Inference in StaRAI

Statistical Relational AI

Tutorial at ICCS 2019

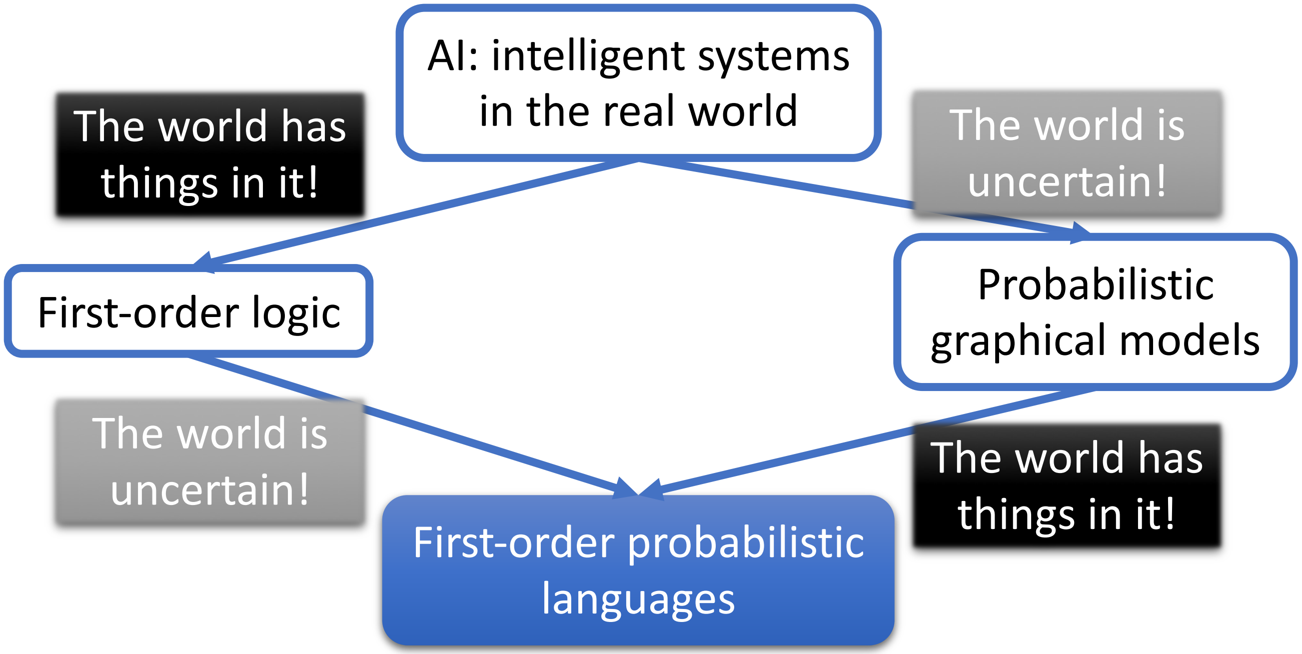
Tanya Braun and Marcel Gehrke, University of Lübeck

UNIVERSITÄT ZU LÜBECK

Thanks to Ralf Möller, Kristian Kersting, and many others for making their slides publicly available

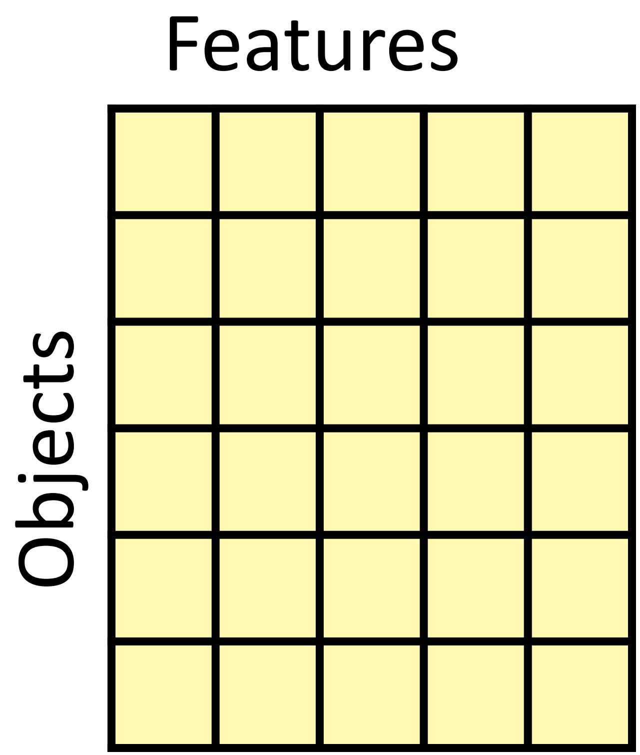
Future of AI?

