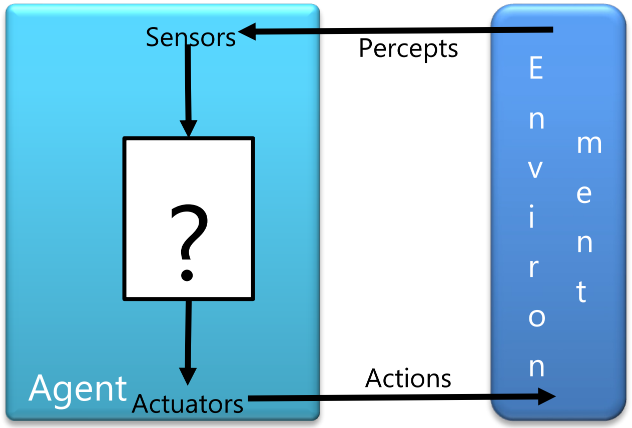
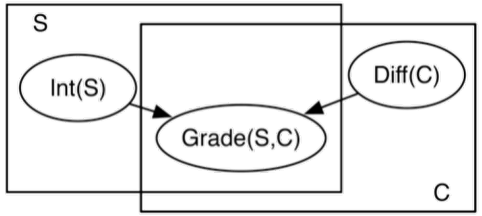
UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR INFORMATIONSSYSTEME

Dynamic Probabilistic Relational Models

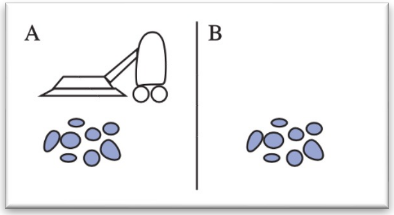


Introduction



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

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



Marcel Gehrke

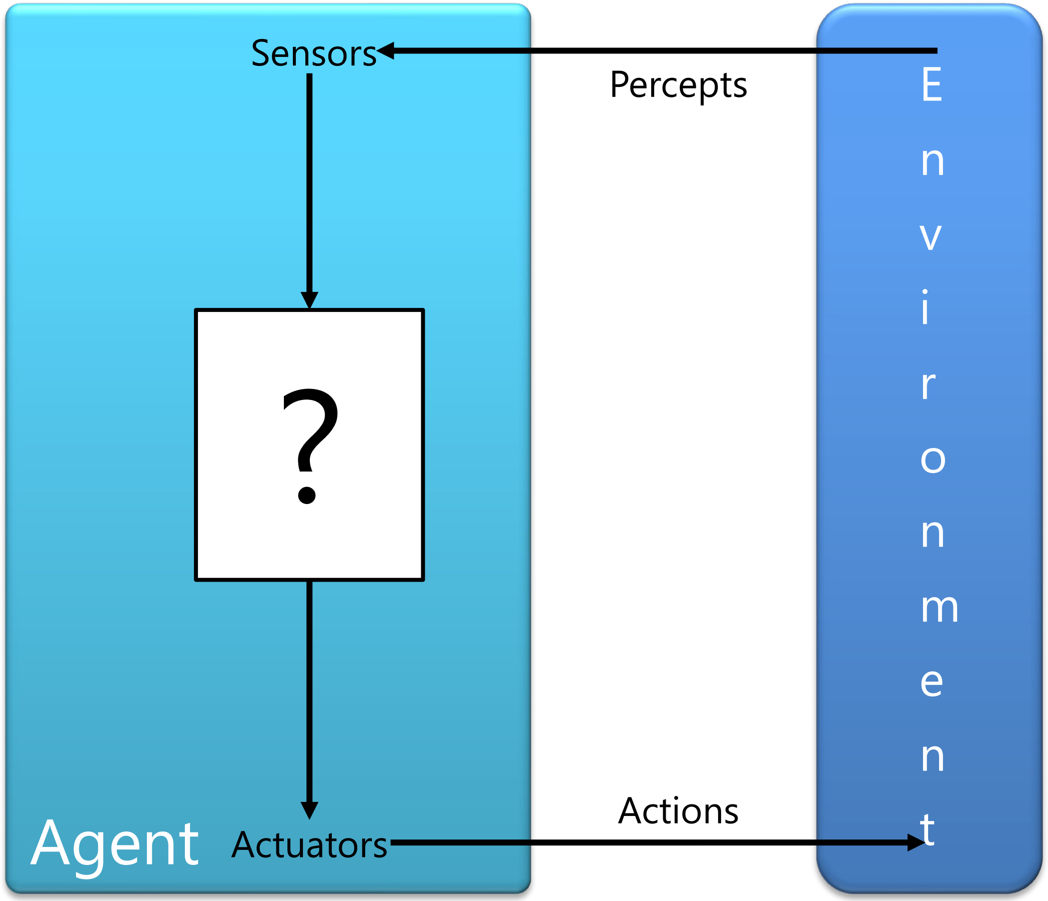
Contents

Overview: 1. Introduction

A. ***Framework: Agent Theory***
   - Agent
B. *Topic: StaRAI*
   - Motivation, context
   - Relational examples, outlook on probabilistic relational models (PRMs)

Agent

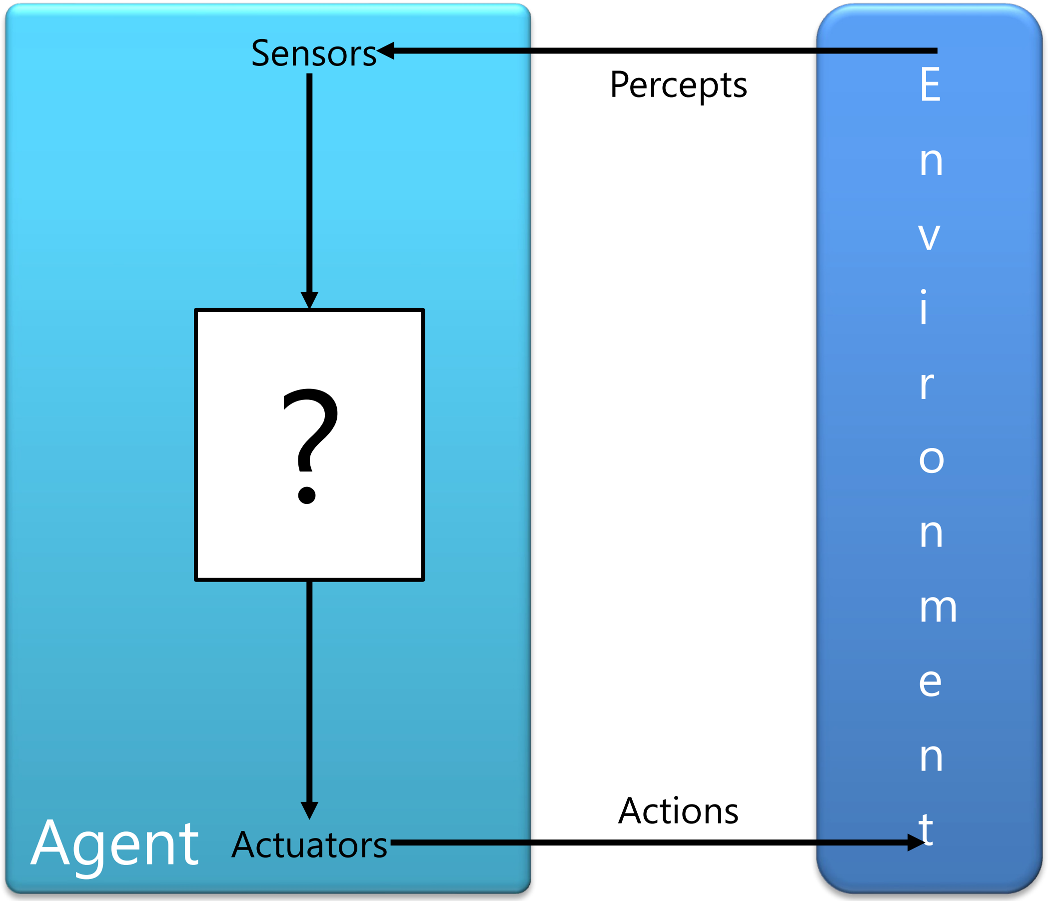
- Something that perceives its environment through sensors and acts through actuators
  - Human agent
    - Sensors: eyes, ears, further organs
    - Actuators: hands, legs, mouth, other body parts
  - Robot agent
    - Sensors: cameras, infra-red sensors, etc.
    - Actuators: motors
  - Software agent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



