Logic, Probability and Computation: Statistical Relational AI and Beyond

David Poole

Department of Computer Science,
University of British Columbia

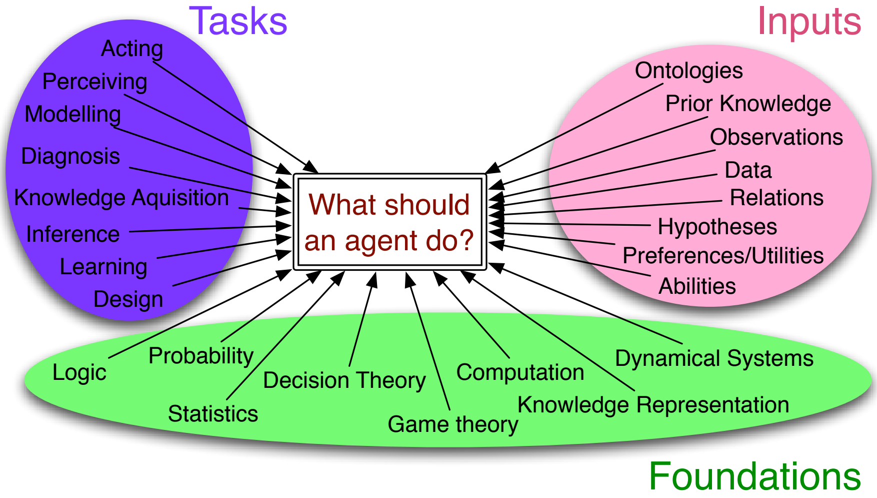
November 2013

For when I am presented with a false theorem, I do not need to examine or even to know the demonstration, since I shall discover its falsity *a posteriori* by means of an easy experiment, that is, by a calculation, costing no more than paper and ink, which will show the error no matter how small it is. . .

And if someone would doubt my results, I should say to him: "Let us calculate, Sir," and thus by taking to pen and ink, we should soon settle the question.

—Gottfried Wilhelm Leibniz [1677]

AI: computational agents that act intelligently



Outline

1. Semantic Science Overview
   - Ontologies
   - Data
   - Hypotheses and Theories
   - Models
2. Making Decisions
3. Relational Probabilistic Models
   - Lifted Inference
4. Existence and Identity Uncertainty

History of AI — a perspective from 2025

- Semantic web has evolved into the world-wide mind (WWM) — a distributed repository of all knowledge, backed up by the best science available.

History of AI — a perspective from 2025

- Semantic web has evolved into the world-wide mind (WWM) — a distributed repository of all knowledge, backed up by the best science available.
- The world-wide mind doesn't just accept new knowledge but critically evaluates it and generates new knowledge.

History of AI — a perspective from 2025

- Semantic web has evolved into the world-wide mind (WWM) — a distributed repository of all knowledge, backed up by the best science available.
- The world-wide mind doesn't just accept new knowledge but critically evaluates it and generates new knowledge.
- Scientists freed from mundane data analysis, develop new hypotheses, interesting questions, and observational data.

History of AI — a perspective from 2025

- Semantic web has evolved into the world-wide mind (WWM) — a distributed repository of all knowledge, backed up by the best science available.
- The world-wide mind doesn't just accept new knowledge but critically evaluates it and generates new knowledge.
- Scientists freed from mundane data analysis, develop new hypotheses, interesting questions, and observational data.
- World-wide mind is the expert on all questions of truth and makes the best predictions. (Using hypotheses provided by a mix of humans and machine learning).

History of AI — a perspective from 2025

- Semantic web has evolved into the world-wide mind (WWM) — a distributed repository of all knowledge, backed up by the best science available.
- The world-wide mind doesn't just accept new knowledge but critically evaluates it and generates new knowledge.
- Scientists freed from mundane data analysis, develop new hypotheses, interesting questions, and observational data.
- World-wide mind is the expert on all questions of truth and makes the best predictions. (Using hypotheses provided by a mix of humans and machine learning).
- Public discourse on values (utilities) to determine the best course of actions for individuals, organizations and society.

