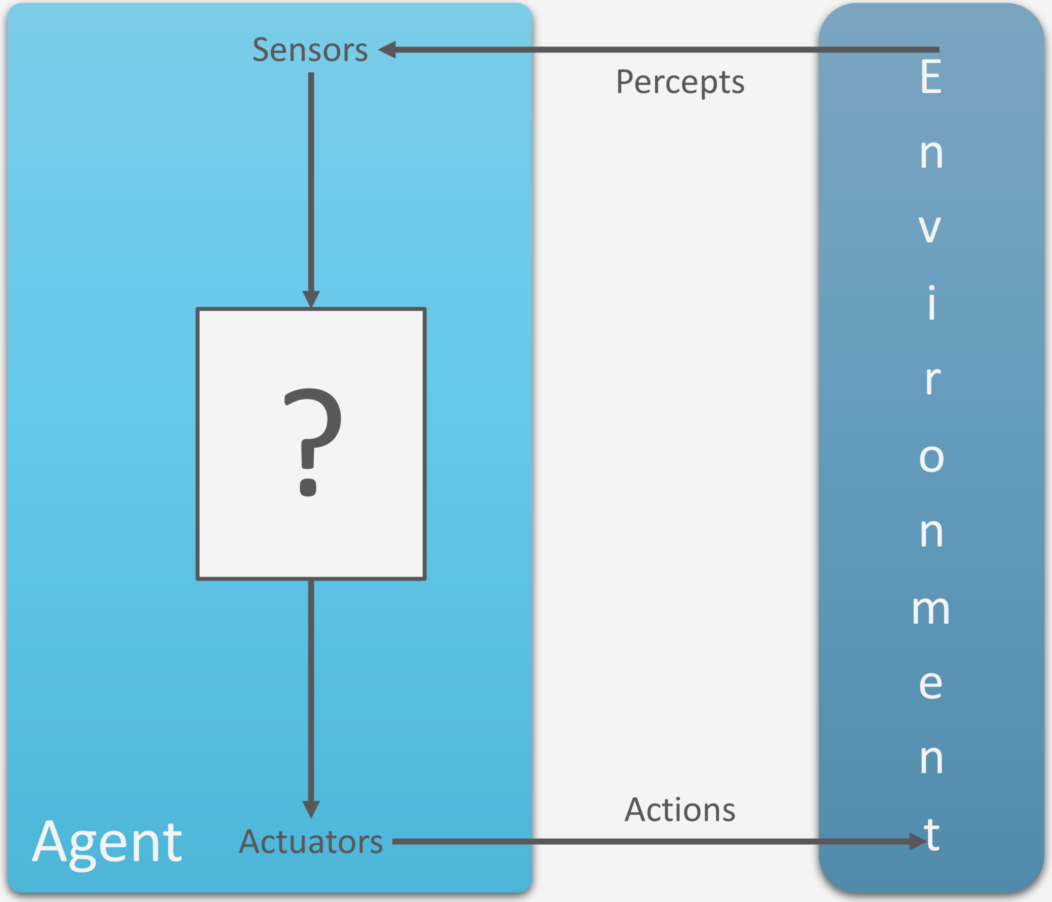


Presentation: https://www.youtube.com/watch?v=QPSgM13hTK8
Slides: https://people.eecs.berkeley.edu/~russell/talks/2020/russell-aaai20-hntdtwwai-4x3.pptx

Implementing Agents

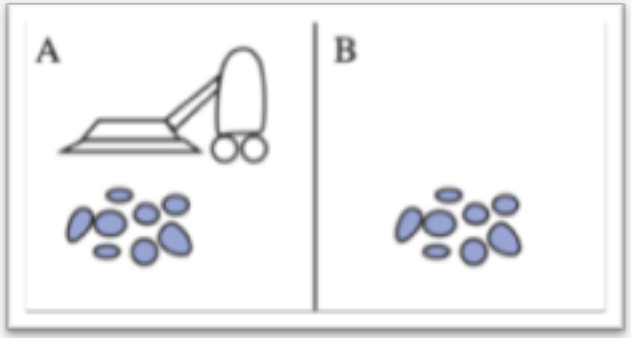
- Agent = architecture + program
  - Architecture
    - Entity with corresponding physical sensors and actuators
      - Entity = computer, robot, software (interfaces as sensors and actuators)
  - Program
    - Implementation of an agent function
    - Runs on an architecture
  - Both depend on task environment

- Implementation necessary, but how?



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Simple Table-driven Agent



- Look up action in table given current percept sequence

function TABLE-DRIVEN-AGENT(*percept*) **returns** an action
  persistent: *percepts*, a sequence, initially empty
    table, a table of actions, indexed by percept sequences

  append *percept* to the end of *percepts*
  action ← LOOKUP(*percepts*, *table*)
  return *action*

Percept sequence	Action
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Vacuum*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Vacuum*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Vacuum*
…	…

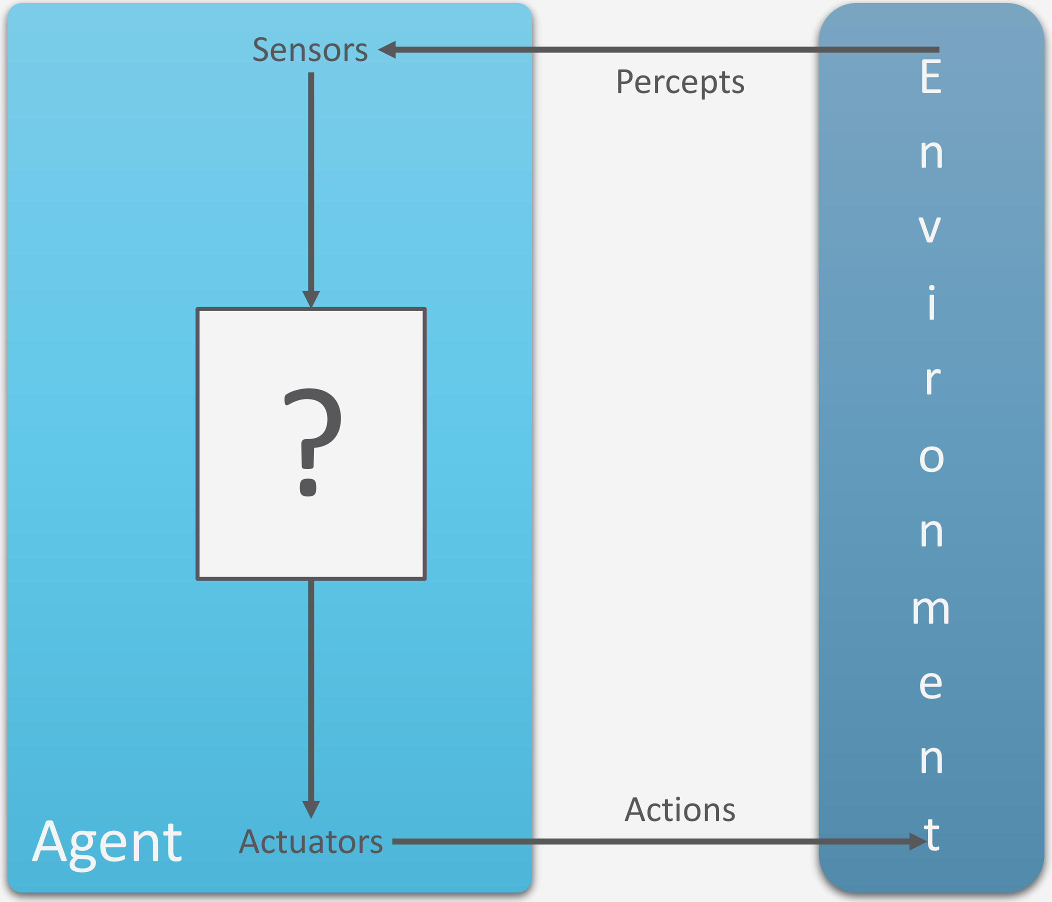
- Disadvantages
  - Table possibly very large: $\sum_{t=1}^{T} |\boldsymbol{P}^*|^t$
    - $\boldsymbol{P}^*$ set of all possible percepts, T life span
  - No autonomy
  - Takes a lot of time to generate table
    - Learn automatically not better (many entries to learn)

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Agent Structure

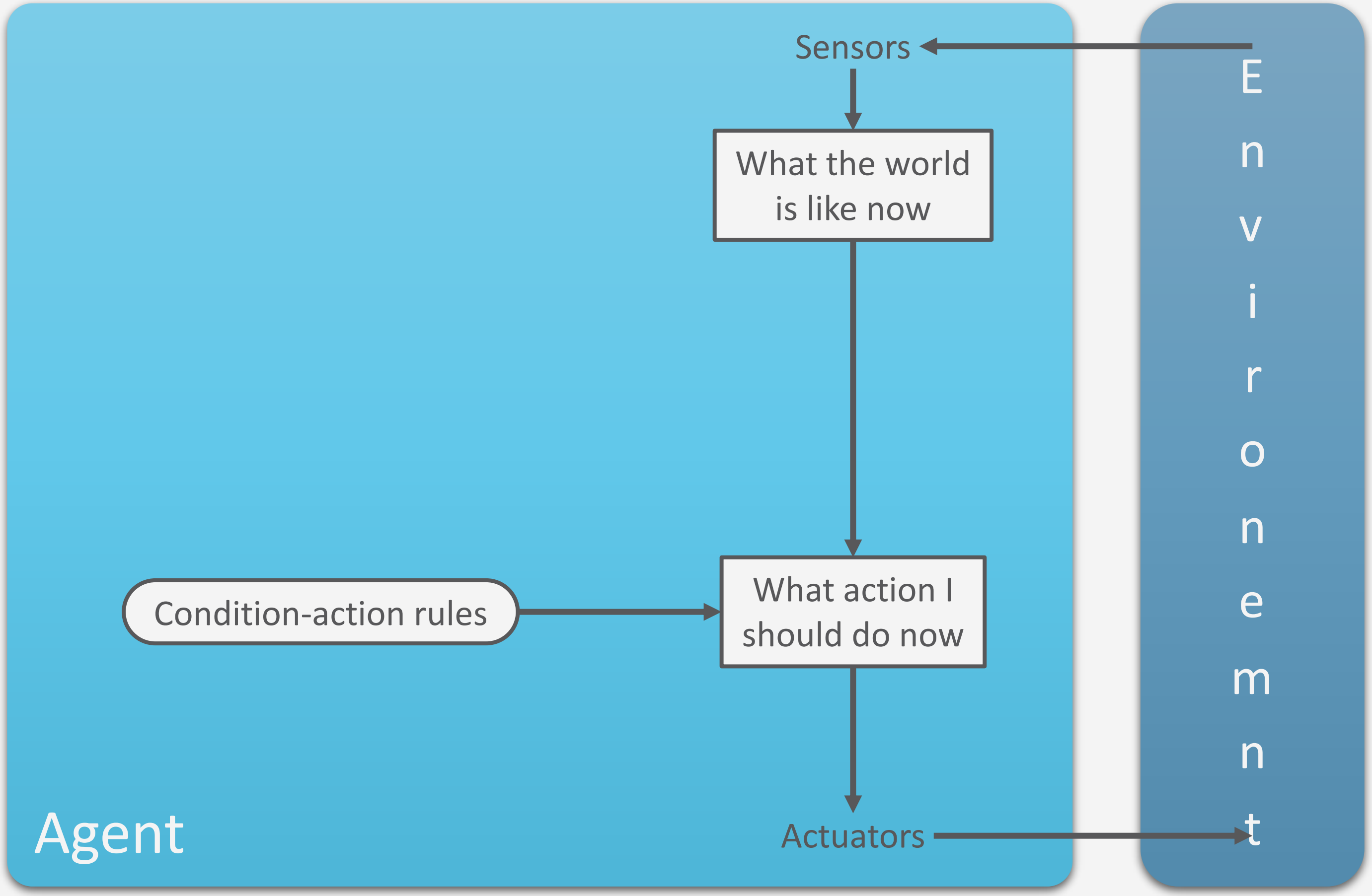
- General agent structures for implementing agent functions:
  - Ordered by increasing generalisation
1. Simple reflex agent
2. Model-based reflex agent
3. Goal-based agent
4. Utility-based agent
5. Learning agent

❖ Human-aware agent (current research)



Agent Structure: Simple Reflex Agent

- Actions chosen based on current percept
  - Ignores previous percepts
  - No modelling of the environemnt
- Only correct decision on action if environment fully observable
  - If partially observable, inifinite loops possible
    - (Partial) solution:
      Choose random action