Stuart Russell

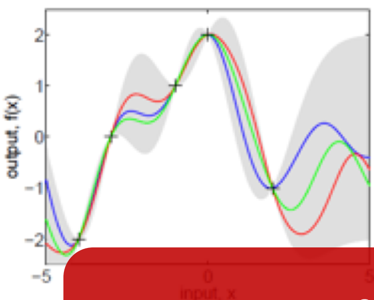


One way of looking at it:
Statistical Relational AI

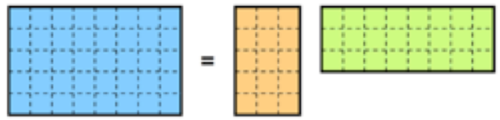
Take your spreadsheet …



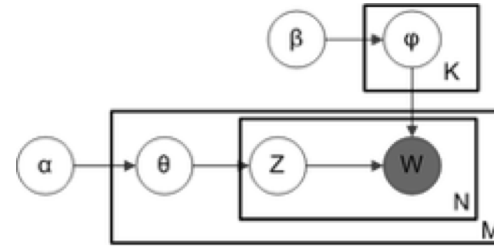
… and apply some AI/ML



Gaussian Processes



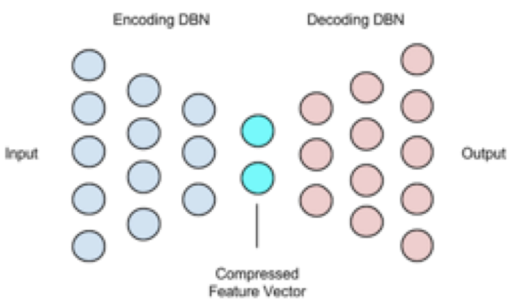
Big Data Matrix Factorization



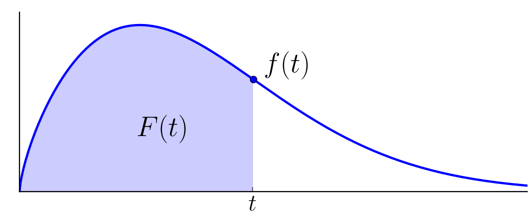
Latent Dirichlet Allocation

Is it really that simple?