Finding information; e.g. diagnosis from symptoms

2013	2025
• need to guess keywords; re-guess until exhaustion	• keywords + context + ontologies → unambiguous query

Finding information; e.g. diagnosis from symptoms

2013	2025
• need to guess keywords; re-guess until exhaustion • what information found is based on popularity and/or appeal to authority	• keywords + context + ontologies → unambiguous query

Finding information; e.g. diagnosis from symptoms

2013	2025
• need to guess keywords; re-guess until exhaustion	• keywords + context + ontologies → unambiguous query
• what information found is based on popularity and/or appeal to authority	• information based on best evidence available in world

Finding information; e.g. diagnosis from symptoms

2013	2025
• need to guess keywords; re-guess until exhaustion	• keywords + context + ontologies → unambiguous query
• what information found is based on popularity and/or appeal to authority	• information based on best evidence available in world
• verify information based on other sites (with different wording)	• information justified by presenting the evidence for and against it

Finding information; e.g. diagnosis from symptoms

2013	2025
• need to guess keywords; re-guess until exhaustion	• keywords + context + ontologies → unambiguous query
• what information found is based on popularity and/or appeal to authority	• information based on best evidence available in world
• verify information based on other sites (with different wording)	• information justified by presenting the evidence for and against it
• extract information from text and graphics to make decisions	• decisions based on evidence and utilities

Believing information

2013	2025
• skeptics throw doubt on science and scientists say "trust us"	• data is available for all to view; all alternative hypotheses can be evaluated

Believing information

2013	2025
• skeptics throw doubt on science and scientists say "trust us"	• data is available for all to view; all alternative hypotheses can be evaluated
• politicians campaign on what is true and what they will do	• politicians campaign on their values

Believing information

2013	2025
• skeptics throw doubt on science and scientists say "trust us"	• data is available for all to view; all alternative hypotheses can be evaluated
• politicians campaign on what is true and what they will do	• politicians campaign on their values
• food shopping is based on price and brands	• food shopping based on optimizing health and well-being (users goals and values, and known risks)

AI Research

2013	2025
• separation of uncertainty and KR issues — ML ignores ontologies — rich representations ignore uncertainty	• uncertainty and ontologies are integral parts of world-wide mind

AI Research

2013	2025
• separation of uncertainty and KR issues — ML ignores ontologies — rich representations ignore uncertainty	• uncertainty and ontologies are integral parts of world-wide mind
• semantic web in its infancy	• world wide mind being used

AI Research

2013	2025
• separation of uncertainty and KR issues — ML ignores ontologies — rich representations ignore uncertainty	• uncertainty and ontologies are integral parts of world-wide mind
• semantic web in its infancy	• world wide mind being used
• relational representations starting to be used in ML	• rich representations with uncertainty ubiquitous

AI Research

2013	2025
• separation of uncertainty and KR issues — ML ignores ontologies — rich representations ignore uncertainty	• uncertainty and ontologies are integral parts of world-wide mind
• semantic web in its infancy	• world wide mind being used
• relational representations starting to be used in ML	• rich representations with uncertainty ubiquitous
• learning based on one or few homogeneous data sets	• learning from all data in world

AI Research

2013	2025
• separation of uncertainty and KR issues — ML ignores ontologies — rich representations ignore uncertainty	• uncertainty and ontologies are integral parts of world-wide mind
• semantic web in its infancy	• world wide mind being used
• relational representations starting to be used in ML	• rich representations with uncertainty ubiquitous
• learning based on one or few homogeneous data sets	• learning from all data in world
• data sets usable only by specialists	• data sets published, available, persistent and interoperable

Science is the foundation of belief

- If system makes a prediction, we should ask: what evidence is there?
- Not all beliefs are equally valid.
- science: We trust scientific conclusions because they are based on evidence.

Science is the foundation of belief

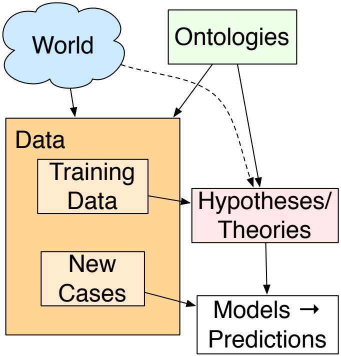
- If system makes a prediction, we should ask: what evidence is there?
- Not all beliefs are equally valid.
- science: We trust scientific conclusions because they are based on evidence.
- semantic web: make all of the world's knowledge accessible to computers.
- semantic science: use scientific method to make informed predictions (conditioning on all information in the world)

Science as the foundation of world-wide mind

Science can be about anything:

- where and when landslides occur
- where to find gold
- what errors students make
- disease symptoms, prognosis and treatment
- what companies will be good to invest in
- what apartment Mary would like
- which celebrities are having affairs