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Implementation of a Simple Reflex Agent

- General representation

function SIMPLE-REFLEX-AGENT(*percept*) **returns** an action
    persistent: *rules*, a set of condition-action rules

    state ⟵ INTERPRET-INPUT(*percept*)
    rule ⟵ RULE-MATCH(*state*, *rules*)
    action ⟵ *rule*.ACTION
    return *action*

- Example vacuum cleaner

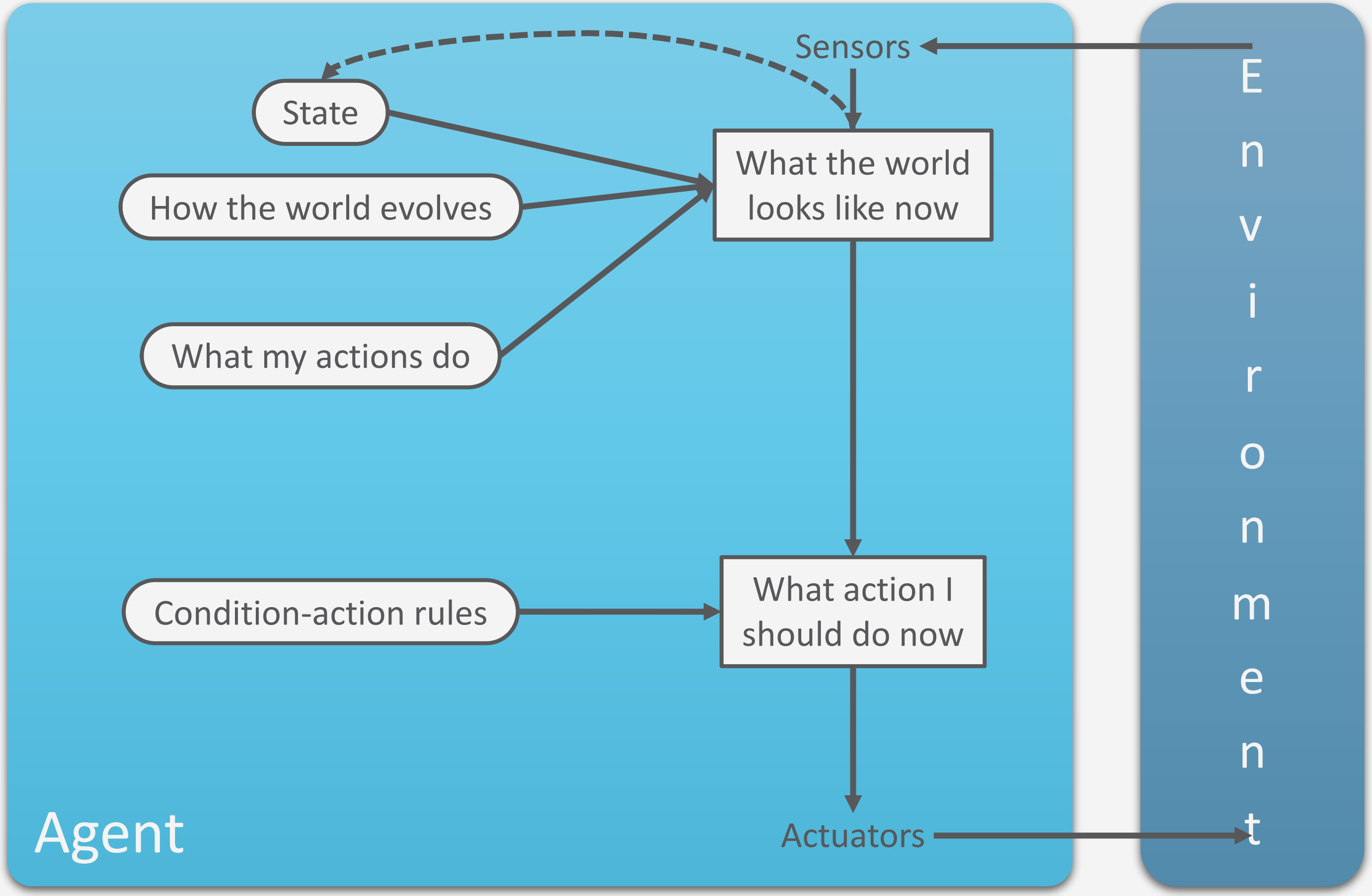
function REFLEX-VACUUM-AGENT([*location*, *status*]) **returns** an action
    persistent: *rules*, a set of condition-action rules

    if *status* = *Dirty* **then return** *Vacuum*
    else if *location* = *A* **then return** *Right*
    else if *location* = *B* **then return** *Left*

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Agent Structure: Model-based Reflex Agent

- Given partial observability, keep track of not observable part
  - Using internal state
    - Depends on percept sequence
    - Encode not observable aspects of the current state
      - Encoding of environment atomic, factorised or structured
- Update internal state with information over
  - How environment evolves independently of agent
  - Effect of actions on environment



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Implementation of a Model-based Reflex Agent

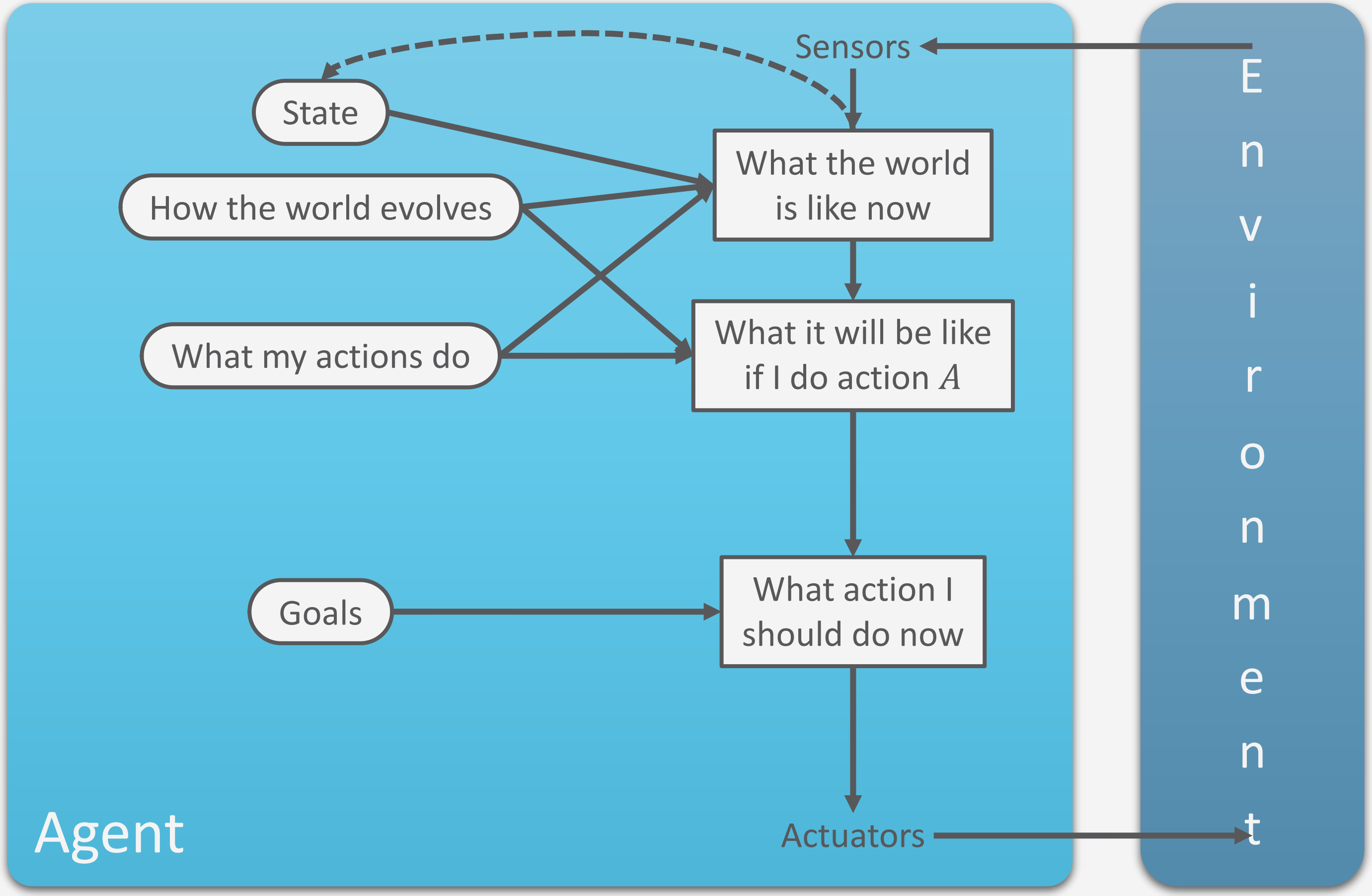
```
function MODEL-BASED-REFLEX-AGENT(percept) returns an action
    persistent: state, the agent's current conception of the world state
                model, a description of how the next state depends on current state and action
                rules, a set of condition-action rules
                action, the most recent action, initially none

    state ⟵ UPDATE-STATE(state, action, percept, model)
    rule ⟵ RULE-MATCH(state, rules)
    action ⟵ rule.ACTION
    return action
```

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Agent Structure: Goal-based Agent

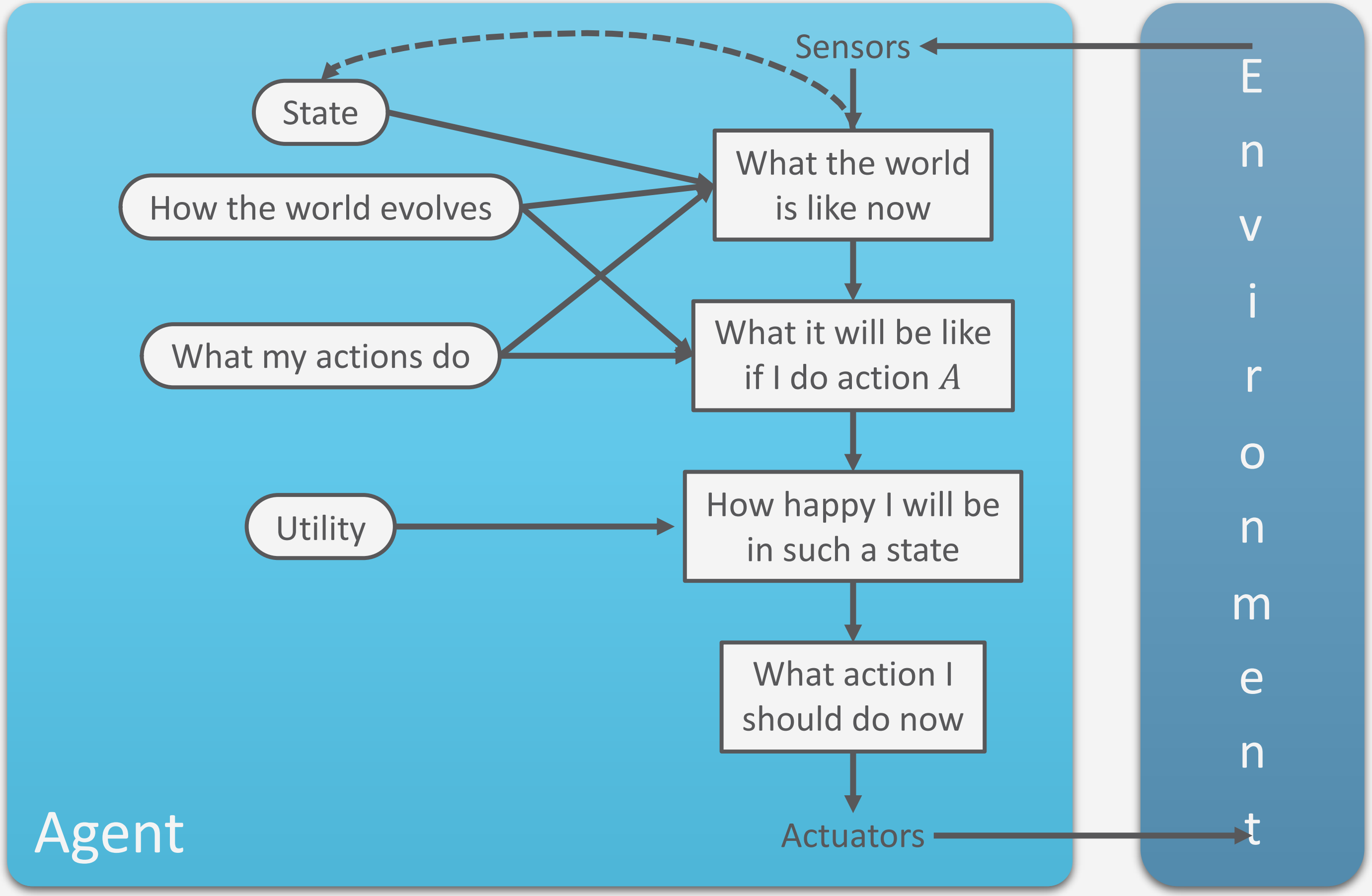
- Goal information useful
  - Description of desirable states
    - Infer from performance measure
    - Conditions for a goal state to fulfil
    - Example: vacuum cleaner
      $\forall x \in Loc : x = clean$
- Combine current state and goal information to choose actions that lead to goal
- Research areas:
  - Search
  - Planning