Decision Trees/Boosting



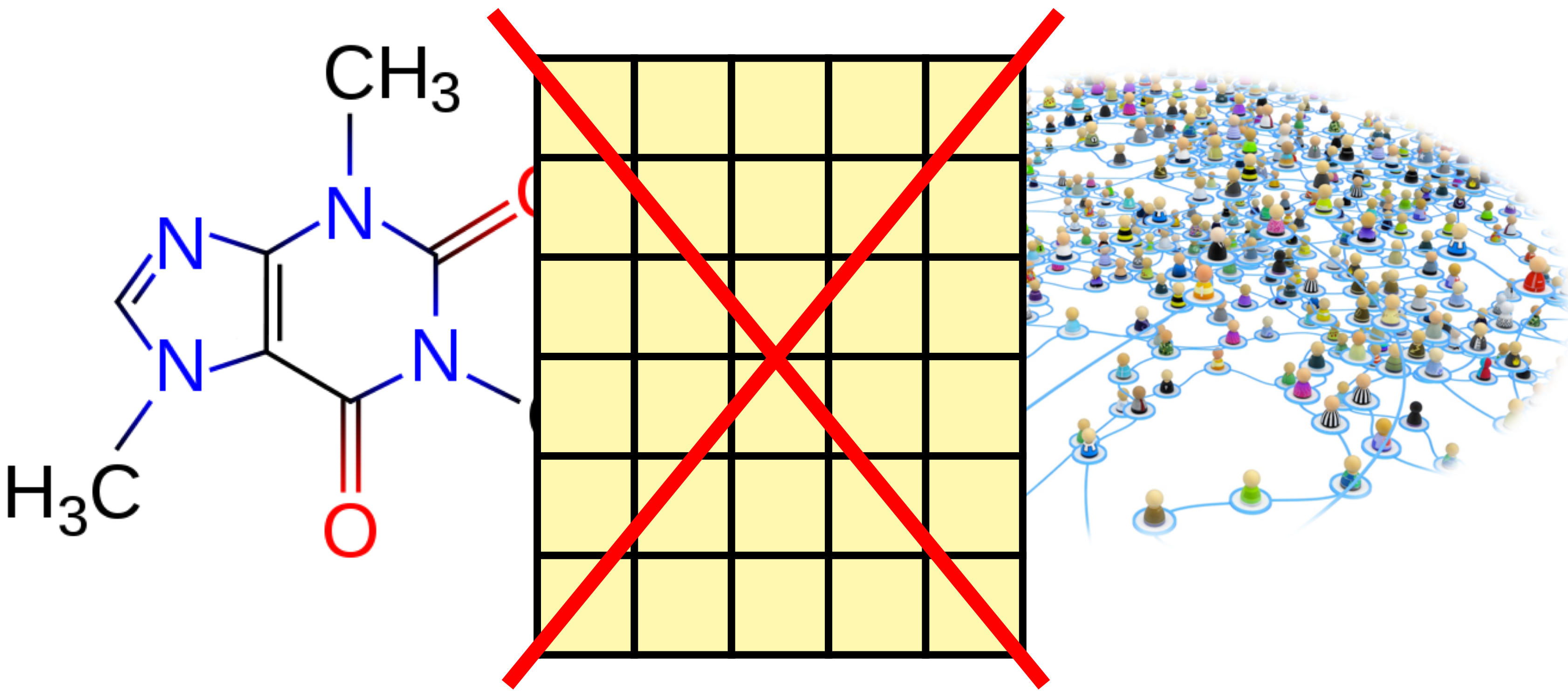
Autoencoder, Deep Learning



Diffusion Models

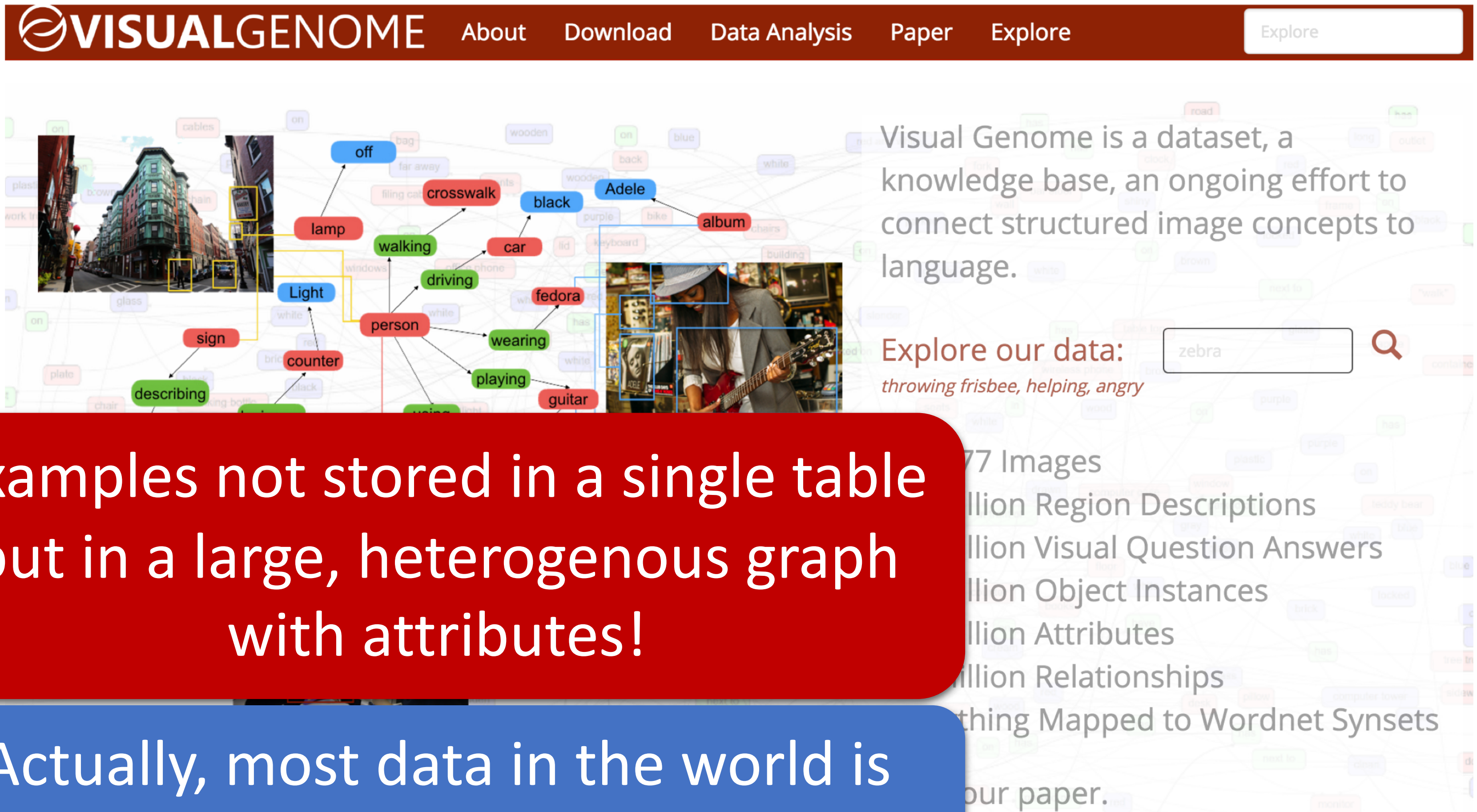
and many more …

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]