Semantic Science



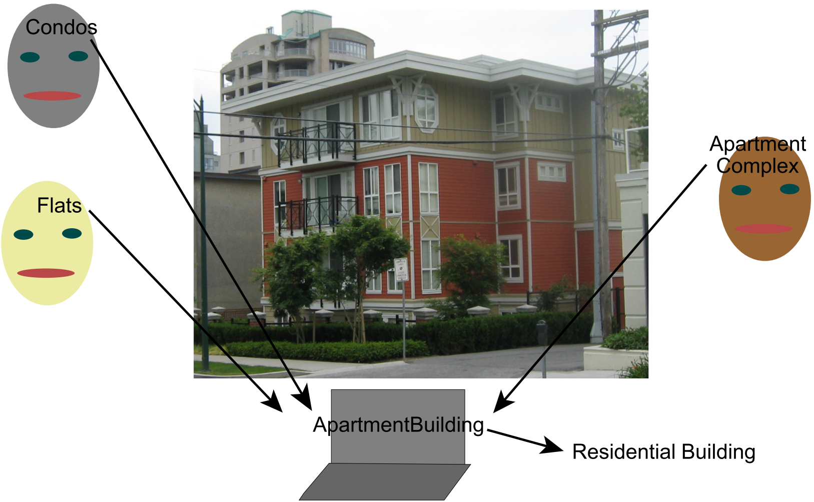
- Ontologies represent the meaning of symbols.
- Data that adheres to ontologies are published.
- Hypotheses that make (probabilistic) predictions on data are published.
- Data used to evaluate hypotheses; the best hypotheses are theories.
- Hypotheses form models for predictions on new cases.
- All evolve in time.

Outline

Ontologies

- In philosophy, ontology the study of existence.
- In CS, an ontology is a (formal) specification of the meaning of the vocabulary used in an information system.
- Ontologies are needed so that information sources can inter-operate at a semantic level.

Ontologies


Condos
Flats
Apartment Complex
ApartmentBuilding
Residential Building

Main Components of an Ontology

- Individuals: the objects in the world (not usually specified as part of the ontology)
- Classes: sets of (potential) individuals
- Properties: between individuals and their values

$\langle Individual, Property, Value \rangle$ triples are universal representations of relations.

Aristotelian definitions

Aristotle [350 B.C.] suggested the definition if a class C in terms of:

- Genus: the super-class
- Differentia: the attributes that make members of the class C different from other members of the super-class

"If genera are different and co-ordinate, their differentiae are themselves different in kind. Take as an instance the genus 'animal' and the genus 'knowledge'. 'With feet', 'two-footed', 'winged', 'aquatic', are differentiae of 'animal'; the species of knowledge are not distinguished by the same differentiae. One species of knowledge does not differ from another in being 'two-footed'."

Aristotle, *Categories*, 350 B.C.

An Aristotelian definition

- An **apartment building** is a **residential building** with **multiple units** and **units are rented**.

$$\begin{aligned} ApartmentBuilding \equiv\ & ResidentialBuilding\& \\ & NumUnits = many\& \\ & Ownership = rental \end{aligned}$$

  $NumUnits : ResidentialBuilding \mapsto \{one, two, many\}$
  $Ownership : Building \mapsto \{owned, rental, coop\}$.

- All classes are defined in terms of properties.

Outline

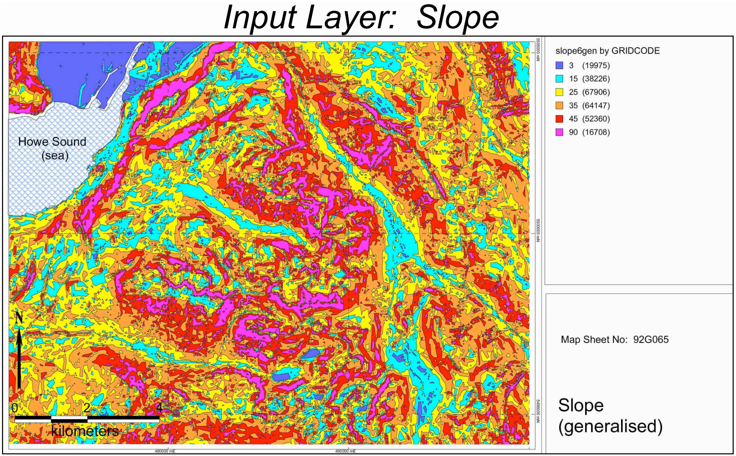
1. Semantic Science Overview
   - Ontologies
   - Data
   - Hypotheses and Theories
   - Models
2. Making Decisions
3. Relational Probabilistic Models
   - Lifted Inference
4. Existence and Identity Uncertainty