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Agent Structure: Utility-based Agent

- Goal-based: binary distinction between *happy* and *unhappy*
- Utility as a distribution over possible states
  - What we look at later in the lecture
  - Essentially an internalisation of the performance measure
    - If internal utility function *agrees with* external performance measure:
    - Agent that chooses actions to maximize its utility will be *rational* according to the external performance measure

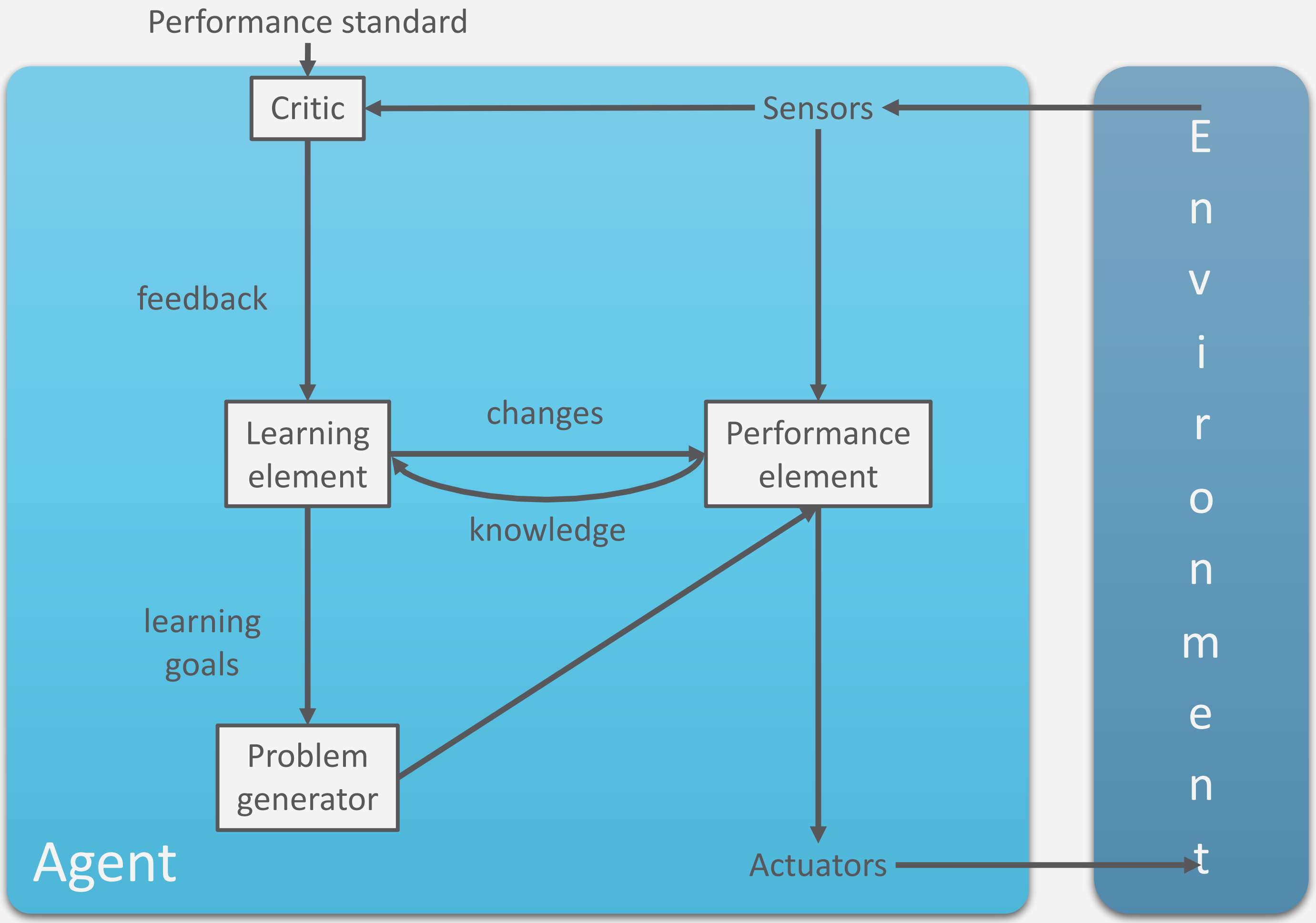


S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Agent Structure: Learning Agent

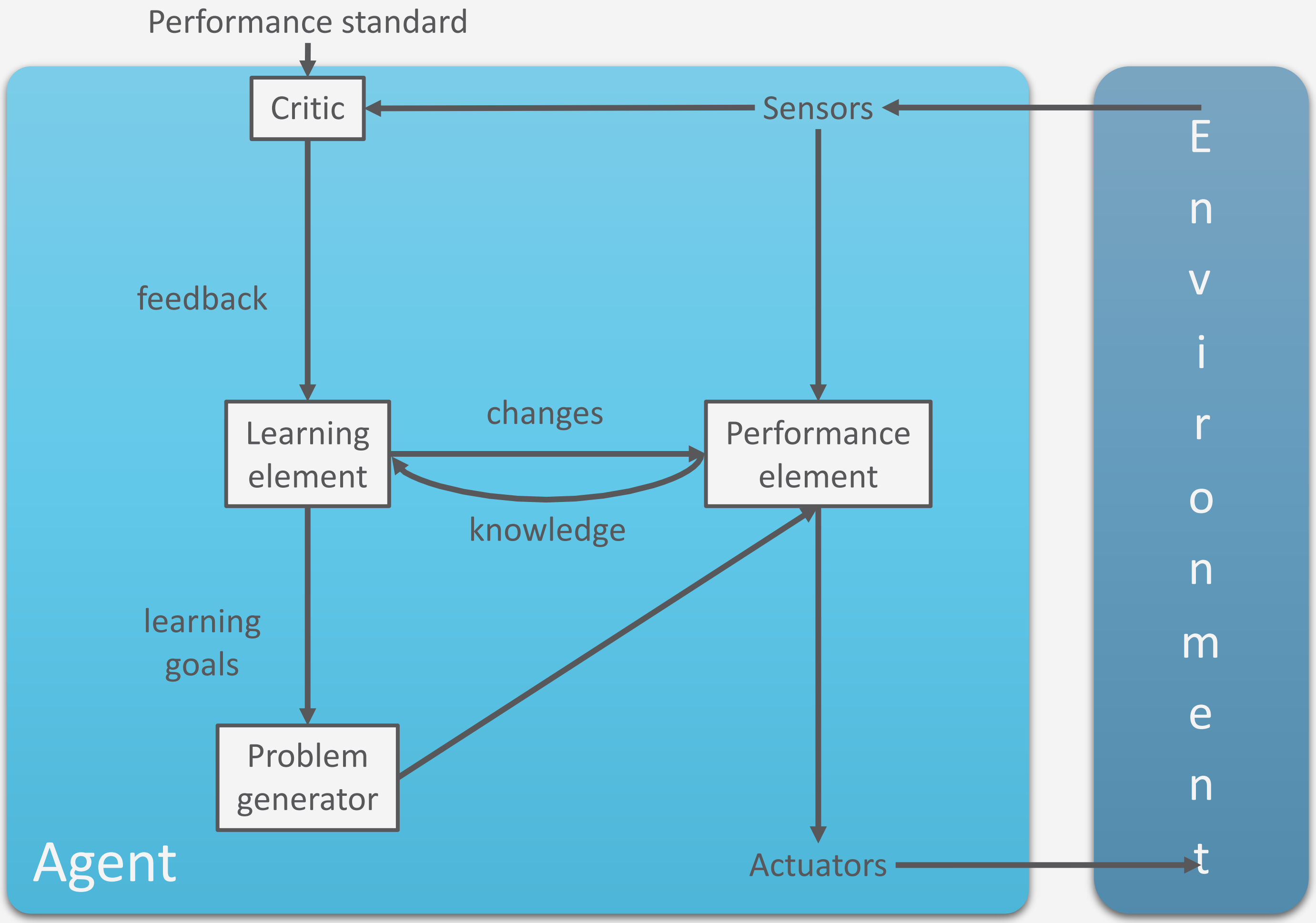
- So far: Agents select actions
  - Make decisions about actions
- How do agent programs emerge?
  - Too laborious to do by hand
- Generate learning agent, let it learn
  - Allows agent to operate in initially unknown environments and to become more competent
  - Research area: *Reinforcement Learning*

Advertisement: Lecture „*Deep Reinforcement Learning*“ (PI, 4) by Malte Schilling this semester



Agent Structure: Learning Agent

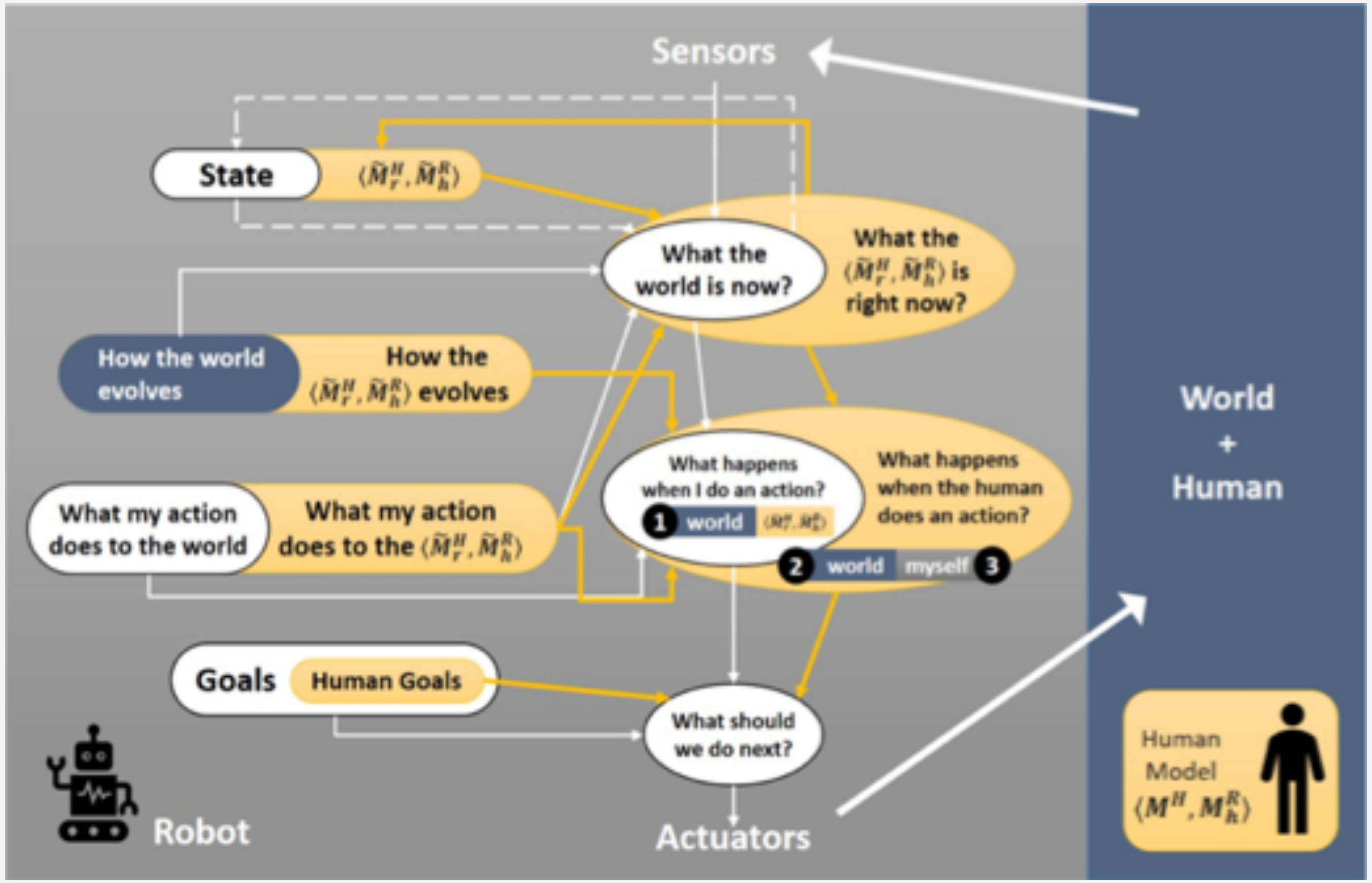
- 4 conceptual components
  - Learning element
    - Make improvements
  - Performance element (agent so far)
    - Select actions
  - Critic
    - Feedback on results
    - Used by learning element to determine how performance element should be modified to do better
  - Problem generator (for exploration)
    - Suggest actions that will lead to new and informative experiences