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

Performance Measure: Good Behaviour

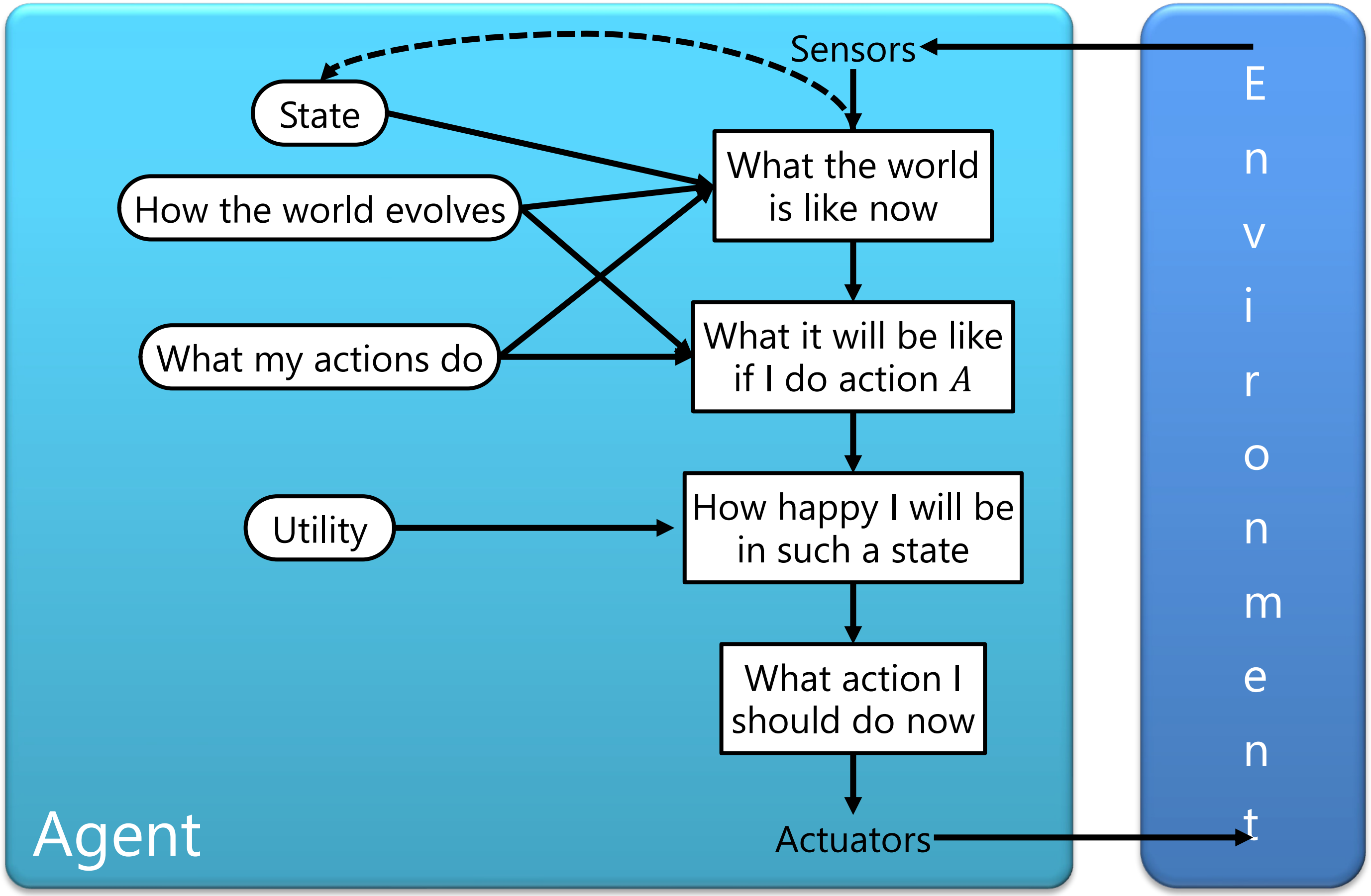
- What is the right thing to do?
  - Agent generates sequence of actions given percepts
  - Causes environment to go through sequence of states
  - If sequence desirable / conforms to our expectations, then agent performed well
- Performance measure that evaluates any given sequence of environment states



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

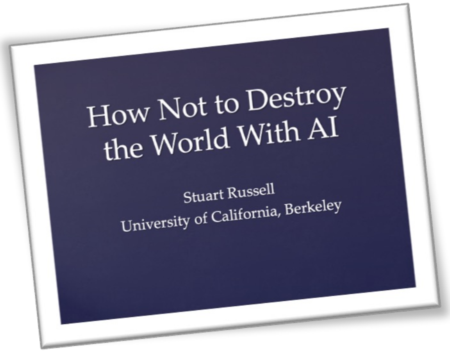
Agent Structure: Utility-based Agent

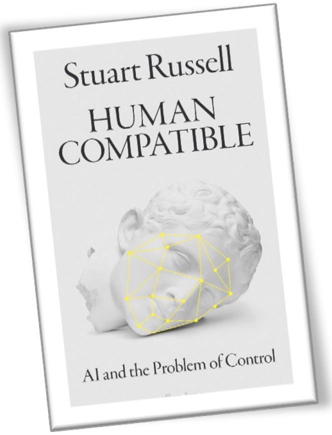
- Utility as a distribution over possible states
  - What we look at later in the lecture
  - Essentially an internalisation of the performance measure
    - If internal utility function *agrees with* external performance measure:
    - Agent that chooses actions to maximize its utility will be *rational*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

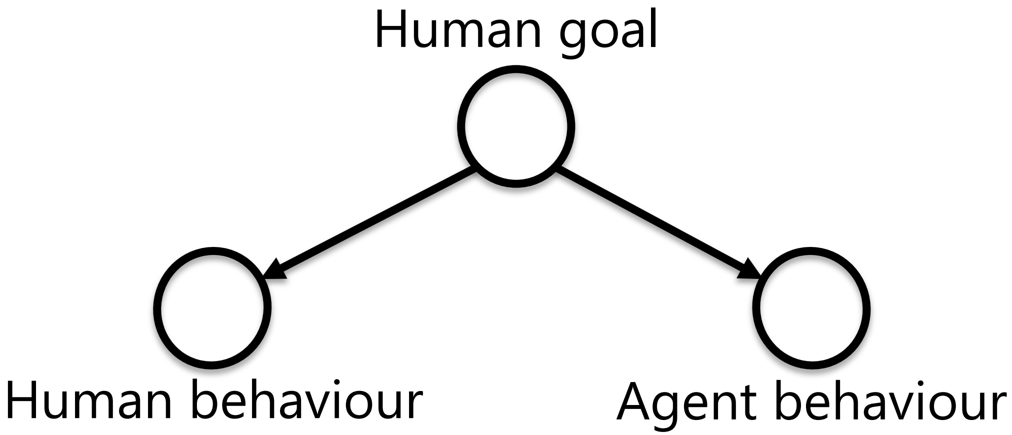


Side Note: Current Research

- Idea:
    - Humans: intelligent to the extent that our actions can be expected to achieve our goals
    - ~~Machines: intelligent to the extent that their actions can be expected to achieve their goals~~
    - Machines 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