Data

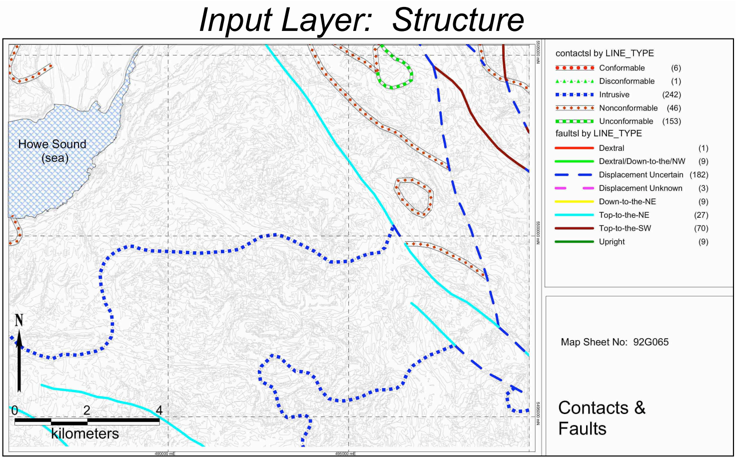
Real data is messy!

- Multiple levels of abstraction
- Multiple levels of detail
- Uses the vocabulary from many ontologies: rocks, minerals, top-level ontology,...
- Rich meta-data:
  - Who collected each datum? (identity and credentials)
  - Who transcribed the information?
  - What was the protocol used to collect the data? (Chosen at random or chosen because interesting?)
  - What were the controls — what was manipulated, when?
  - What sensors were used? What is their reliability and operating range?

Example Data, Geology



Example Data, Geology



http://www.vsto.org/

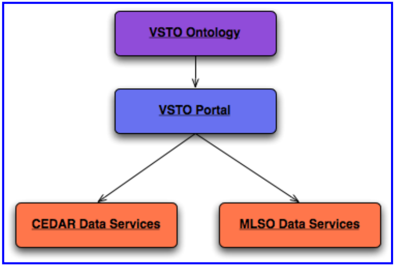
Welcome to the Virtual Solar Terrestrial Observatory

The Virtual Solar Terrestrial Observatory (VSTO) is a unified semantic environment serving data from diverse data archives in the fields of solar, solar-terrestrial, and space physics (SSTSP), currently:

- Upper atmosphere data from the **CEDAR** (Coupling, Energetics and Dynamics of Atmospheric Regions) archive
- Solar corona data from the **MLSO** (Mauna Loa Solar Observatory) archive

The VSTO portal uses an underlying ontology (i.e. an organized knowledge base of the SSTSP domain) to present a general interface that allows selection and retrieval of products (ascii and binary data files, images, plots) from heterogenous external data services.

VSTO Data Access



Data is theory-laden

- Sapir-Whorf Hypothesis [Sapir 1929, Whorf 1940]: people's perception and thought are determined by what can be described in their language. (Controversial in linguistics!)

Data is theory-laden

- Sapir-Whorf Hypothesis [Sapir 1929, Whorf 1940]: people's perception and thought are determined by what can be described in their language. (Controversial in linguistics!)
- A stronger version for information systems:

  What is stored and communicated by an information system is constrained by the representation and the ontology used by the information system.

- Ontologies come logically prior to the data.
- Data can't make distinctions that can't be expressed in the ontology.
- Different ontologies result in different data.

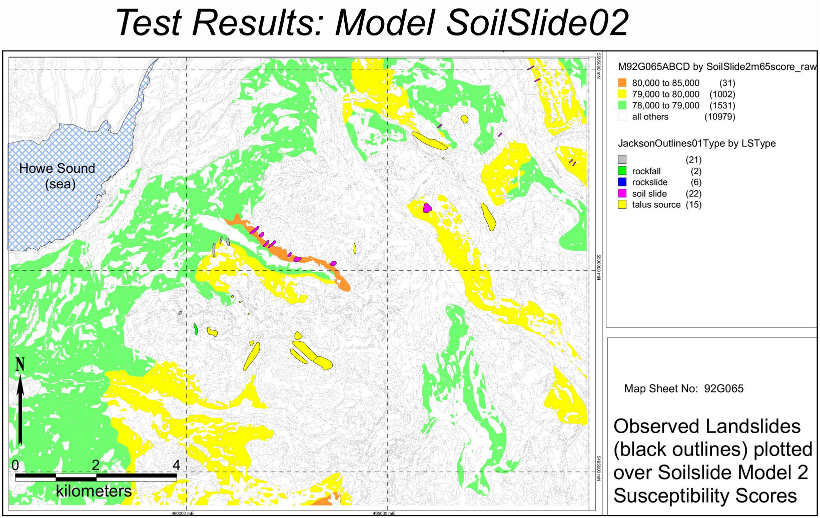
Outline

Hypotheses make predictions on data

Hypotheses are procedures that make prediction on data.
Theories are hypotheses that best fit the observational data.