S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

Side Note: Current Research

- *Human-aware agent*
  - Interaction with a human during acting (*human-in-the-loop*)
  - Example: *Urban Search and Rescue* with a robot
- Different models considered
  - Original agent model $\mathcal{M}^R$
  - Human model $\mathcal{M}^H$
  - Agent model $\widetilde{\mathcal{M}}_r^H$ of human model $\mathcal{M}^H$
  - Agent model $\widetilde{\mathcal{M}}_h^R$ of human model $\mathcal{M}_h^R$, which human has of $\mathcal{M}^R$



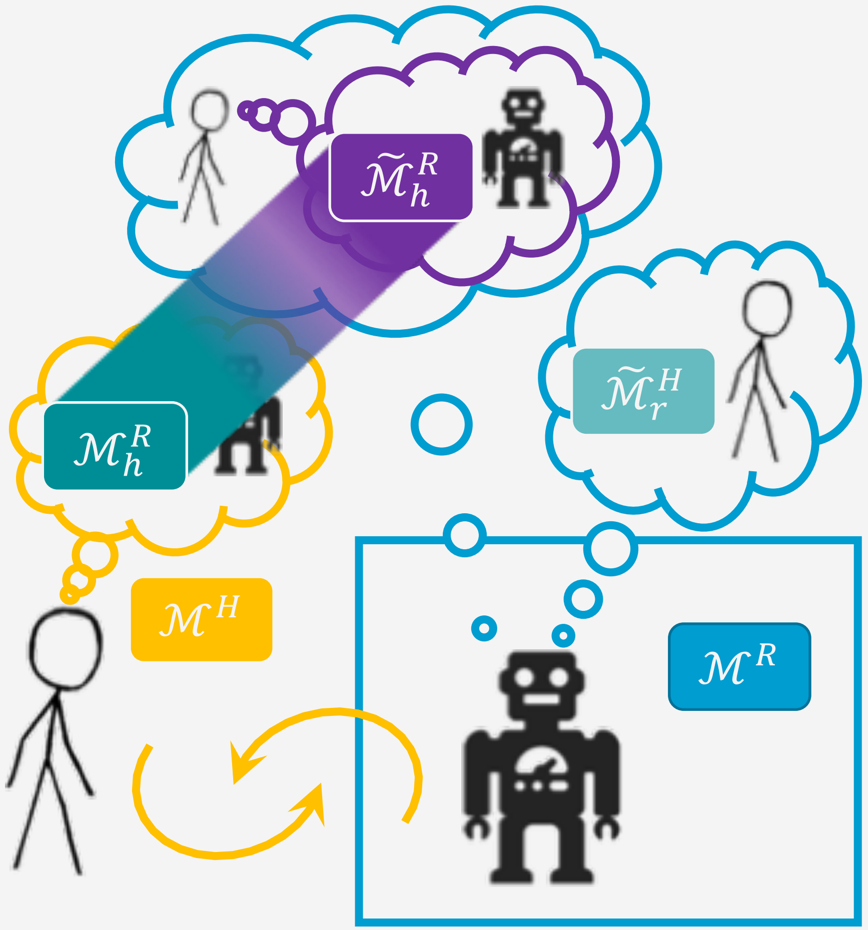
http://bit.ly/3bno2io

Side Note: Current Research

- Different models considered:
  - Original agent model $\mathcal{M}^R$
  - Human model $\mathcal{M}^H$
  - Agent model $\widetilde{\mathcal{M}}_r^H$ of human model $\mathcal{M}^H$
  - Agent model $\widetilde{\mathcal{M}}_h^R$ of human model $\mathcal{M}_h^R$, which human has of $\mathcal{M}^R$

- $\widetilde{\mathcal{M}}_r^H$ allows for anticipating human behaviour
- $\widetilde{\mathcal{M}}_h^R$ allows for conforming to human expectations

- See, e.g., Subbarao (Rao) Kambhampati
  http://rakaposhi.eas.asu.edu



Interim Summary

- Agent = architecture + agent program
- Task environment
  - Description: PEAS
    - Performance measure not easy to formalise
  - Properties of task environment
    - Fully observable (vs. partially observable); single agent (vs. multiple agents); deterministic (vs. stochastic), strategic; episodic (vs. sequential); static (vs. dynamic), semi-dynamic; discrete (vs. continuous); special: known (vs. unknown)
  - Environment encoding
    - Atomic, factorised, structured
- Agent structure
  - Simple / model-based reflex agent, goal- / utility-based agent; learning agent

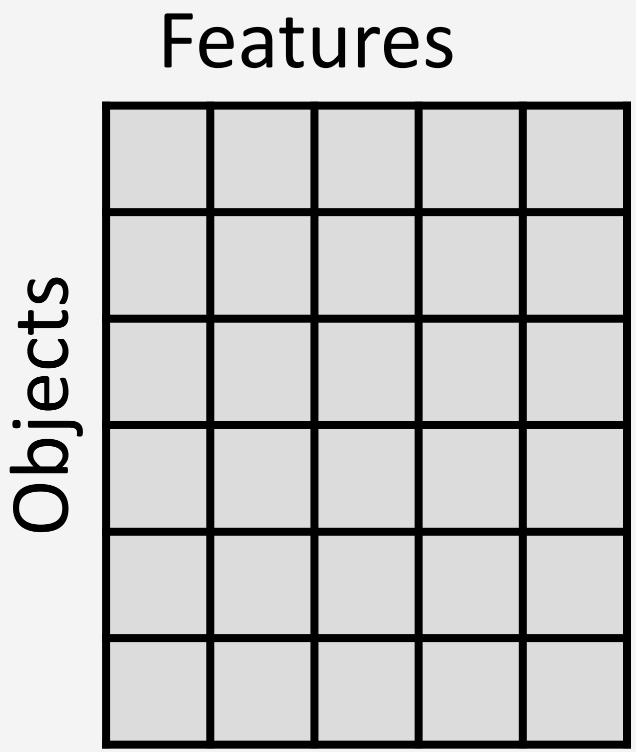
Overview: 1. Introduction

A. *Artificial Intelligence*
   - Approaches: thinking / acting humanly / rationally

B. *Framework: Agent Theory*
   - Agent
   - Task environment
   - Agent structure

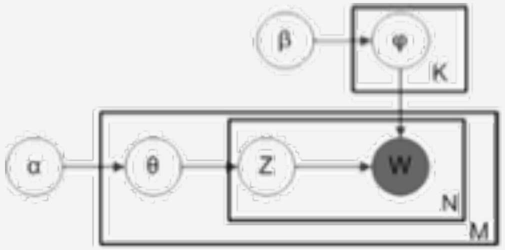
C. ***Topic: StaRAI***
   - Motivation, context
   - Relational examples, outlook on probabilistic relational models (PRMs)

Take Your Spreadsheet …



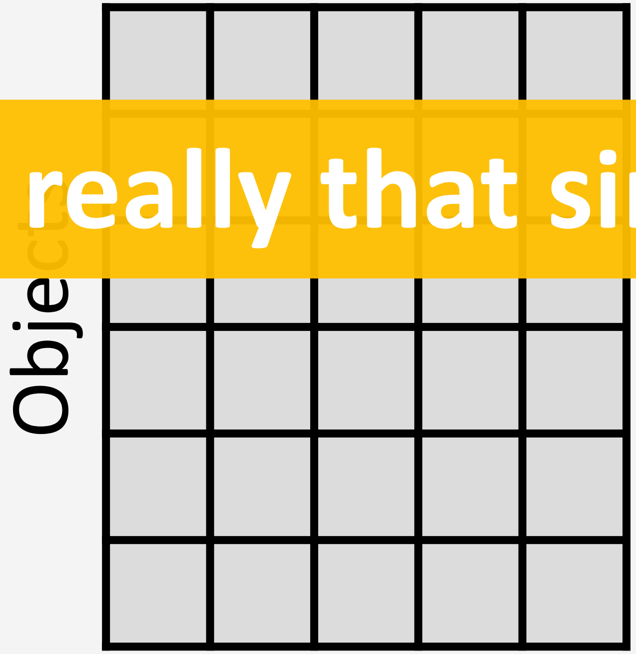
… and Apply Some AI/Machine Learning Methods

Big Data Matrix Factorization

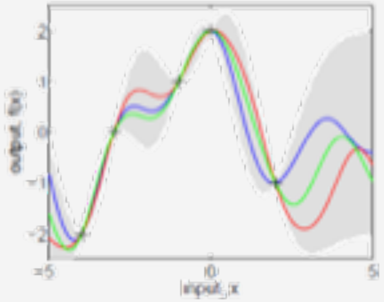


Latent Dirichlet Allocation

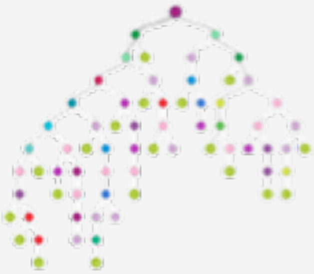
Features



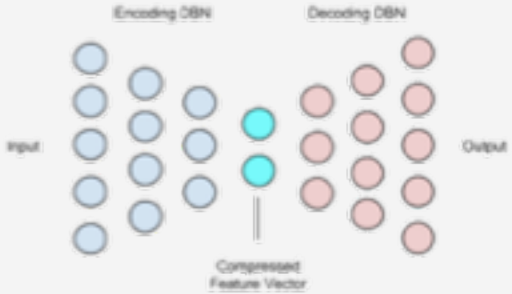
Is it really that simple?