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

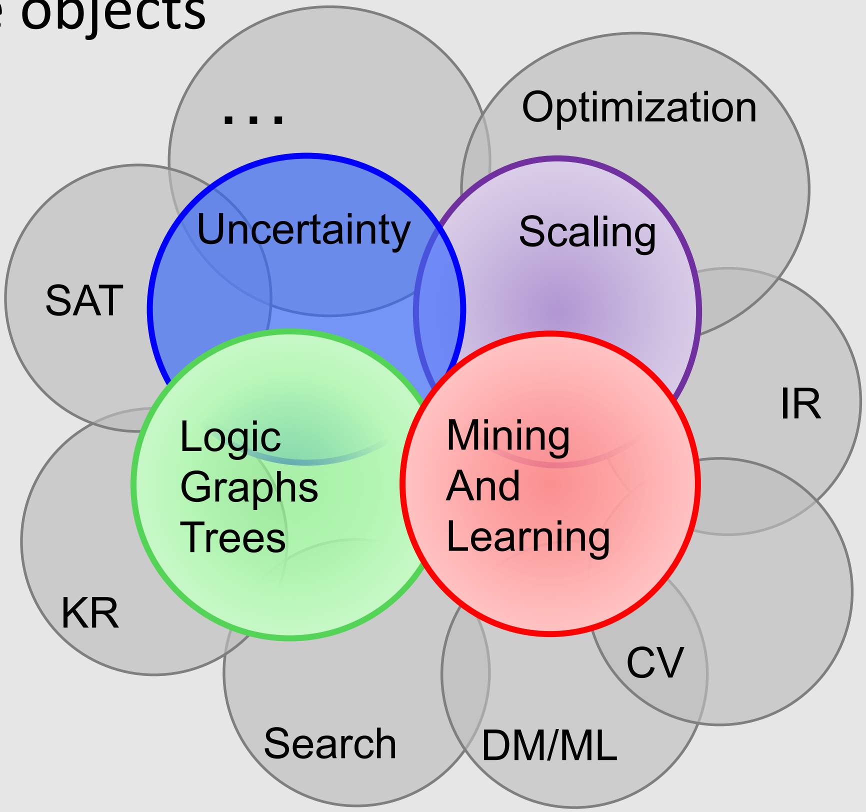
Heart diseases and strokes – cardiovascular disease – are expensive for the world