- Hypotheses can make various predictions about data:
  - point probabilities: probability you will be run over tomorrow is 0.002
  - ...
- Probabilistic predictions are what is needed for decision making and can be learned from data.

Example Prediction from a Hypothesis



Applying hypotheses to new cases

- Hypotheses are often narrow, e.g., prognosis of people with a lung cancer.
- Hypotheses are general in the sense that they can be adapted to different cases.

Applying hypotheses to new cases

- Hypotheses are often narrow, e.g., prognosis of people with a lung cancer.
- Hypotheses are general in the sense that they can be adapted to different cases.
- How can we compare hypotheses that differ in their generality?
- Hypothesis A makes predictions about all cancers. Hypothesis B makes predictions about lung cancers. Should the comparison between A and B take into account A's predictions on non-lung cancer?

Applying hypotheses to new cases

- Hypotheses are often narrow, e.g., prognosis of people with a lung cancer.
- Hypotheses are general in the sense that they can be adapted to different cases.
- How can we compare hypotheses that differ in their generality?
- Hypothesis A makes predictions about all cancers. Hypothesis B makes predictions about lung cancers. Should the comparison between A and B take into account A's predictions on non-lung cancer?
- What about C: *if lung cancer, use B's prediction, else use A's prediction*?

Models

- A model is an ensemble of hypotheses applied to a particular case.
  - E.g., *if lung cancer, use B's prediction, else use A's prediction*
  - Can use sophisticated methods to determine which hypothesis to use.
  - Judge hypotheses by how well they fit into models.
  - Models can be judged by simplicity.
  - Hypothesis designers don't need to game the system by manipulating the generality of hypotheses

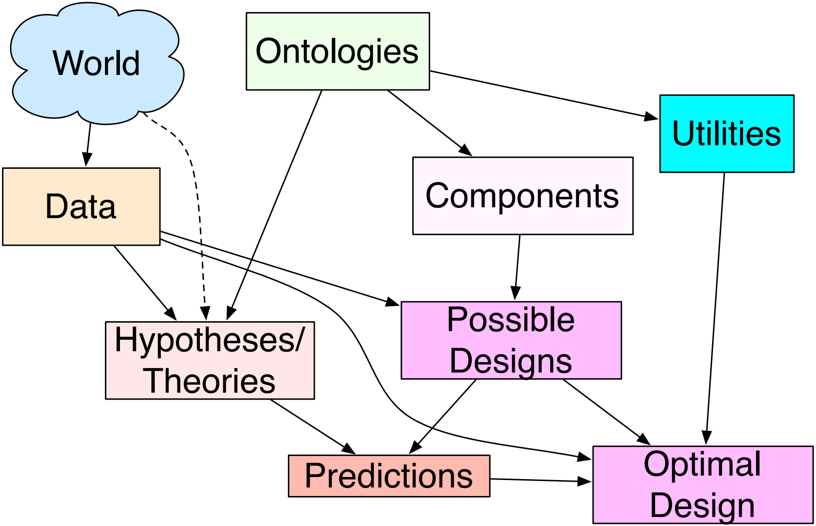
Dynamics of Semantic Science

- New data and hypotheses are continually added.
- Anyone can design their own ontologies.
  — People vote with their feet what ontology they use.
  — Need for semantic interoperability leads to ontologies with mappings between them.
- Hypotheses engineered + learned
- Ontologies evolve with hypotheses:
  A hypothesis learns useful unobserved features
  $\longrightarrow$ add these to an ontology
  $\longrightarrow$ other researchers can refer to them
  $\longrightarrow$ reinterpretation of data
- Ontologies can be judged by the predictions of the hypotheses that use them
  — role of a vocabulary is to describe useful distinctions.