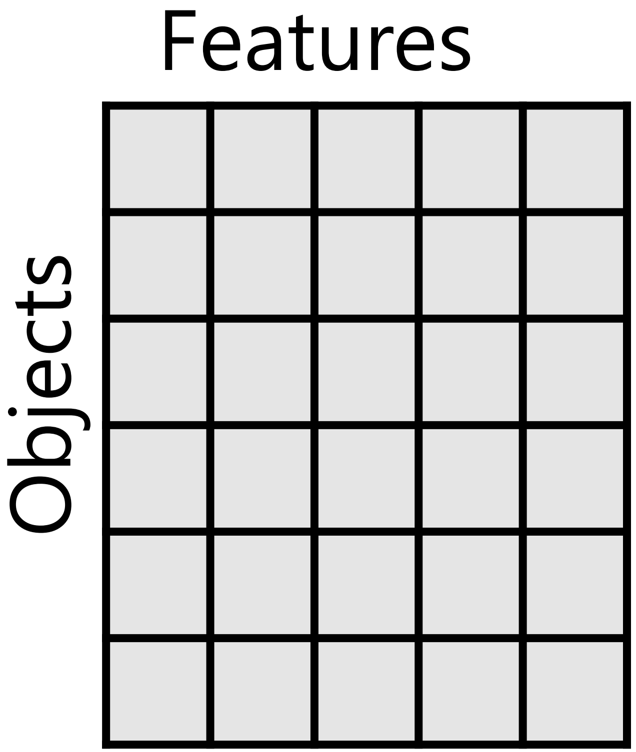


Overview: 1. Introduction

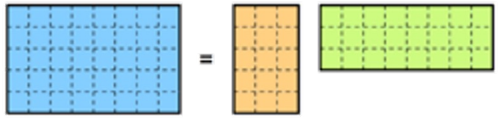
A. *Framework: Agent Theory*
   - Agent

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

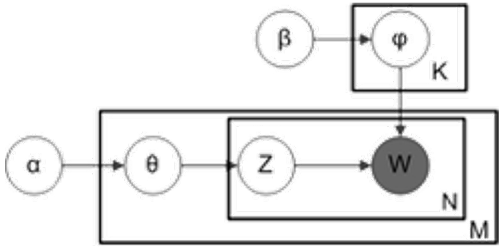
Take Your Spreadsheet …



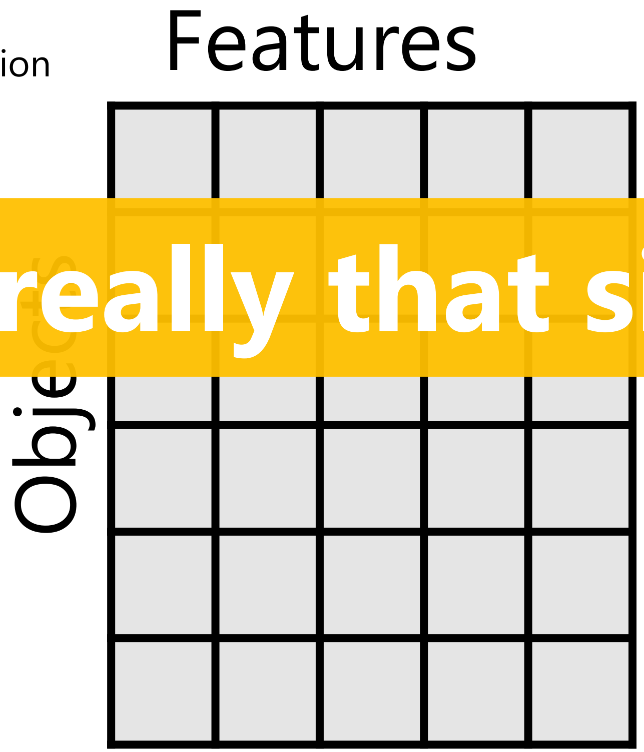
... and Apply Some AI/Machine Learning Methods

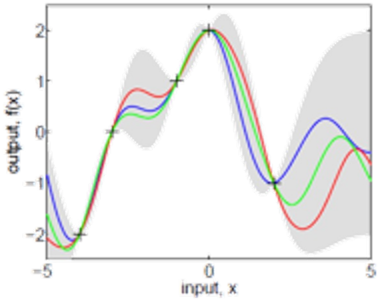


Big Data Matrix Factorization



Latent Dirichlet Allocation

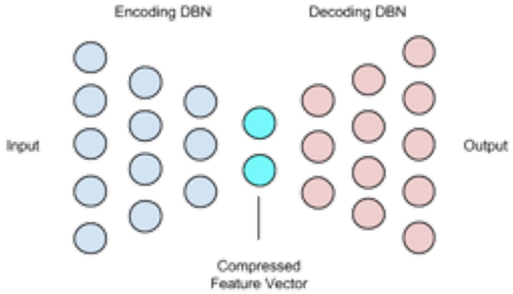




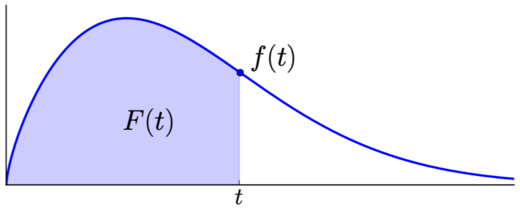
Gaussian Processes

Is it really that simple?