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

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

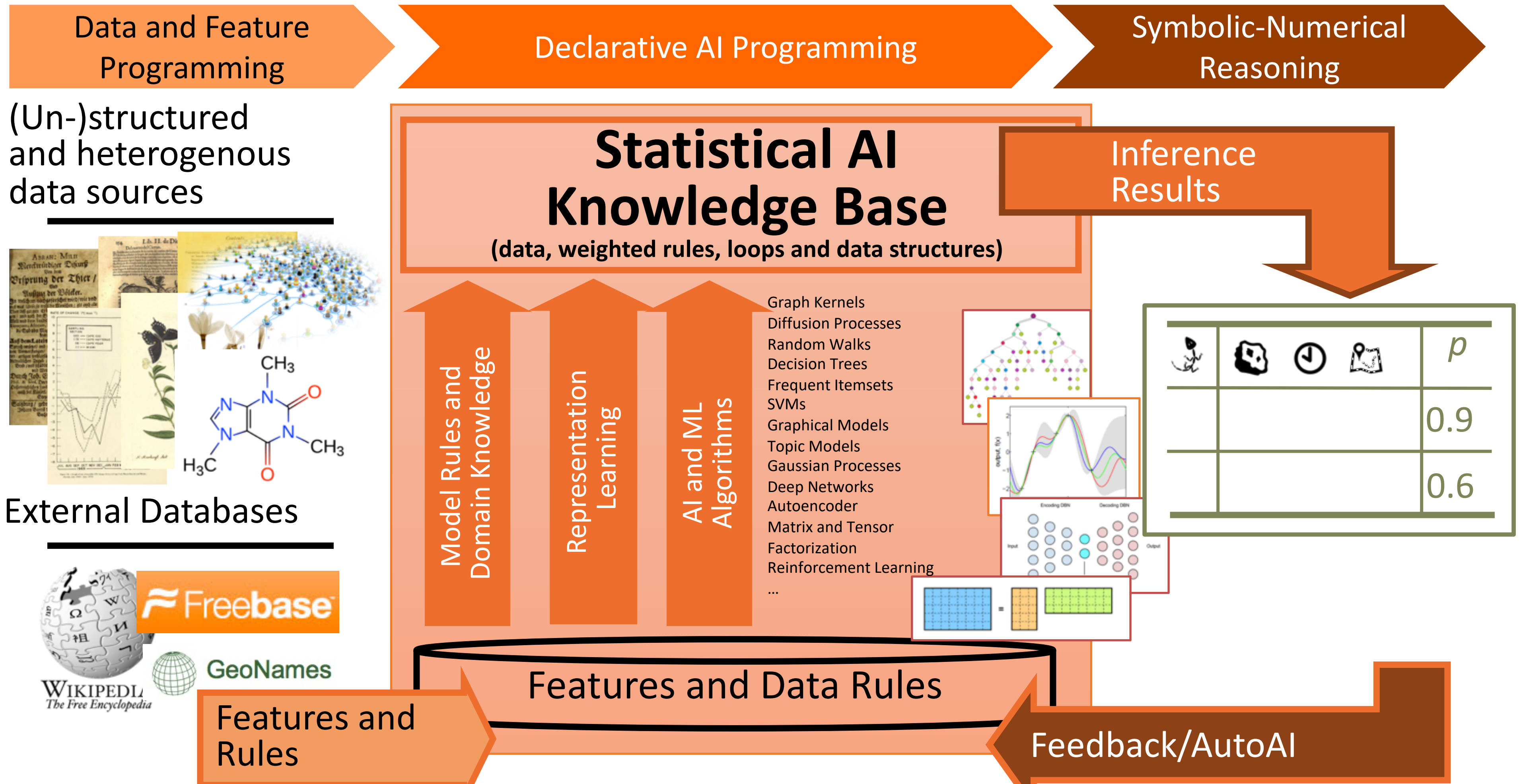
Statistical Relational Learning/AI

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

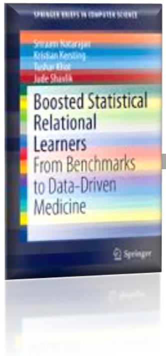


[Getoor, Taskar MIT Press ’07; De Raedt, Frasconi, Kersting, Muggleton, LNCS’08; Domingos, Lowd Morgan Claypool ’09; Natarajan, Kersting, Khot, Shavlik Springer Brief’15; Russell CACM 58(7): 88-97 ’15, Gogate, Domingos CACM 59(7):107-115 ´16]

This establishes a novel “Deep AI”