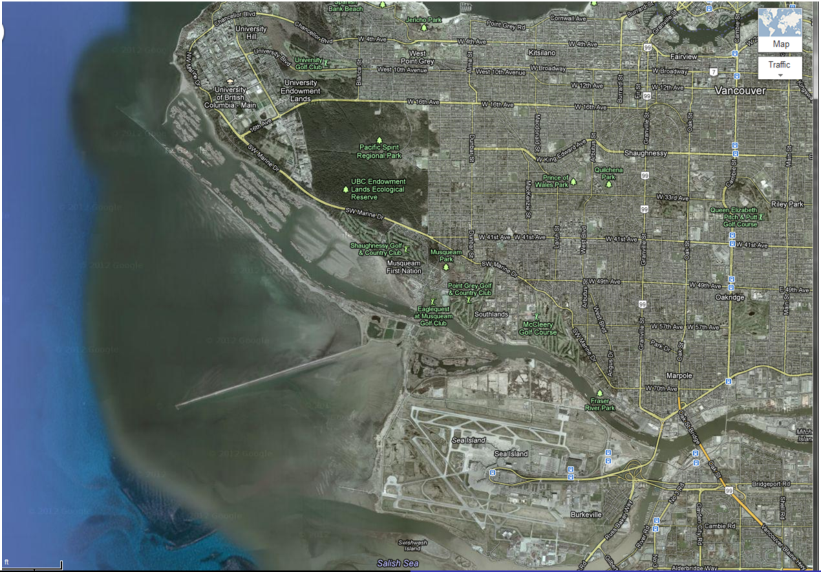
Outline

1. Semantic Science Overview
   - Ontologies
   - Data
   - Hypotheses and Theories
   - Models
2. Making Decisions
3. Relational Probabilistic Models
   - Lifted Inference
4. Existence and Identity Uncertainty

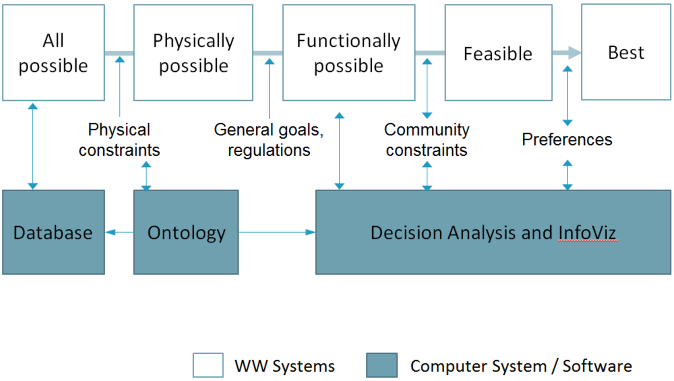
Decision Making



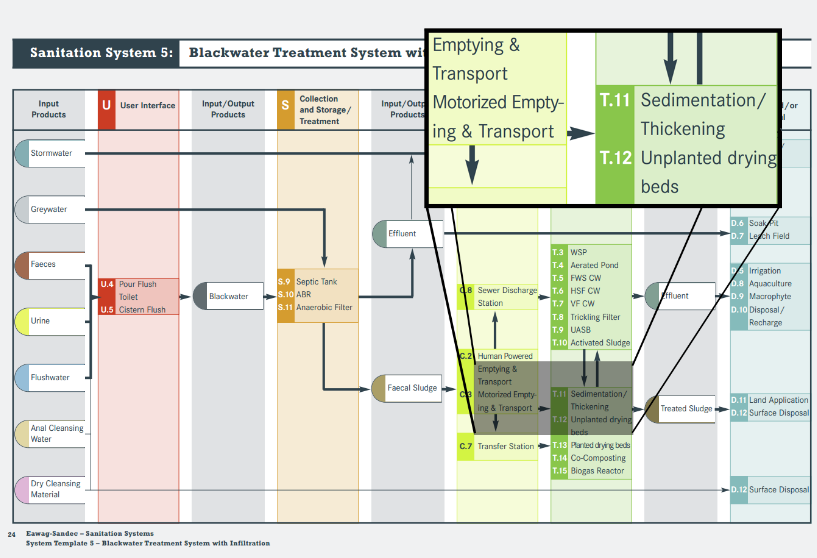
Wastewater Management



Decision Making applied to Wastewater Management



Traditional Design Space



Visualizing Design Space

Desiderata:

- hierarchical — can drill down into details but are not overwhelmed by details
- show the diversity of possible solutions.
- as simple as possible but no simpler
- explore feasible solutions and infeasible solutions

Traditional Utility Tradeoffs

Fidelis Resource Group, Integrated Resource Recovery Study, for Metro Vancouver, 2011