Decision Trees/Boosting



Autoencoder, Deep Learning

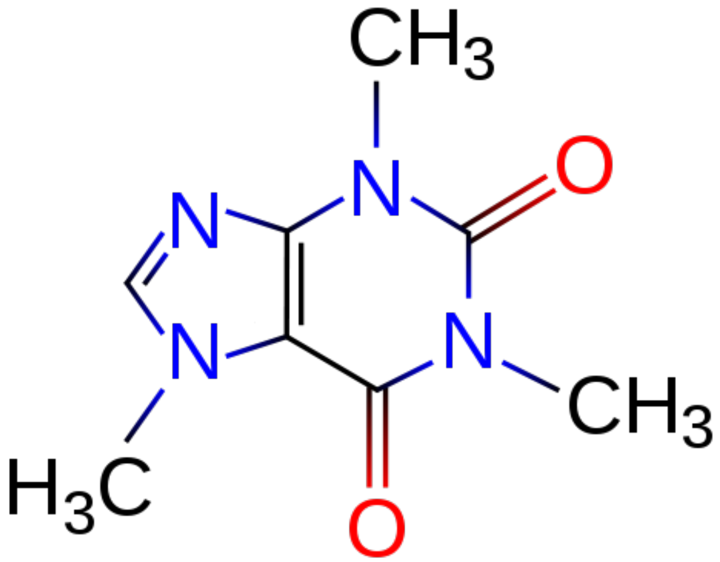


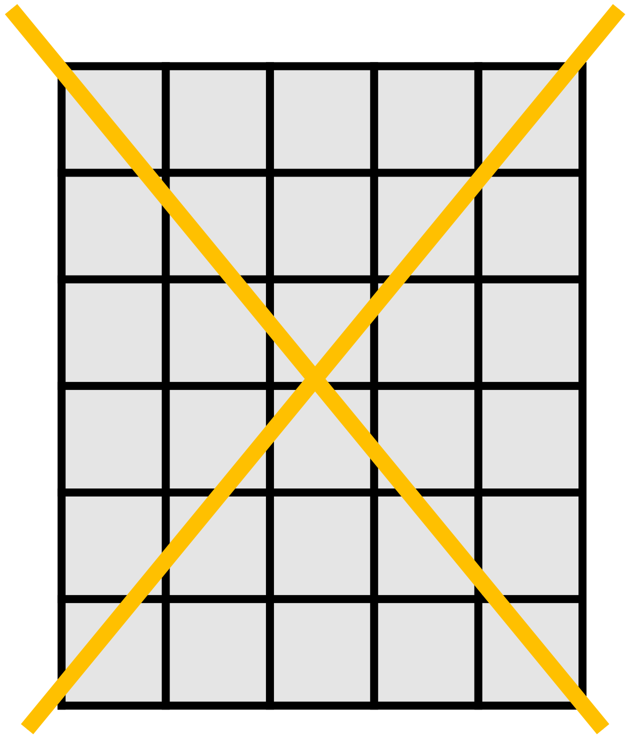
Diffusion Models

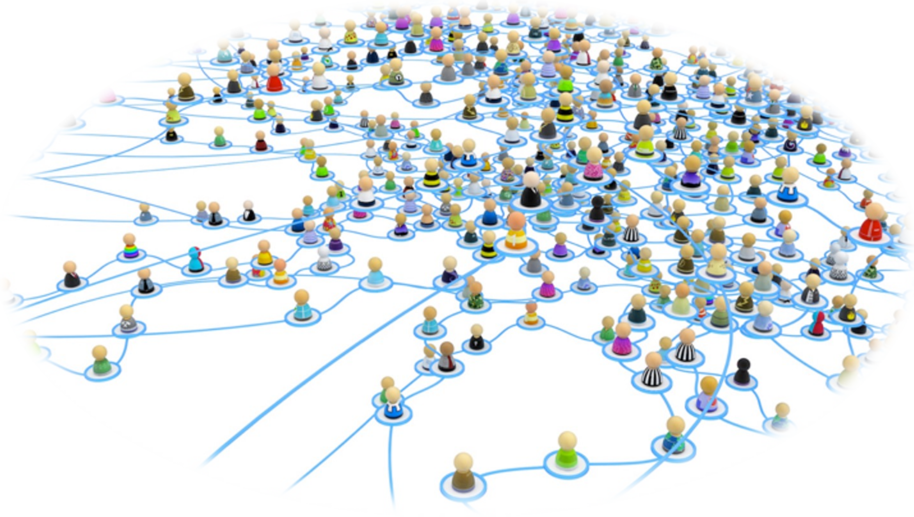
and many more ...

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



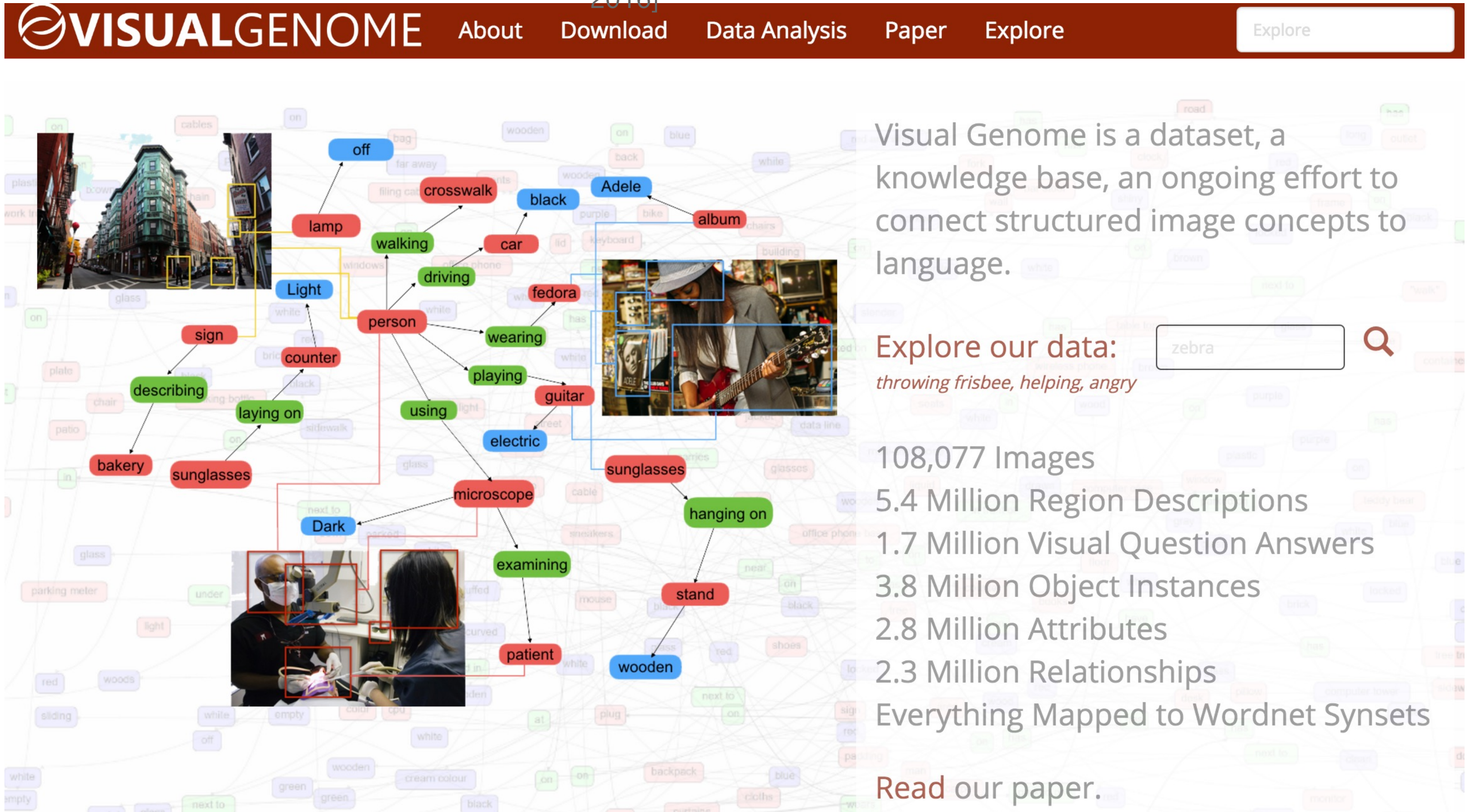




Complex data networks!