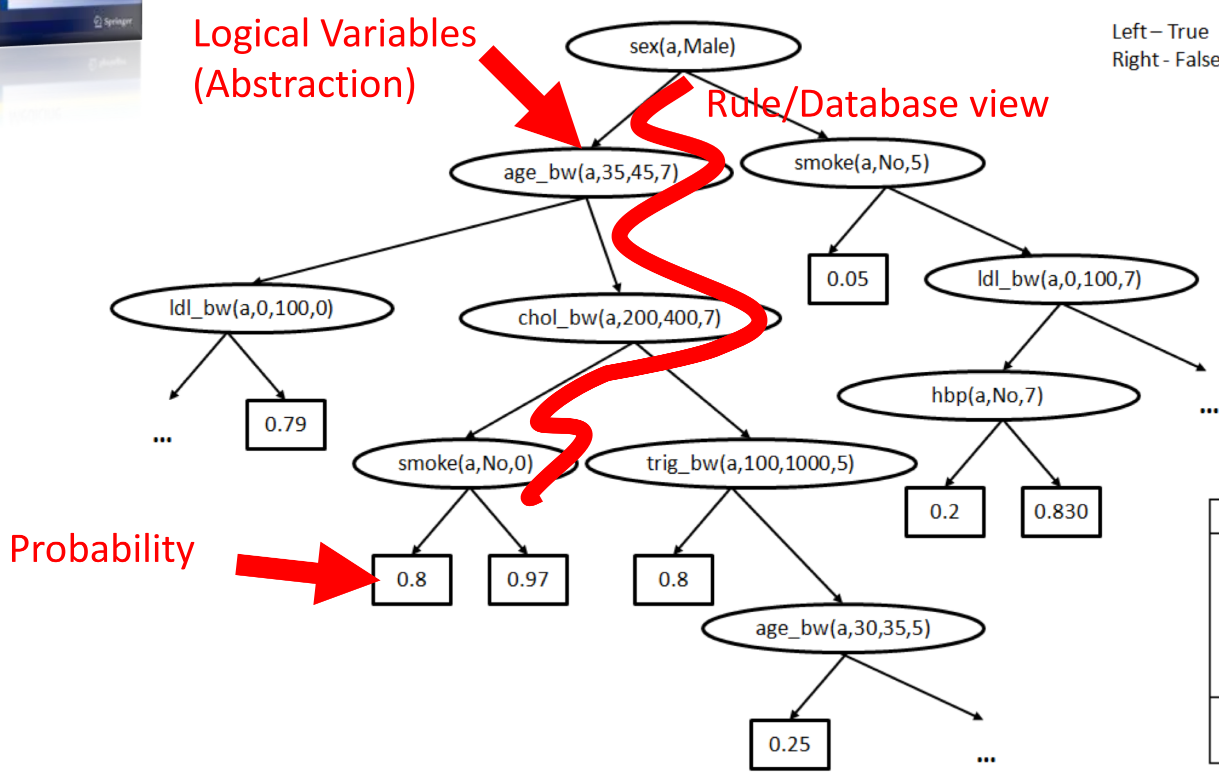


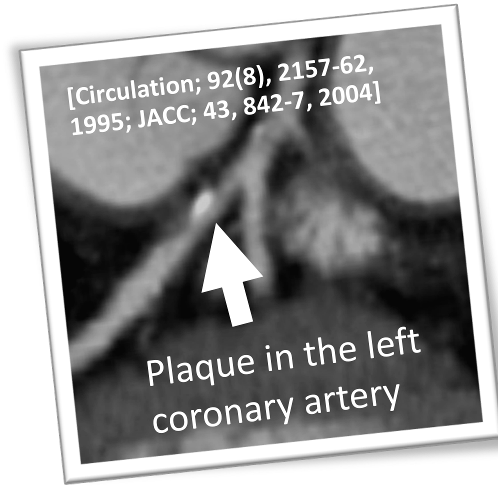
[Ré, Sadeghian, Shan, Shin, Wang, Wu, Zhang IEEE Data Eng. Bull.’14; Natarajan, Picado, Khot, Kersting, Ré, Shavlik ILP’14; Natarajan, Soni, Wazalwar, Viswanathan, Kersting Solving Large Scale Learning Tasks’16, Mladenov, Heinrich, Kleinhans, Gonsior, Kersting DeLBP’16, Kordjamshidi, Roth, Kersting IJCAI-ECAI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



This "Deep AI" can understand EHRs

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





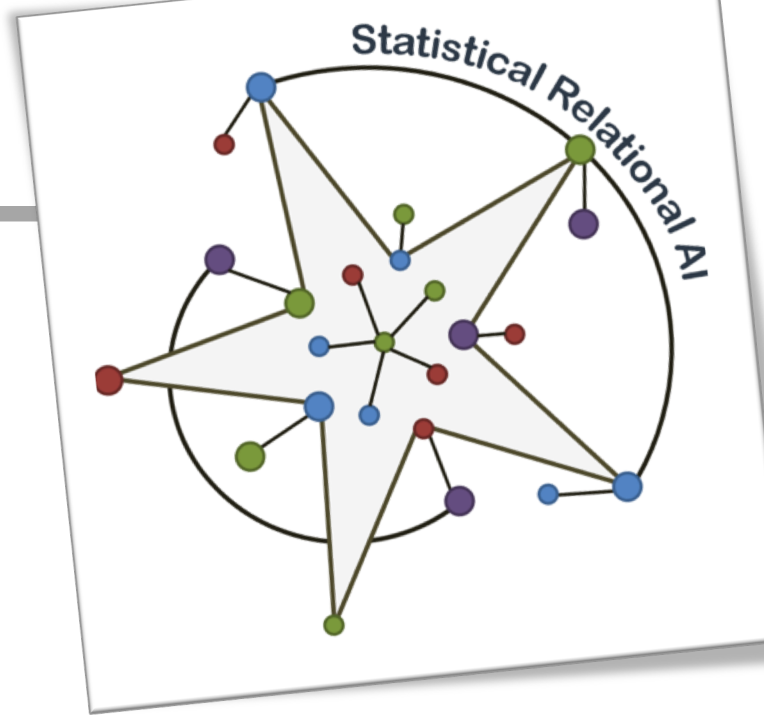
Algorithm	Accuracy	AUC-ROC
J48	0.667	0.607
SVM	0.667	0.5
AdaBoost	0.667	0.608
Bagging	0.677	0.613
NB	0.75	0.653
RPT	0.669*	0.778
RFGB	0.667*	0.819

The higher, the better

25%

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

[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]

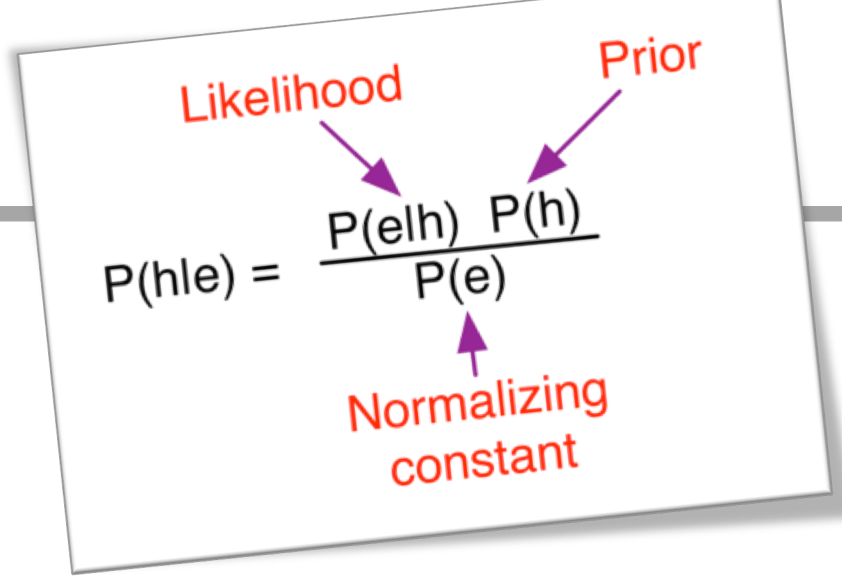


Mission for today

Providing an overview and an introduction into probabilistic inference with a focus on StaRAI

Let’s consider some more gentle examples

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?