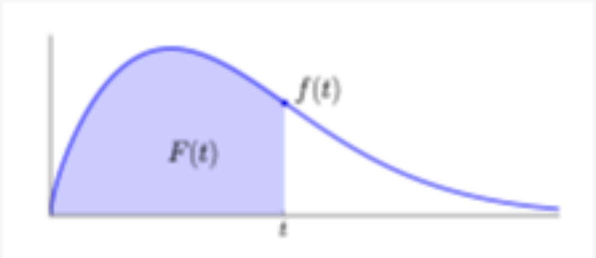
Gaussian Processes



Decision Trees/Boosting



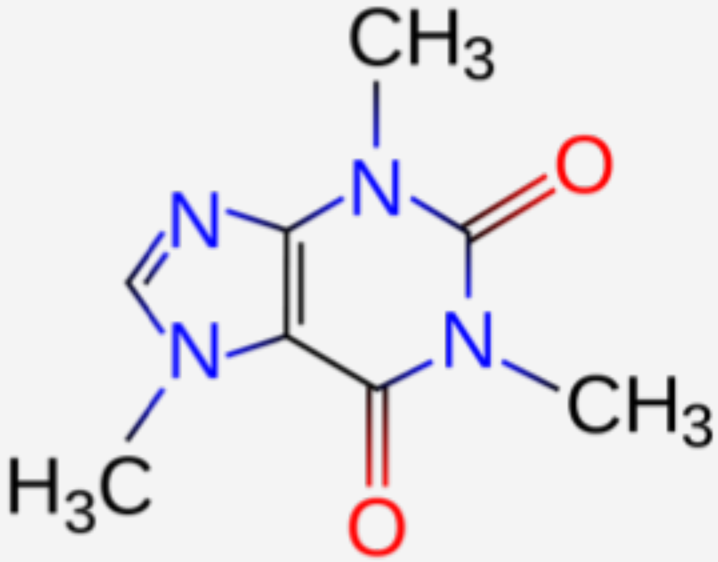
Autoencoder, Deep Learning

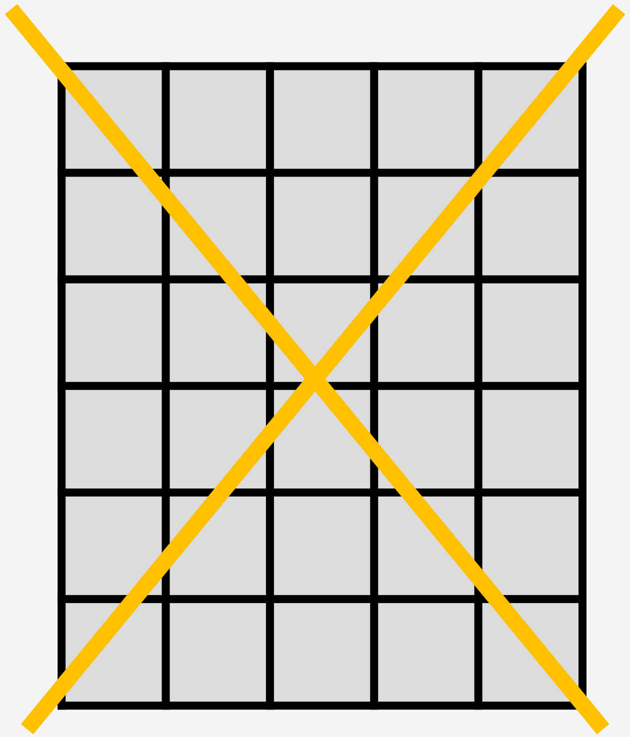


Diffusion Models

and many more …

Learning and Mining with Graphs



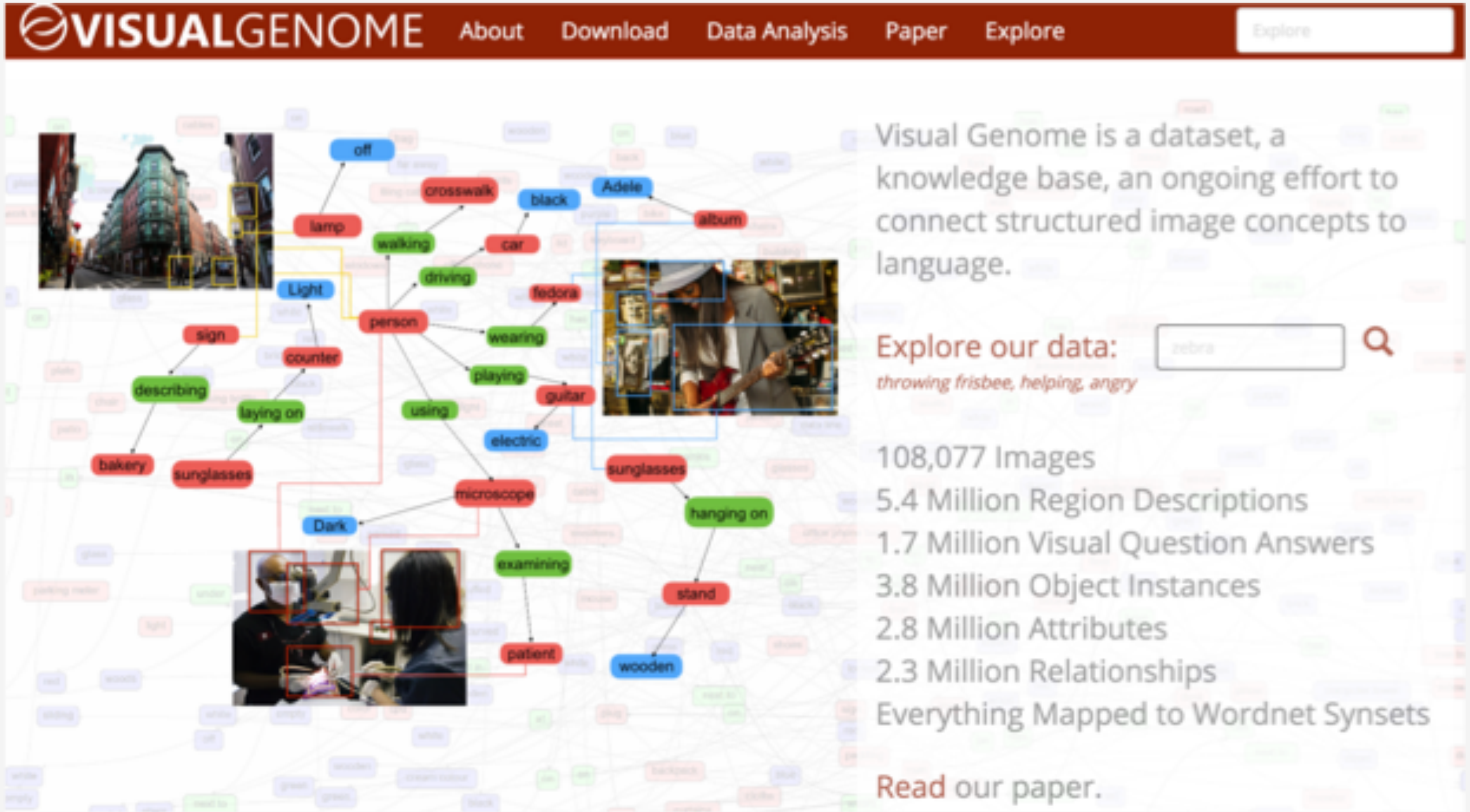


Haussler ’99, Gärtner, Flach, Wrobel COLT’03, Vishwanathan, Schraudolph, Kondor, Borgwardt JMLR’10, Shervashidze, Schweitzer, van Leeuwen, Mehlhorn, Borgwardt JMLR’11, Neumann, Garnett, Bauckhage, Kersting MLJ’16, Morris, Kersting, Mutzel, ICDM´17, and many more

Complex data networks!

[Lu, Krishna, Bernstein, Fei-Fei „Visual Relationship Detection“ CVPR 2016]

- Examples not stored in a single table but in a large, heterogenous graph with attributes!
  - Actually, most data in the world is stored in relational databases



Heart diseases and strokes – cardiovascular disease – are expensive for the world

According to the World Heart Federation, cardiovascular disease cost the European Union €169 billion in 2003 and the USA about €310.23 billion in direct and indirect annual costs. By comparison, the estimated cost of all cancers is €146.19 billion and HIV infections, €22.24 billion

Nat Rev Genet. 2012 May 2;13(6):395-405

Electronic Health Records
A New Opportunity for AI
to Save Our Lives

Connection to AI, Agents, and Environments

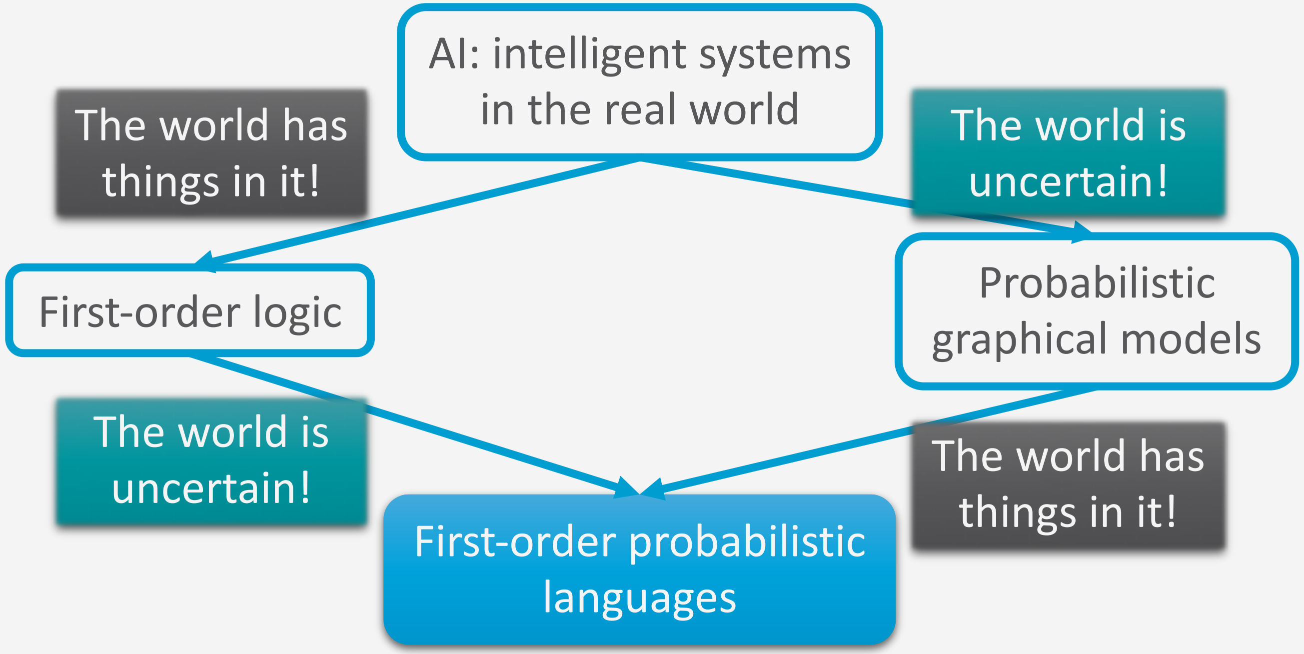


Figure based on Stuart Russell

Connection to AI, Agents, and Environments

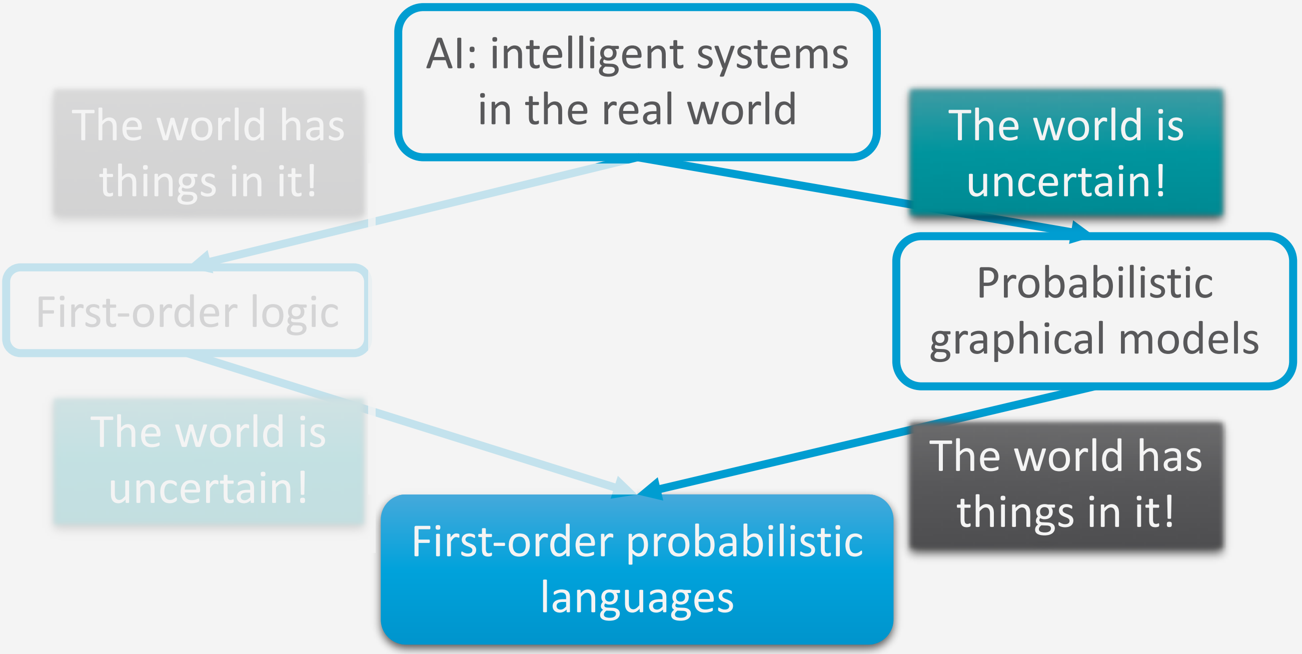
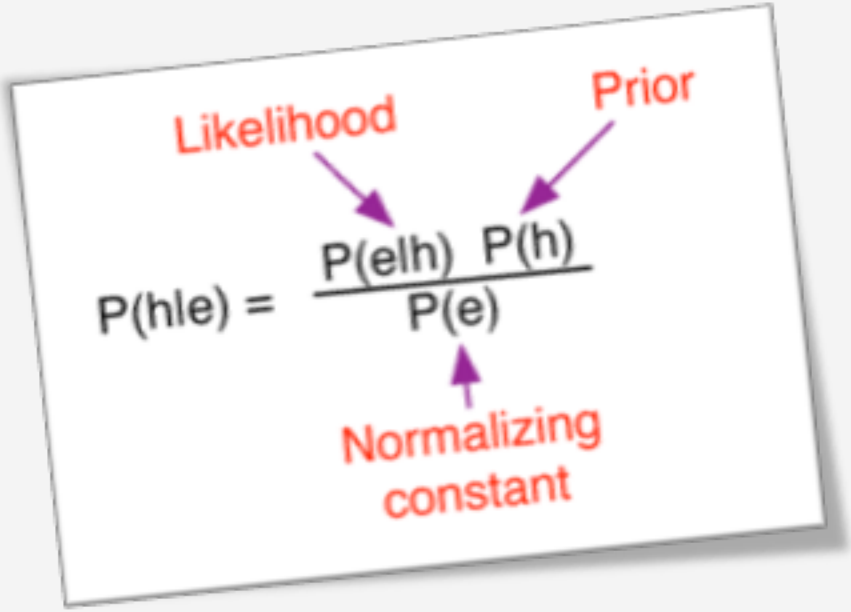


Figure based on Stuart Russell

Example: Bayes' Rule

- What if
  h is the effect of a drug on a particular patient, and
  e is the patient's electronic health record?