- 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

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



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

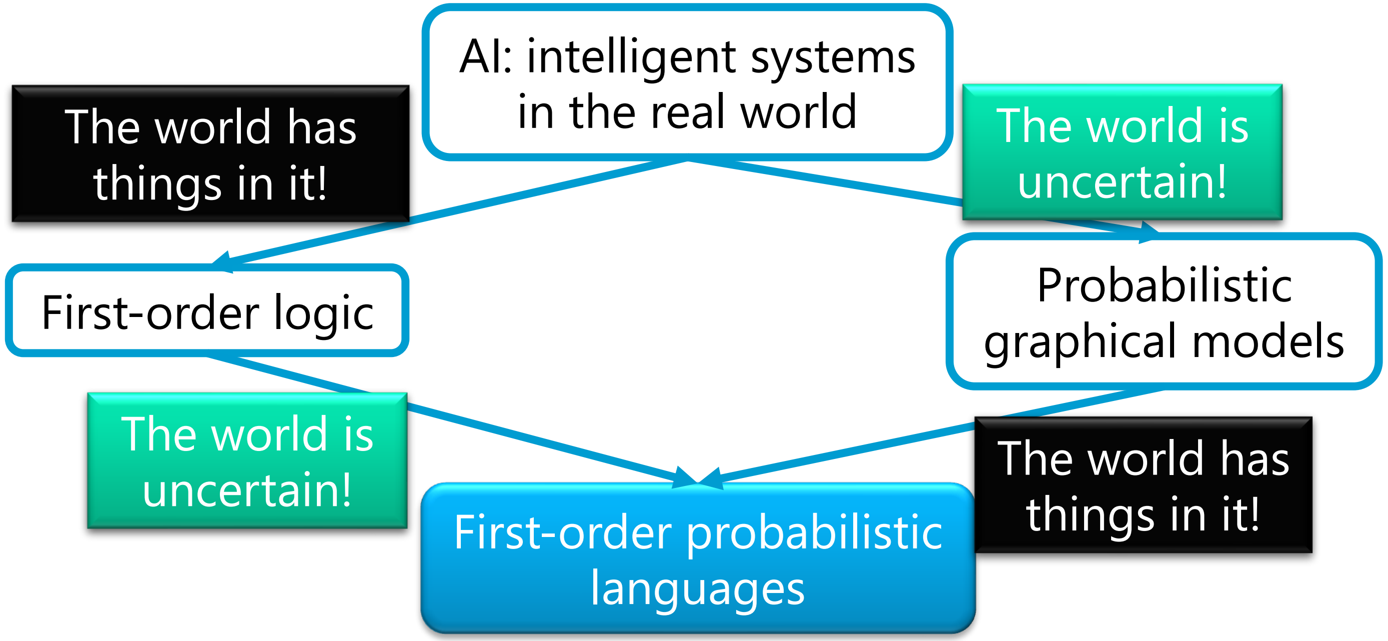
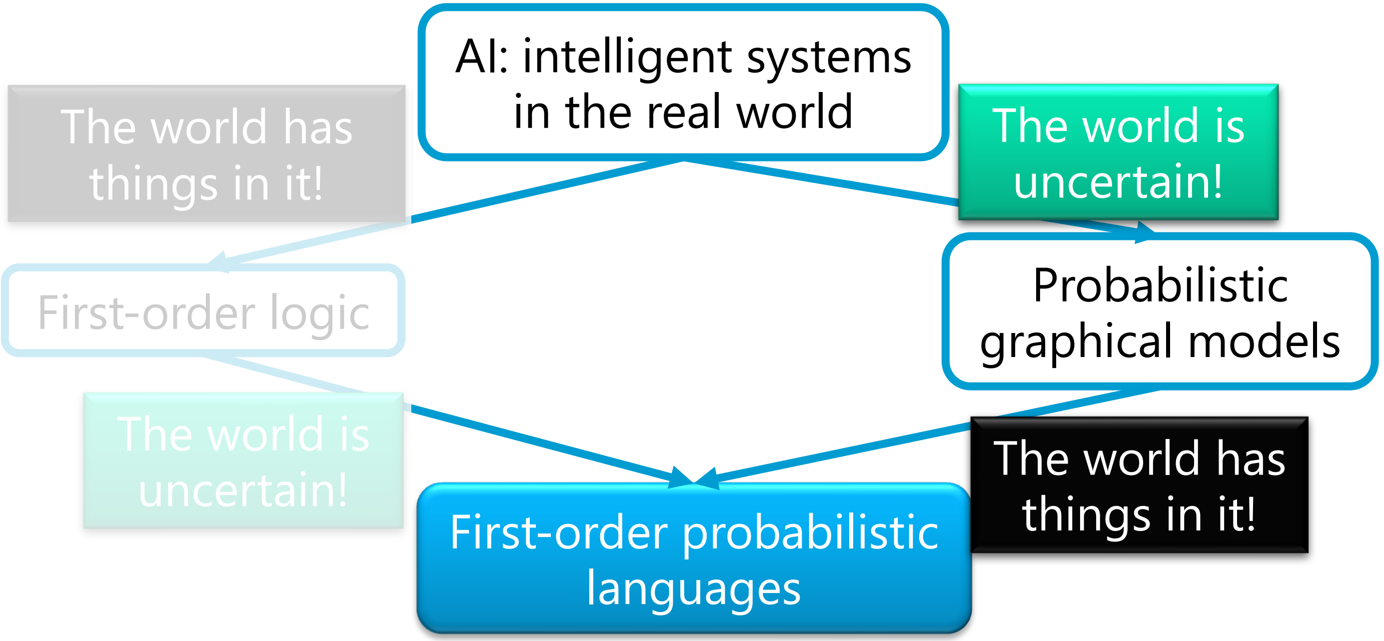


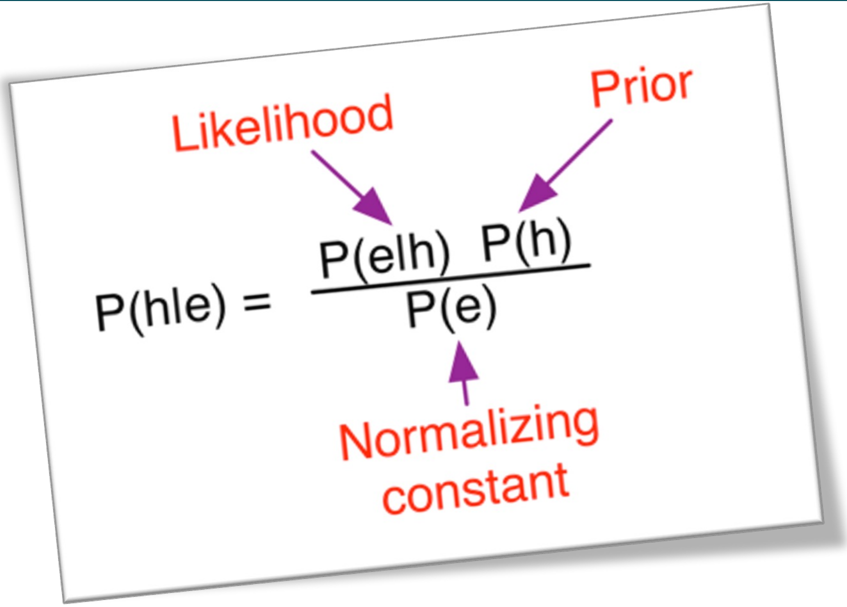
Figure based on Stuart Russell

Connection to AI, Agents, and Environments



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?