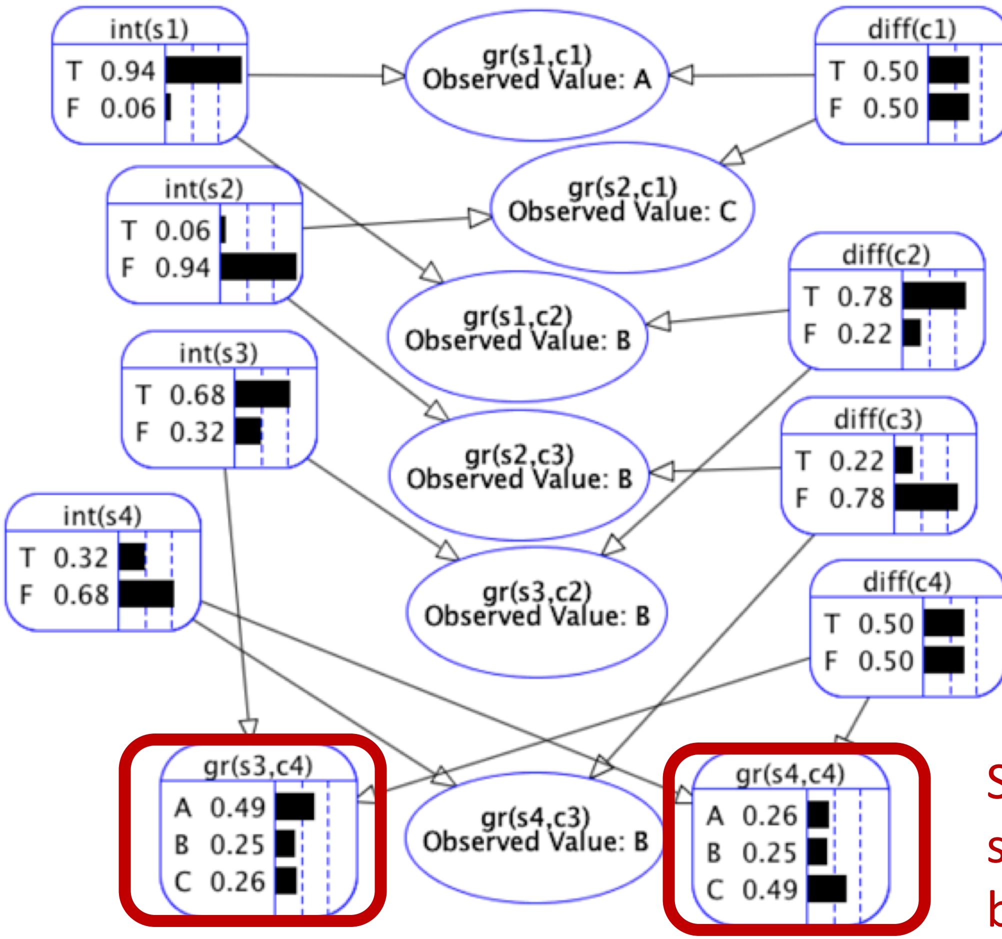
Predicting Grades

- 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	?