Table 17: Triple Bottom Line Evaluation

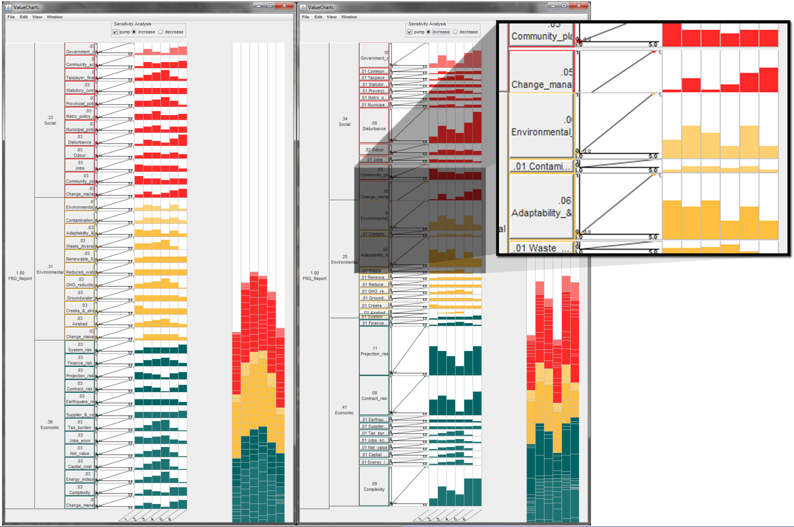
Net totals:	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Economic	17▼	—	5▲	7▲	5▼	13▼
Rank	#6	#3	#2	#1	#4	#5
Environmental	7▼	—	3▲	2▲	3▼	21▼
Rank	#5	#3	#1	#2	#4	#6
Social	13▼	—	—	4▼	1▲	—
Rank	#6	#2	#2	#5	#1	#2
▼	45▼	—	11▼	29▼	13▼	52▼
▲	8▲	—	19▲	34▲	6▲	18▲
Net total	37▼	—	8▲	5▲	7▼	34▼
Rank	**#6**	**#3**	**#1**	**#2**	**#4**	**#5**

As at: 1 Mar 2001

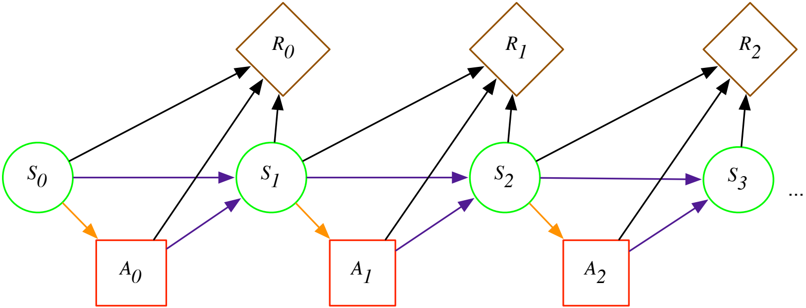
	Evaluation item	*Primary account*	1	2	3	4	5	6
1	Change management	Economic	▼▼	—	▼	▼▼▼	▲	▲▲
2	Complexity	Economic	▼▼▼	—	▼	▼▼	▲▲	▲▲
3	Energy independence	Economic	▼▼	—	▲	▲▲▲▲	▼	▼▼▼▼
4	Capital cost	Economic	▼▼▼	—	▲▲	▲▲▲▲	▼▼	▼▼▼▼
5	Net value	Economic	▼▼	—	▲	▲▲▲	▼	▼▼▼▼
6	Jobs	Economic	▲	—	▲	▲▲	▼	▼▼
7	Tax burden	Economic	▼▼	—	▲▲▲	▲▲▲▲	▼▼	▼▼▼▼
8	Supplier & competitive readiness	Economic	—	—	—	—	—	▲
9	Earthquake risk	Economic	—	—	—	—	—	▼▼
10	Contract risk	Economic	▼	—	▼	▼▼▼▼	▼	▲
11	Projection risk	Economic	▲	—	▼	▼▼▼	—	▲
12	Finance risk	Economic	▼▼	—	▲	▲▲	—	▼▼
13	System risk	Economic	▼▼	—	—	—	—	▲▲
14	Change management	Environmental	▼▼▼	—	—	—	—	▼▼
15	Airshed	Environmental	▼	—	▲	▲▲	▼	▼▼
16	Creeks & streams	Environmental	▲▲	—	—	—	—	—
17	Groundwater	Environmental	—	—	—	—	—	▼
18	GHG reduction	Environmental	▼	—	▲	▲▲▲	—	▼▼▼▼
19	Reduced water consumption	Environmental	—	—	—	—	—	—
20	Renewable fuel use	Environmental	▼	—	—	—	—	—