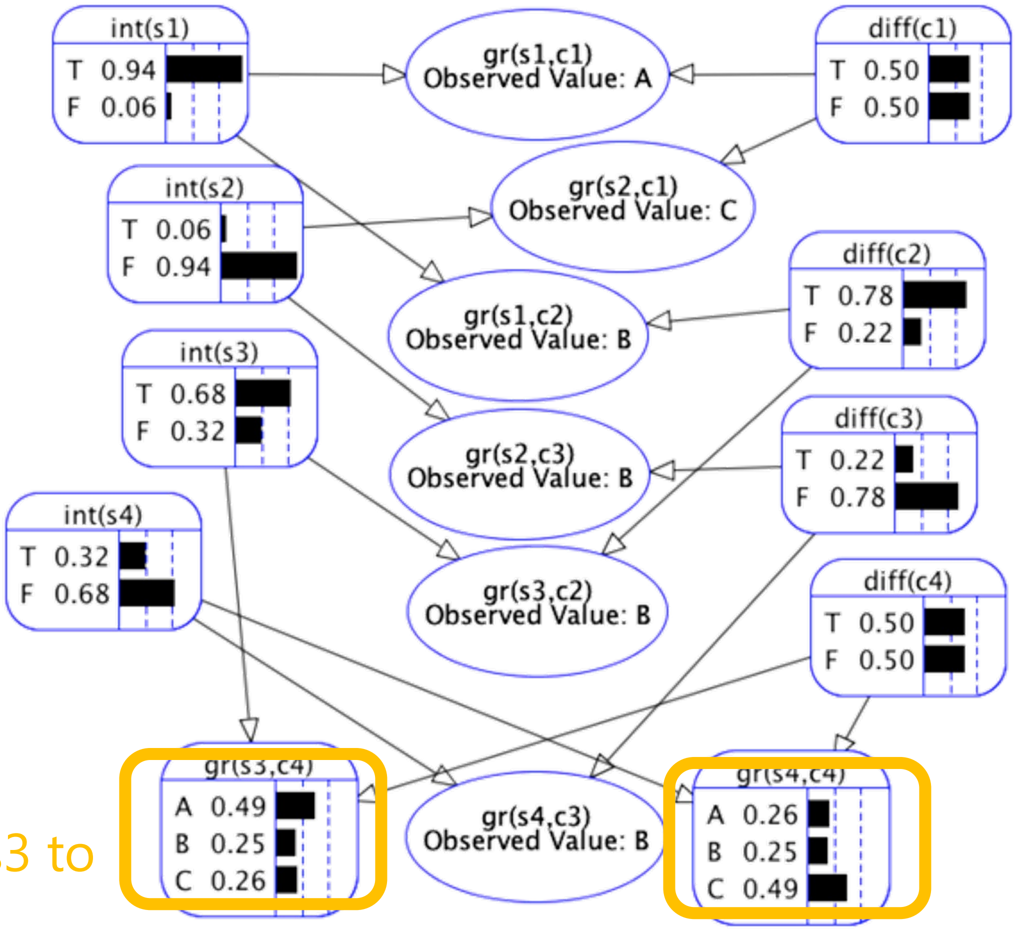
Example: Predicting Grades

- 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

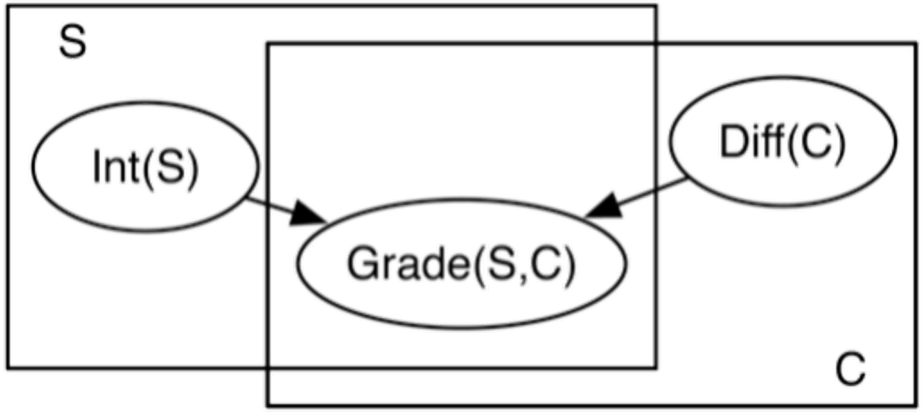
- Rigid and large graphical model
- Available features
  - Student s intelligence: $Int(s)$
    - Discrete (Boolean) range: T, F
  - Course c difficulty: $Diff(c)$
    - Discrete (Boolean) range: T, F
  - Student s grade in course c: $Gr(s, c)$
    - Discrete range: A, B, C



So, we should expect student s3 to perform better

Example: Predicting Grades

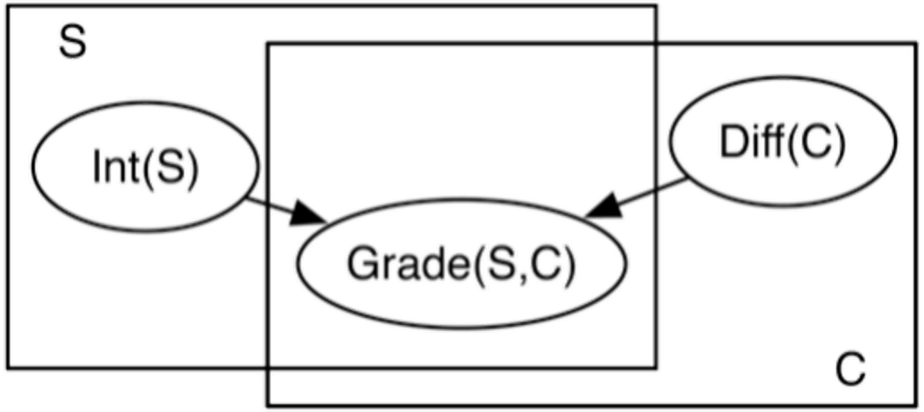
- Relational models: more flexible and compact way
- Program abstraction
  - S, C logical variable representing students, courses
  - Set of individuals of a type is called a domain or population
  - $Int(S), Grade(S, C), Diff(C)$ are parameterized random variables
- Grounding
  - for every student s,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



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

$Int(S)$	$Diff(C)$	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

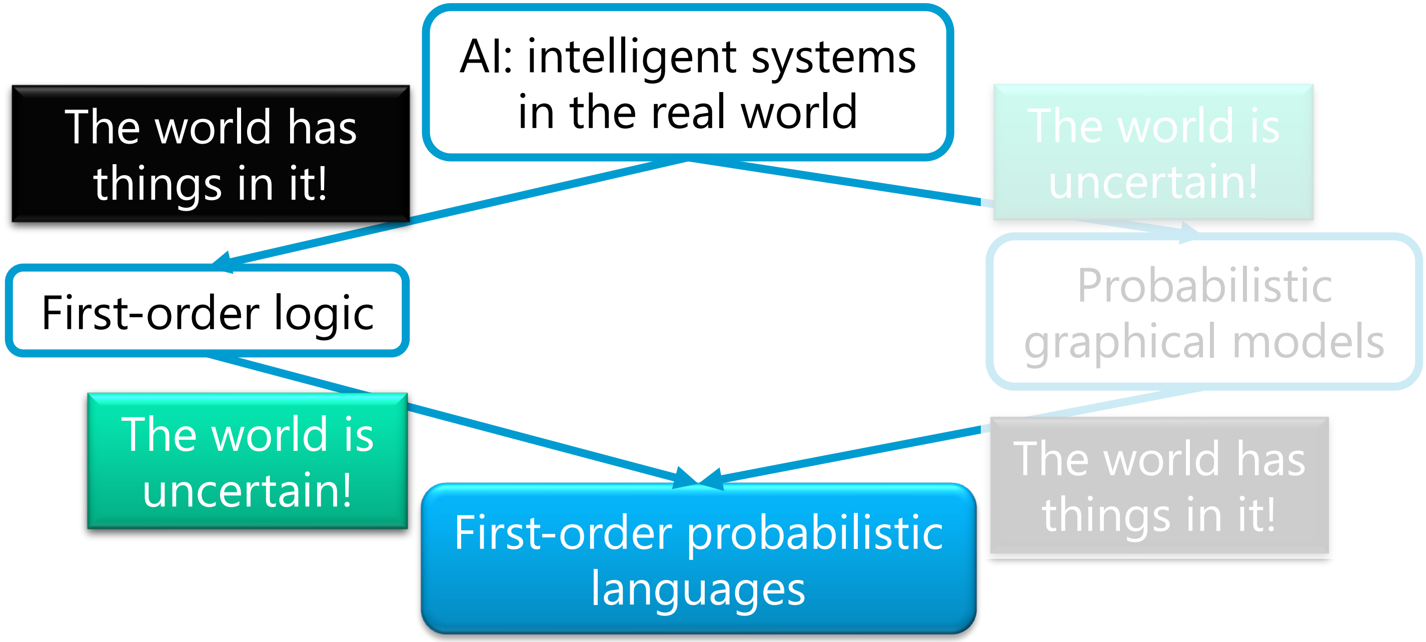


Figure based on Stuart Russell

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 th
  - 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
  - e will see more of them later