- What if e is the electronic health records for all of the people in the world?
- What if e is a collection of student records in a university?
- What if e is a description of everything known about the geology of Earth?



Example: Predicting Grades

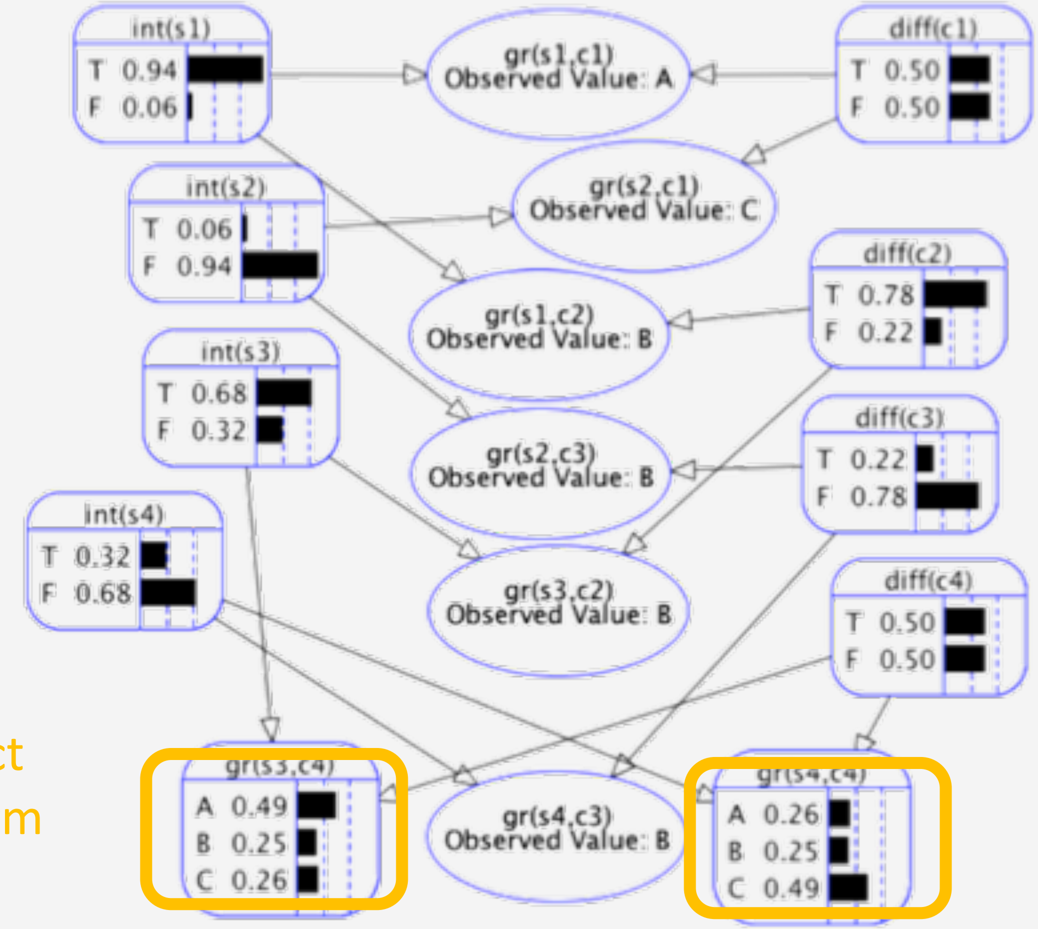
- Students s3 and s4 have the same averages, on courses with the same averages.
- Which student would you expect to do better?

Student	Course	Grade
s_1	c_1	A
s_2	c_1	C
s_1	c_2	B
s_2	c_3	B
s_3	c_2	B
s_4	c_3	B
s_3	c_4	?
s_4	c_4	?

Example: Predicting Grades

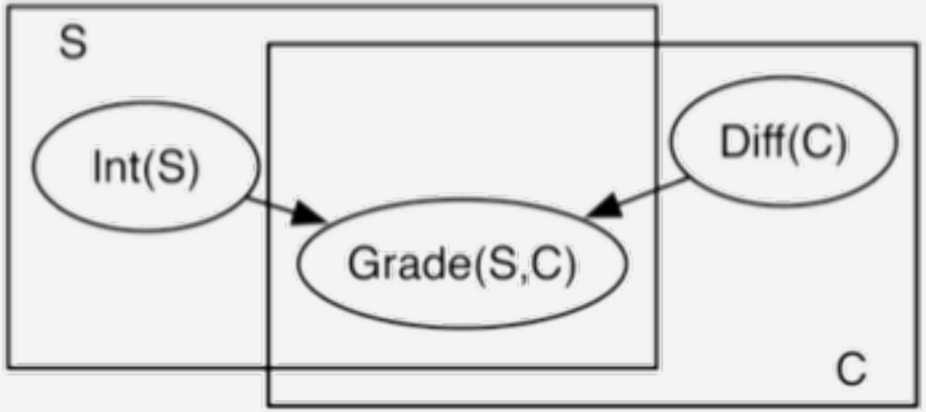
- Rigid and large graphical model
- Available features
  - Student s intelligence: $Int(s)$
    - Discrete (Boolean) range: T, F
  - Course c difficulty: $Diff(c)$
    - Discrete (Boolean) range: T, F
  - Student s grade in course c: $Gr(s, c)$
    - Discrete range: A, B, C

So, we should expect student s3 to perform better



Example: Predicting Grades

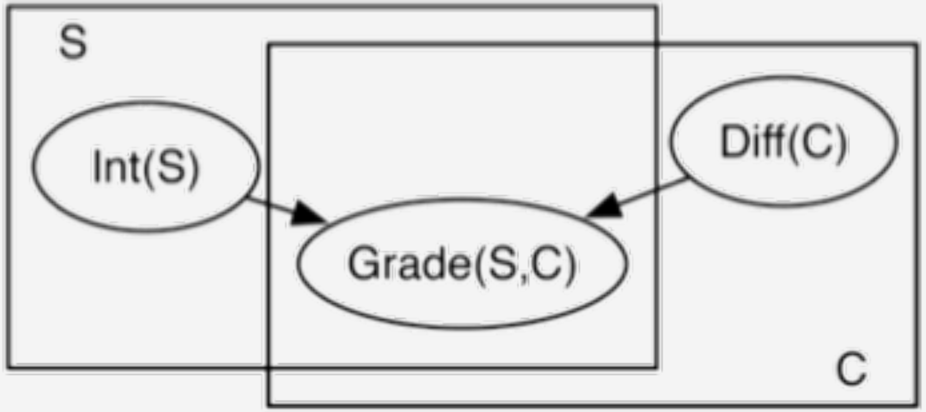
- Relational models: more flexible and compact way
- Program abstraction
  - S, C logical variable representing students, courses
  - Set of individuals of a type is called a domain or population
  - $Int(S), Grade(S, C), Diff(C)$ are parameterized random variables
- Grounding
  - for every student s, there is a random variable $Int(s)$
  - for every course c, there is a random variable $Diff(c)$
  - for every s, c pair there is a random variable $Grade(s, c)$
  - all instances share the same structure and parameters



Called plate notation, plates are pictured as boxes, denoting logical variables, types, groups

Example: Predicting Grades

- If there were 1000 students and 100 courses:
  - Grounding contains
    - 1000 $Int(s)$ variables
    - 100 $Diff(c)$ variables
    - 100000 $Grade(s,c)$ variables
  - **Total: 101100 variables**
- Numbers to be specified to define the probabilities = **10 parameters**
  - 1 for $Int(S)$,
  - 1 for $Diff(C)$,
  - 8 for $Grade(S,C)$
    - Idea of parfactor models as we will see later



$$P_v = P(Gr(S,C) = v | Int(S), Diff(C))$$

$Int(S)$	$Diff(C)$	P_A	P_B	P_C
$false$	$false$	p_{11}	p_{12}	$1 - p_{11} - p_{12}$
$false$	$true$	p_{21}	p_{22}	$1 - p_{21} - p_{22}$
$true$	$false$	p_{31}	p_{32}	$1 - p_{31} - p_{32}$
$true$	$true$	p_{41}	p_{42}	$1 - p_{41} - p_{42}$

Not necessary due to probability distributions adding to 1

Connection to AI, Agents, and Environments

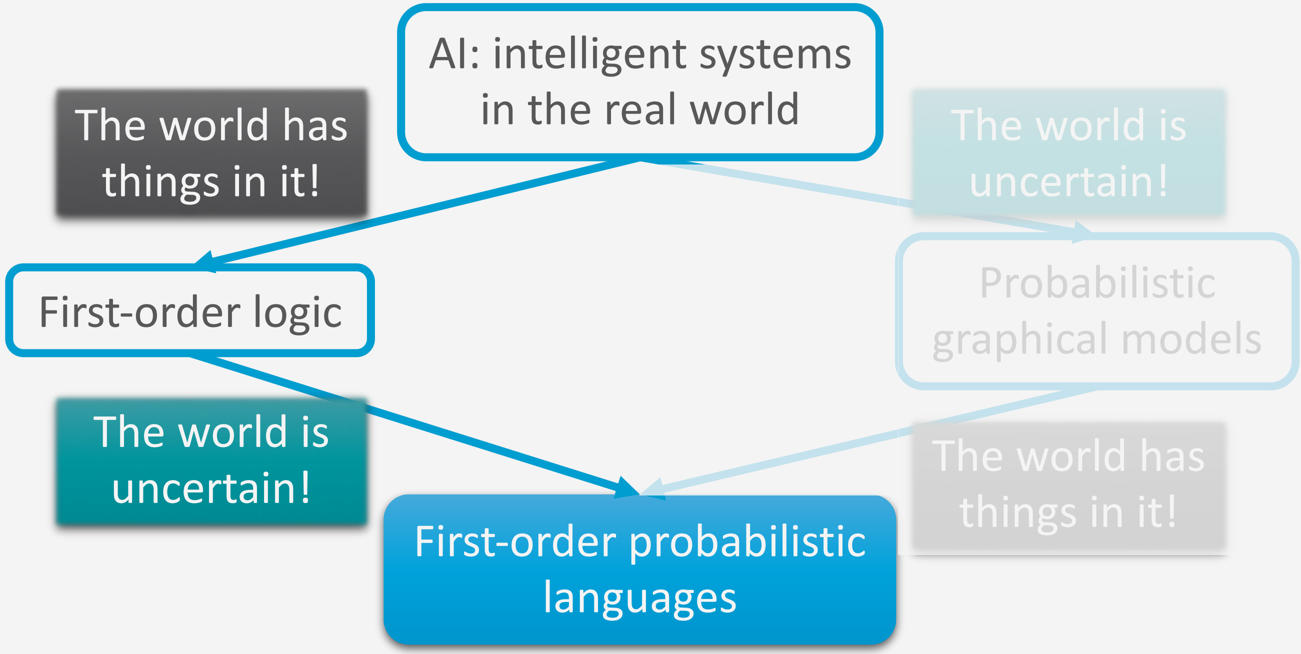


Figure based on Stuart Russell

From Logics to Probabilistic Relational Models

- Propositional: Descriptions about world
  - Form of constraints on an environment

$$Presents \Rightarrow Attends$$

$$Presents(eve, paper1, IJCAI) \Rightarrow Attends(eve, IJCAI)$$

- First-order: Objects, relations among them
  - Groundings to get to propositional logic

$$Presents(X, P, C) \Rightarrow Attends(X, C)$$

- Denote either true or false statements

- Approach to soften constraints: Introduce weights to denote that statements hold to a certain degree over all possible worlds
  - Example: Markov Logic Network
    - We will see more of them later