Predicting Grades

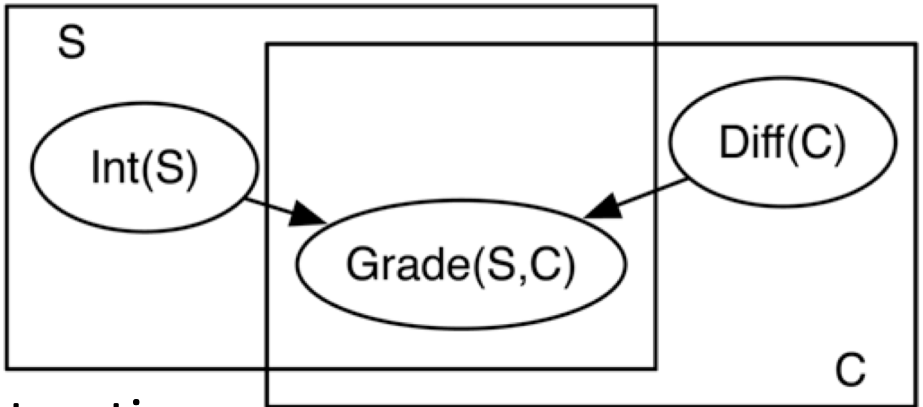
- Rigid and Large graphical model



So, we should expect student s3 to perform better

Predicting Grades

- Relational models: more flexible and compact way

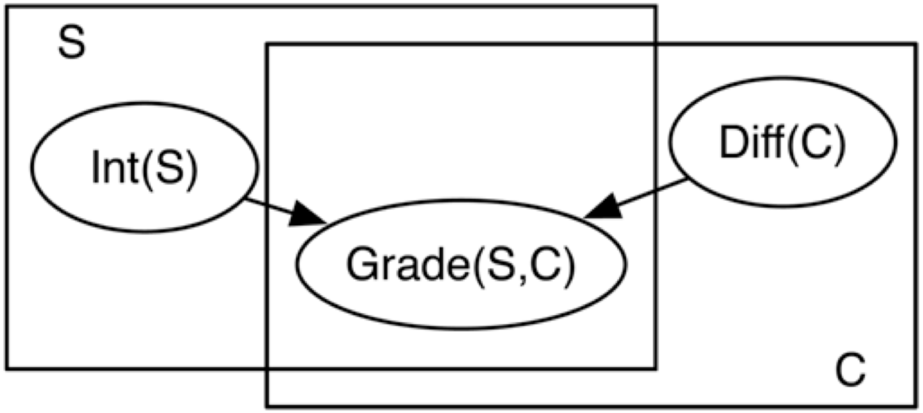


Using plate notation, one can captures the regularities

- Program abstraction
  - S, C logical variable representing students, courses
  - Set of individuals of a type is called a 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

Predicting Grades

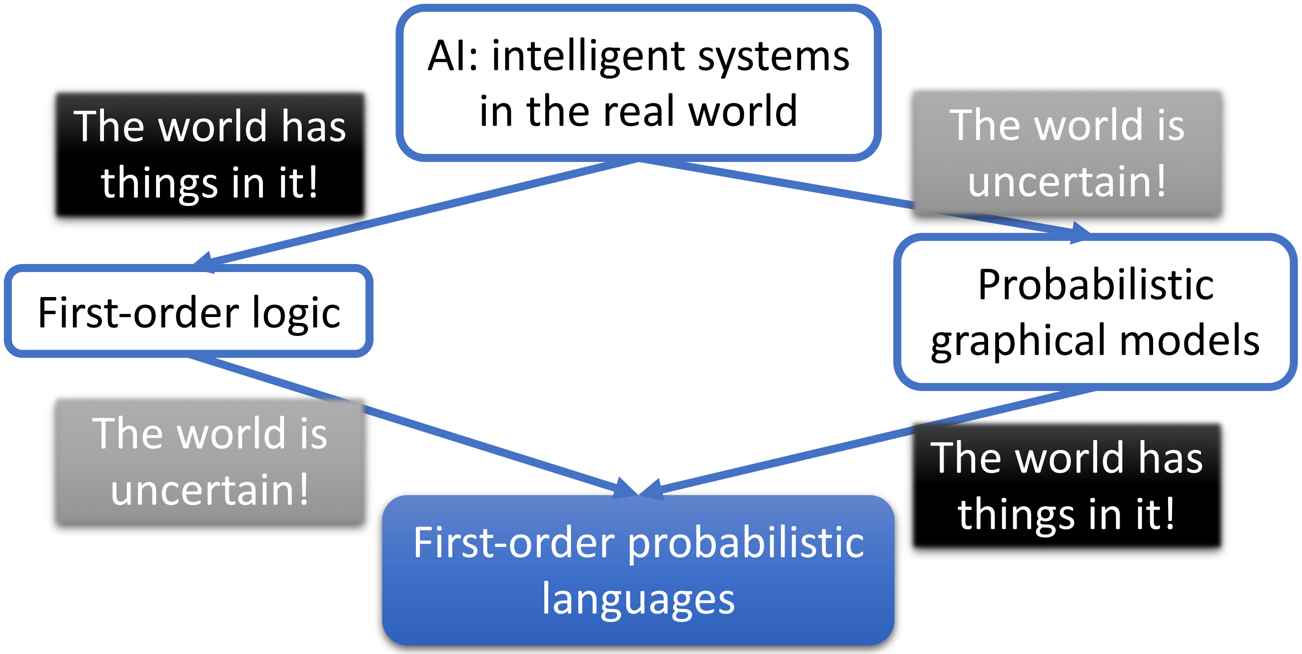
- Relational models: more flexible and compact way



Using plate notation, one can captures the regularities

- If there were 1000 students and 100 courses:
  - Grounding contains
    - 1000 Int(s) variables
    - 100 D(c) variables
    - 100000 Grade(s,c) variablest
  - **total: 101100 variables**
- Numbers to be specified to define the probabilities 1 for I (S), 1 for D(C), 8 for Gr(S,C) = **10 parameters**.

Probabilistic Relational Models



One way of looking at it:
Statistical Relational AI

Probabilistic Relational Models

Probabilistic Relational Models

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

Mission and Schedule of the Tutorial*

Providing an introduction into inference in StaRAI

- Introduction — 20 min ✓
  - StaR AI
- Overview: Probabilistic relational modeling — 30 min
  - Semantics (grounded-distributional, maximum entropy)
  - Inference problems and their applications
  - Algorithms and systems
- Scalable static inference — 40 + 30 min
  - Exact propositional inference
  - Exact lifted inference
- Scalable dynamic inference — 50 min
  - Exact propositional inference
  - Exact lifted inference
- Summary — 10 min

*Thank you to the SRL/StaRAI crowd for all their exciting contributions! The tutorial is necessarily incomplete. Apologies to anyone whose work is not cited