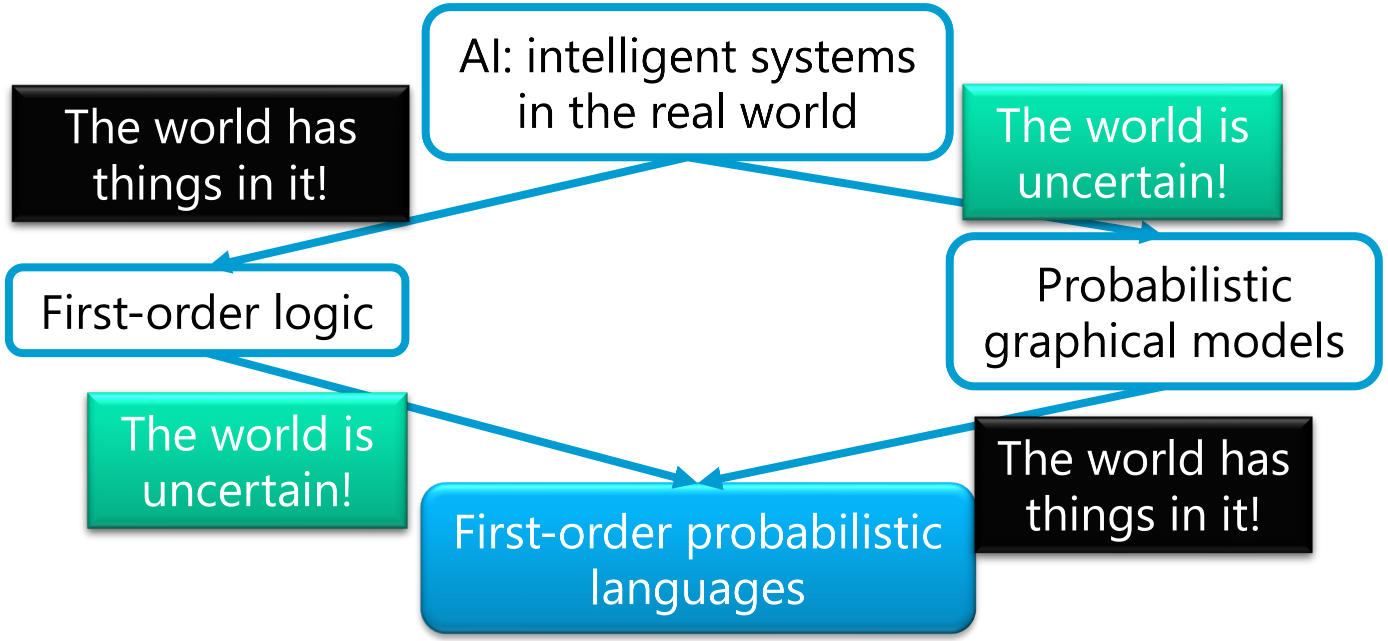
Hard constraint

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

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

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

Connection to AI, Agents, and Environments



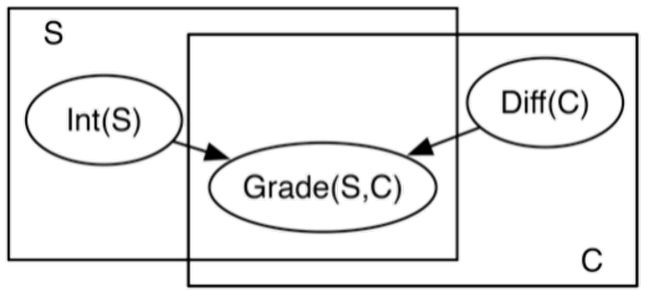
One way of looking at it: Statistical Relational AI

Probabilistic relational models

Figure based on Stuart Russell

Probabilistic Relational Models (PRMs)

- 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.



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

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

- Uncertainty about:
  - Properties of individuals
  - Relationships among individuals
  - Identity (equality) of individuals
  - Existence (and number) of 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

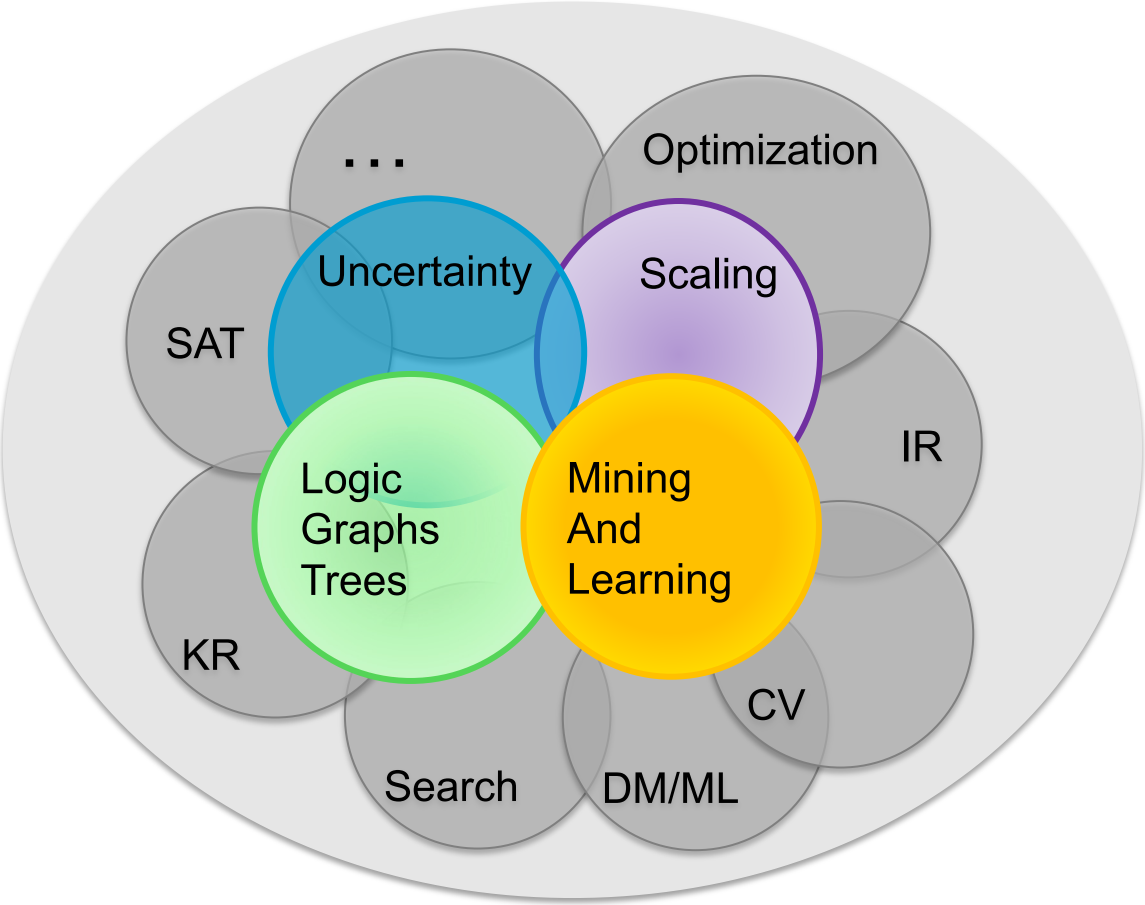
- Approaches exist to deal with such environments to a certain degree