Hard constraint

$$\infty \; Presents(X, P, C) \Rightarrow Attends(X, C)$$

$$3.75 \; Publishes(X, C) \land FarAway(C) \Rightarrow Attends(X, C)$$

Soft constraint,
weight = $\exp(3.75)$

Connection to AI, Agents, and Environments

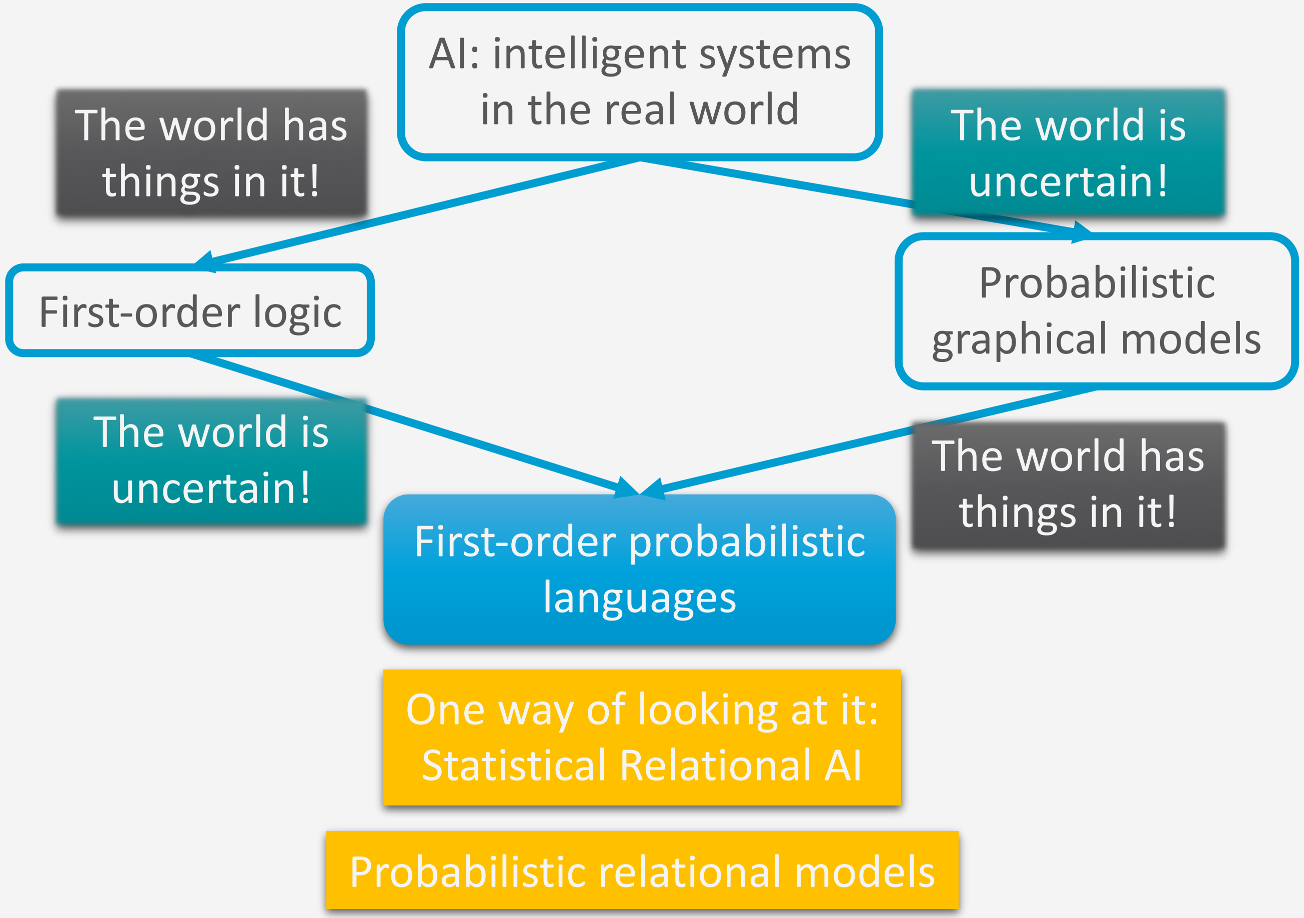
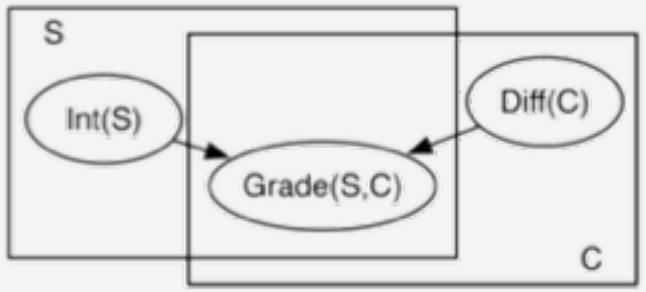


Figure based on Stuart Russell

Probabilistic Relational Models (PRMs)



$\infty\ Presents(X, P, C) \Rightarrow Attends(X, C)$

$3.75\ Publishes(X, C) \land FarAway(C) \Rightarrow Attends(X, C)$

- Random variables for combinations of individuals in populations
  - Build a probabilistic model before knowing (all of) the individuals
  - Learn the model for one set of individuals
  - Apply the model to existing and new individuals
  - Allow complex relationships between individuals
- Exchangeability:
  - Before we know anything about individuals, they are *indistinguishable*, and so should be treated identically.
- Uncertainty about:
  - Properties of individuals
  - Relationships among individuals
  - Identity (equality) of individuals
  - Existence (and number) if individuals
- Depicted formalisms: Parfactor graphs, Markov logic networks

PRMs for Modelling the Environment in this Lecture

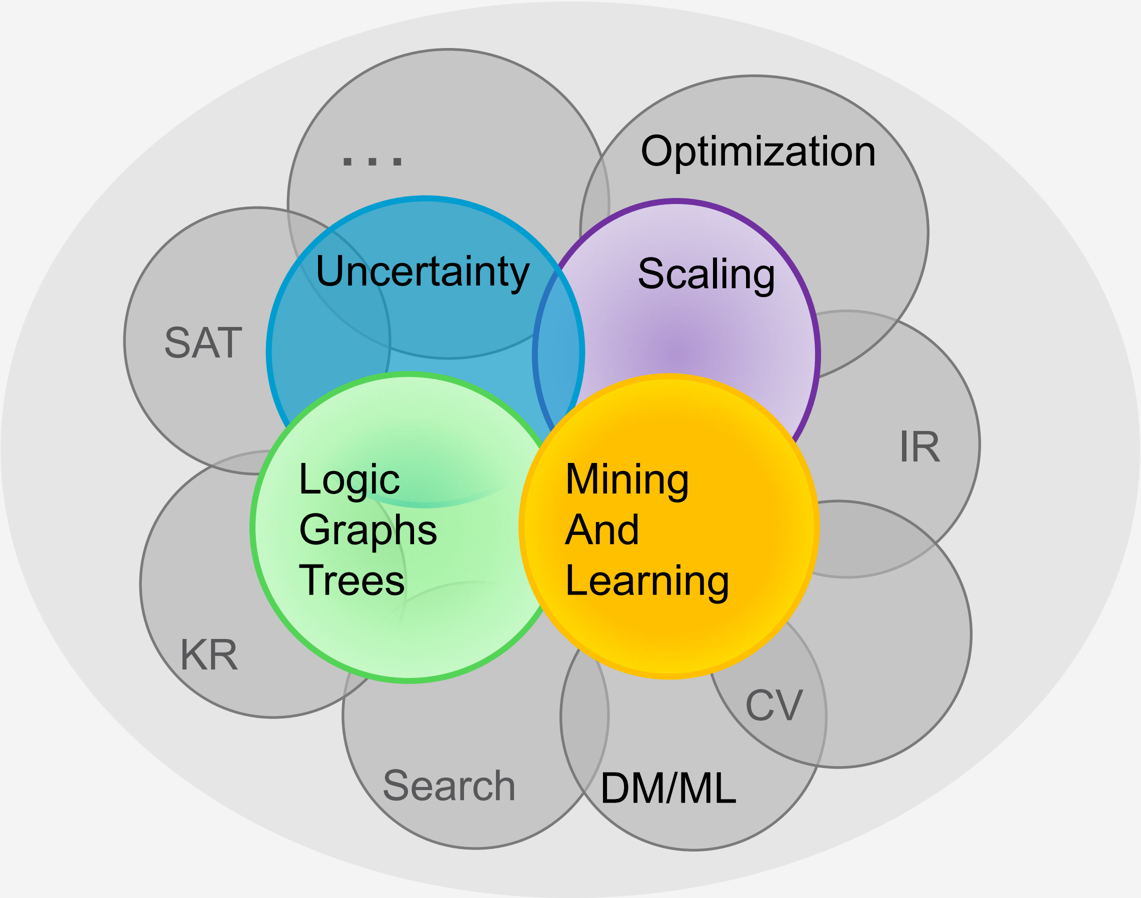
- Different types of PGMs for a (relational) factorised representation of the environment and decision making
- Possible to model environment with following properties
  - Fully or partially observable
  - Single agent
  - Stochastic
  - Episodic or sequential
  - Static
  - Discrete or continuous
- Not considered in this lecture
  - Multiple agents
  - Deterministic, strategic
  - Dynamic
- Approaches exist to deal with such environments to a certain degree

Attention: In the PGM literature, the notions of *static* and *dynamic* are used instead of the notions of *episodic* and *sequential*, while the notions of *static* and *dynamic* are covered by the so-called *Markov assumption* or *(Non-) Markovian abstraction*

S. Russel and P. Norvig: Artificial Intelligence – A Modern Approach, 2020.

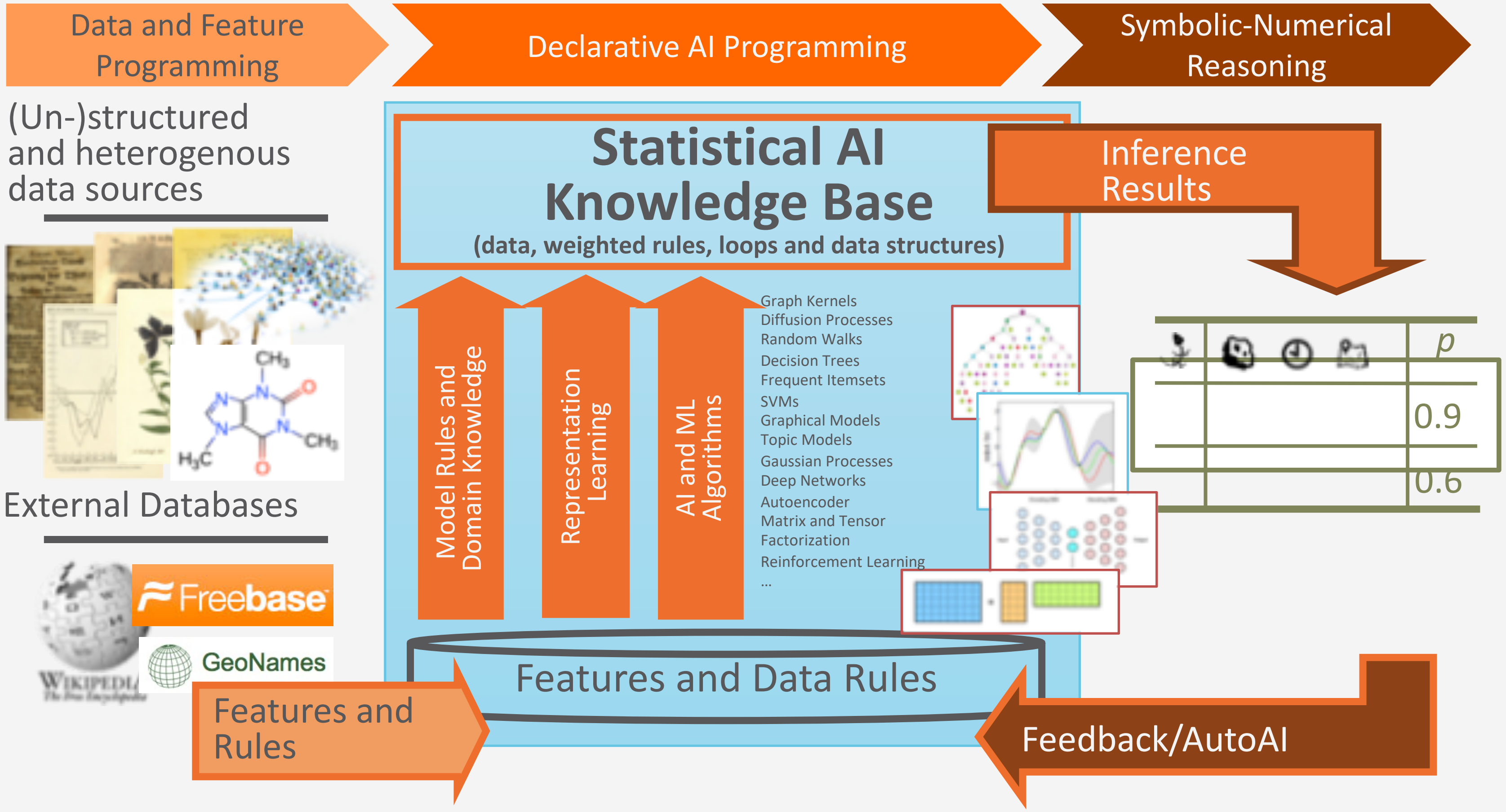
The Larger Scope: Statistical Relational Learning (SRL) & StaRAI