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 & StaRAI

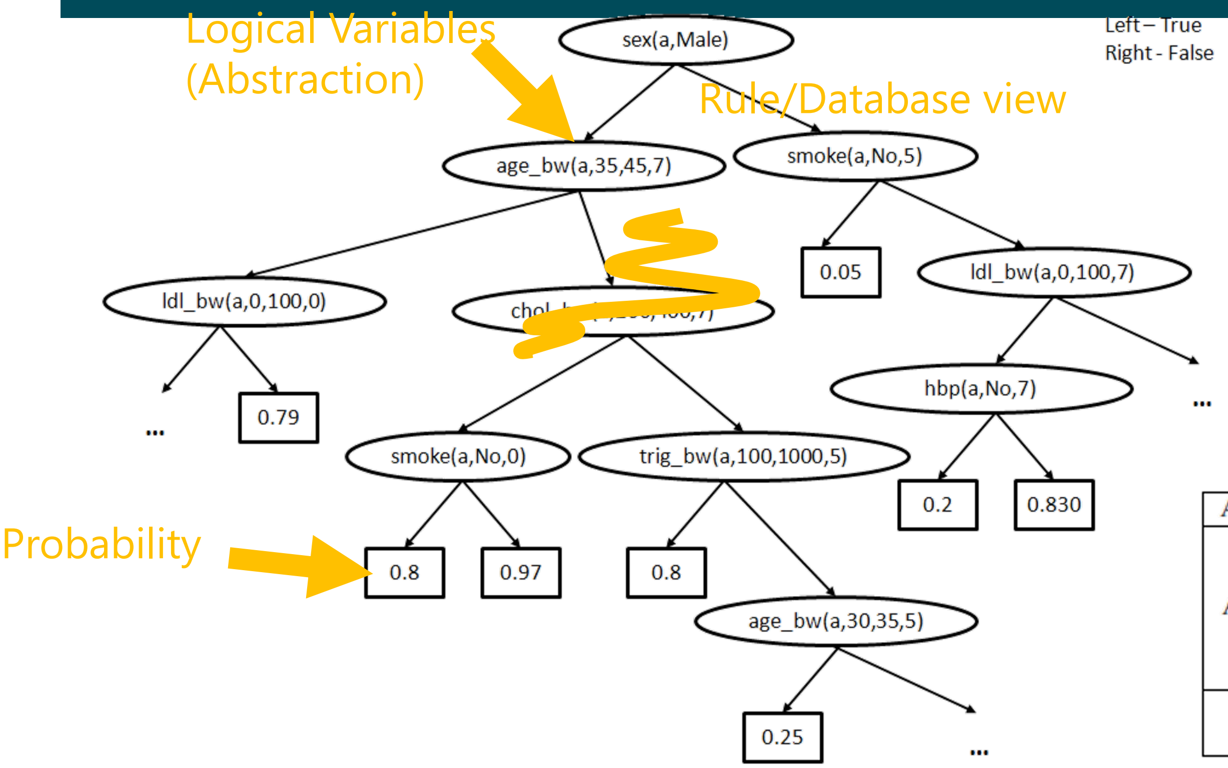
- Study and design
  - intelligent agents
  - that reason about and
  - act in noisy worlds
  - composed of objects and relations among the objects

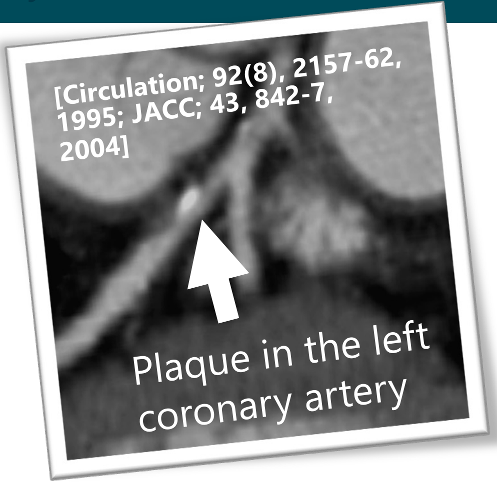


[Getoor, Taskar MIT Press '07; De Raedt, Frasconi, Kersting, Muggleton, LNCS'08; Domingos, Lowd Morgan Claypool '09; Natarajan, Kersting, Khot, Shavlik Springer Brief'15; Russell CACM 58(7): 88-97 '15, Gogate, Domingos CACM 59(7):107-115 ´16]

This “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

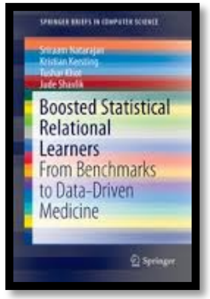
The higher, the better

25%

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

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]

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

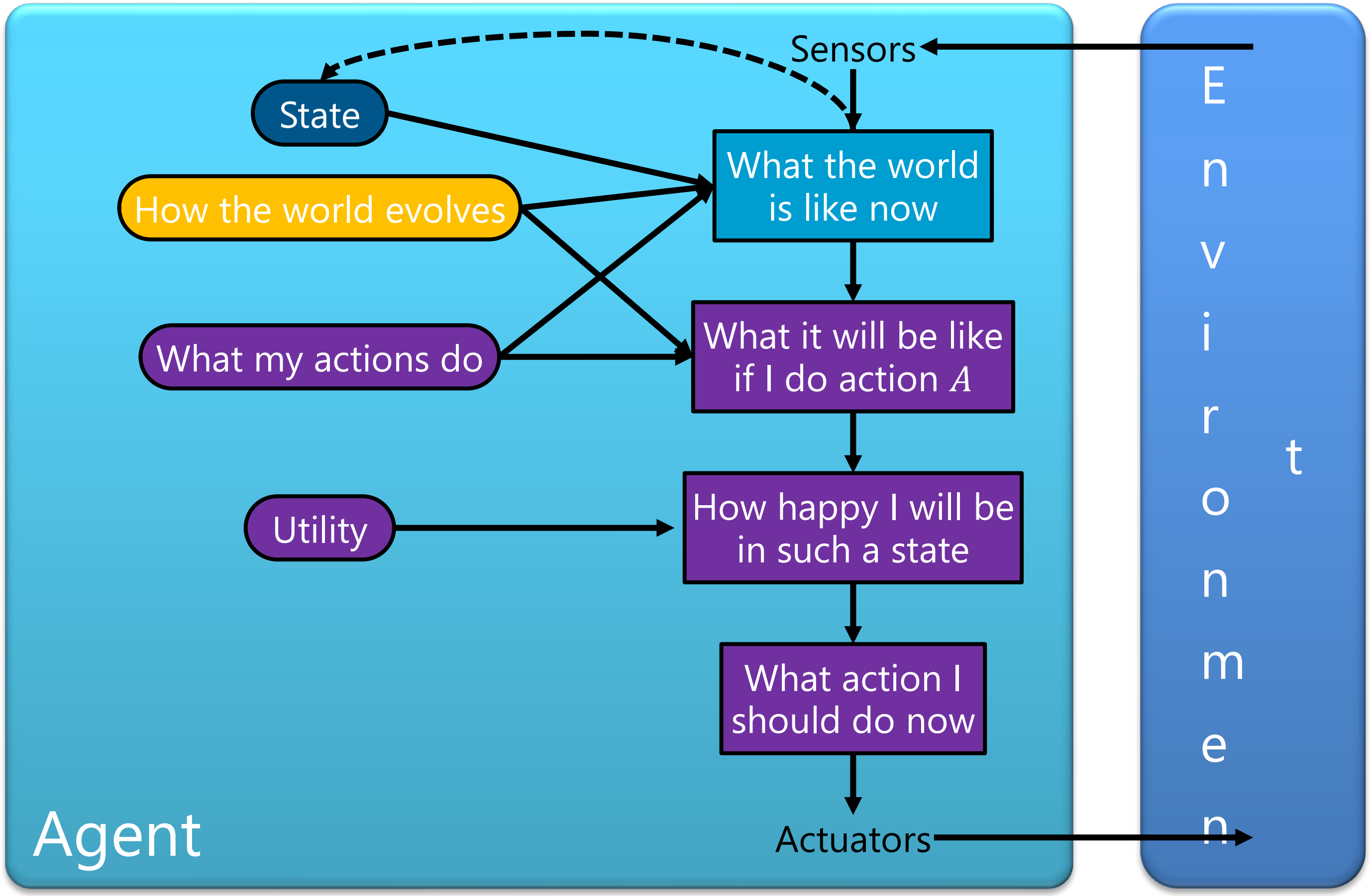
Interim Summary

- 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 sequential PRMs and inference
  6. Lifted decision making
  7. Lifted learning (of episodic PRMs)



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