- Study and design
  - intelligent agents
  - that reason about and
  - act in noisy worlds
  - composed of objects and relations among the objects



[Getoor, Taskar MIT Press ’07; De Raedt, Frasconi, Kersting, Muggleton, LNCS’08; Domingos, Lowd Morgan Claypool ’09; Natarajan, Kersting, Khot, Shavlik Springer Brief’15; Russell CACM 58(7): 88-97 ’15, Gogate, Domingos CACM 59(7):107-115 ´16]

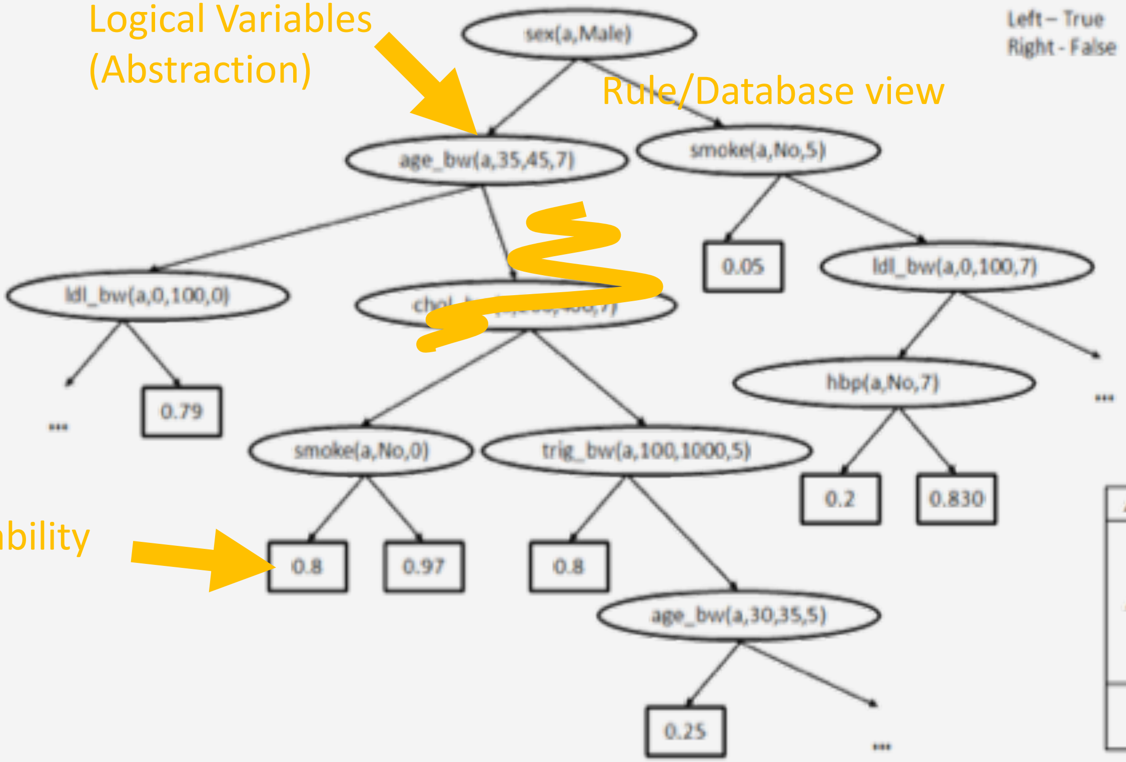
This Establishes a Novel “Deep AI”

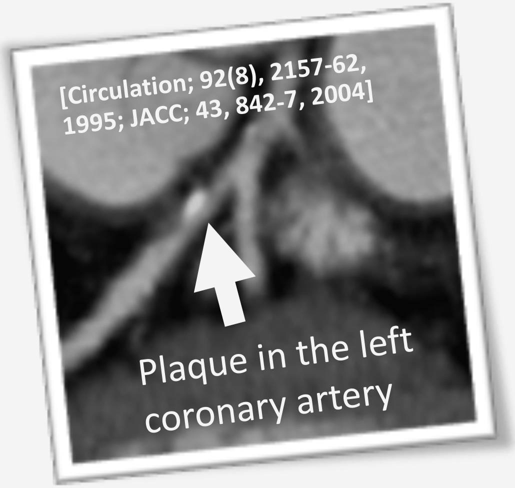


[Ré, Sadeghian, Shan, Shin, Wang, Wu, Zhang IEEE Data Eng. Bull.’14; Natarajan, Picado, Khot, Kersting, Ré, Shavlik ILP’14; Natarajan, Soni, Wazalwar, Viswanathan, Kersting Solving Large Scale Learning Tasks’16, Mladenov, Heinrich, Kleinhans, Gonsior, Kersting DeLBP’16, Kordjamshidi, Roth, Kersting IJCAI-ECAI 2018, …]

This “Deep AI” Can Understand EHRs

Atherosclerosis is the cause of the majority of Acute Myocardial Infarctions (heart attacks)





Algorithm	Accuracy	AUC-ROC
J48	0.667	0.607
SVM	0.667	0.5
AdaBoost	0.667	0.608
Bagging	0.677	0.613
NB	0.75	0.653
RPT	0.669*	0.778
RFGB	0.667*	0.819

The higher, the better

25%

We will come back to this later in the lecture.

Natarajan, Khot, Kersting, Shavlik. Boosted Statistical Relational Learners. Springer Brief 2015

[Kersting, Driessens ICML´08; Karwath, Kersting, Landwehr ICDM´08; Natarajan, Joshi, Tadepelli, Kersting, Shavlik. IJCAI´11; Natarajan, Kersting, Ip, Jacobs, Carr IAAI `13; Yang, Kersting, Terry, Carr, Natarajan AIME ´15; Khot, Natarajan, Kersting, Shavlik ICDM´13, MLJ´12, MLJ´15]

Algorithm for Mining Markov Logic Networks	Likelihood The higher, the better	AUC-ROC The higher, the better	AUC-PR The higher, the better	Time The lower, the better
Boosting	**0.81**	**0.96**	**0.93**	9s
LSM	0.73	0.54	0.62	93 hrs

11% 78% 50% 37200x faster

Kristian Kersting. Statistical Relational AI. Tutorial at KI 2018.

Heart diseases and strokes – cardiovascular disease – are expensive for the world

According to the World Heart Federation, cardiovascular disease cost the European Union €169 billion in 2003 and the USA about €310.23 billion in direct and indirect annual costs. By comparison, the estimated cost of all cancers is €146.19 billion and HIV infections, €22.24 billion

Nat Rev Genet. 2012 May 2;13(6):395-405

Human-centred approaches needed to ensure ethical behaviour, transparent reasoning, explainable results, privacy, …

Electronic Health Records
A New Opportunity for AI
to Save Our Lives

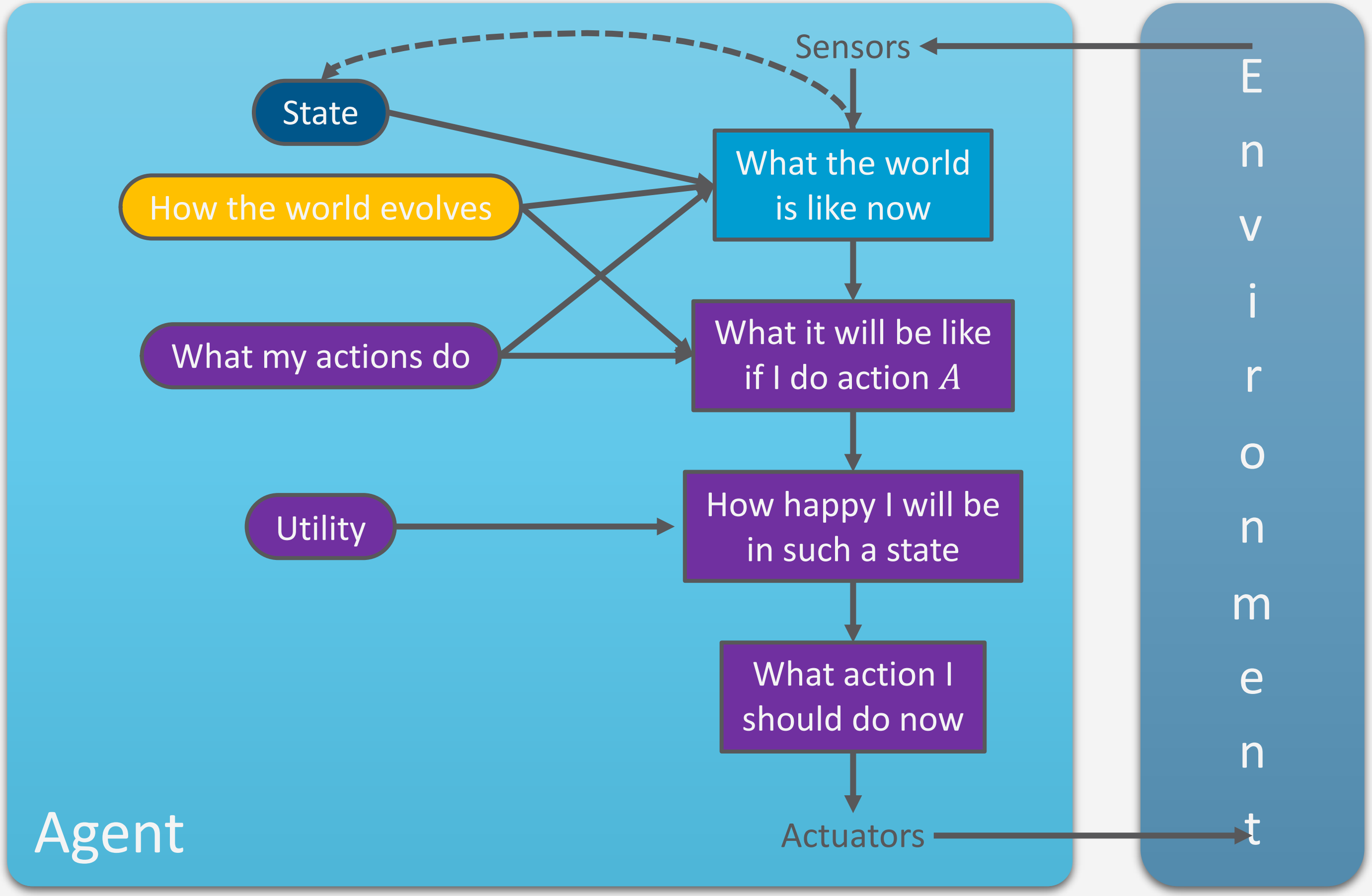
Interim Summary

- Environment characterised by
  - Objects and relations between them
  - Uncertainty
- PRMs combine both logic and probability theory
  - Models covered here can model the following environment properties
    - Fully or partially observable
    - Single agent
    - Stochastic
    - Episodic or sequential
    - Static
    - Discrete or continuous

Contents

Contents in this Lecture Related to *Utility-based Agents*

- Further topics
  3. (Episodic) PRMs
  4. Lifted inference (in episodic PRMs)
  5. Lifted learning (of episodic PRMs)
  6. Lifted sequential PRMs and inference
  7. Lifted decision making
  8. Continuous space and lifting



Overview: 1. Introduction

A. *Artificial Intelligence*
  - Approaches: thinking / acting humanly / rationally

B. *Framework: Agent Theory*
  - Agent
  - Task environment
  - Agent structure

C. *Topic: StaRAI*
  - Motivation, context
  - Relational examples, outlook on probabilistic relational models (PRMs)

→ Foundations: Logic, Probability Theory, & PGMs