Finding Optimal Designs

Value Charts: [Jeanette Bautista and Giuseppe Carenini, 2008]

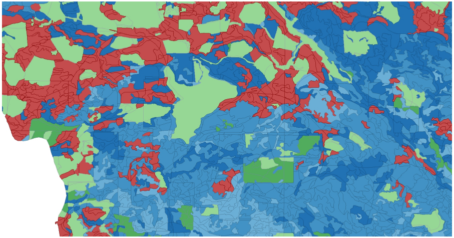


Markov Decision Processes



Planning in Forestry

Example Map of Age Feature



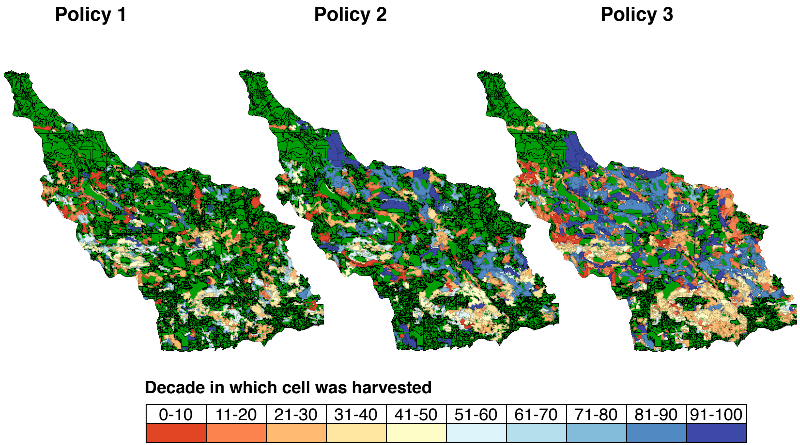
Age of trees in cell.

0-25	26-50	51-75	76-100	101-150	150-

Scale for 10 Binary Features and Binary Actions

Number of ...	at each cell	entire landscape
actions	2	$2^{1000} \approx 10^{300}$
states	2^{10}	$(2^{10})^{1000} \approx 10^{3000}$

Resulting Plans



Outline

1. Semantic Science Overview
   - Ontologies
   - Data
   - Hypotheses and Theories
   - Models
2. Making Decisions
3. Relational Probabilistic Models
   - Lifted Inference
4. Existence and Identity Uncertainty

Relational Learning

- Often the values of properties are not meaningful values but names of individuals.
- It is the properties of these individuals and their relationship to other individuals that needs to be learned.
- Relational learning has been studied under the umbrella of "Inductive Logic Programming" as the representations are often logic programs.

Example: trading agent

What does Joe like?

Individual	Property	Value
joe	*likes*	*resort_14*
joe	*dislikes*	*resort_35*
...	...	...
resort_14	*type*	*resort*
resort_14	*near*	*beach_18*
beach_18	*type*	*beach*
beach_18	*covered_in*	*ws*
ws	*type*	*sand*
ws	*color*	*white*
...	...	...

Values of properties may be meaningless names.

Example: trading agent

Possible theory that could be learned:

$$
\begin{aligned}
& prop(joe, likes, R) \leftarrow \\
& \qquad prop(R, type, resort) \land \\
& \qquad prop(R, near, B) \land \\
& \qquad prop(B, type, beach) \land \\
& \qquad prop(B, covered\_in, S) \land \\
& \qquad prop(S, type, sand).
\end{aligned}
$$

Joe likes resorts that are near sandy beaches.

- But we want probabilistic predictions.

Independent Choice Logic

- A language for first-order probabilistic models.
- Idea: combine logic and probability, where all uncertainty in handled in terms of Bayesian decision theory, and a logic program specifies consequences of choices.
- Parametrized random variables are represented as logical atoms, and plates correspond to logical variables.

Independent Choice Logic

- An alternative is a set of ground atomic formulas. $\mathcal{C}$, the choice space is a set of disjoint alternatives.
- $\mathcal{F}$, the facts is a logic program that gives consequences of choices.
- P_0 a probability distribution over alternatives:

$$\forall A \in \mathcal{C} \ \sum_{a \in A} P_0(a) = 1.$$

Meaningless Example

$$\mathcal{C} = \{\{c_1, c_2, c_3\}, \{b_1, b_2\}\}$$

$$\begin{aligned}\mathcal{F} = \{\ & f \leftarrow c_1 \wedge b_1, && f \leftarrow c_3 \wedge b_2, \\ & d \leftarrow c_1, && d \leftarrow {\sim} c_2 \wedge b_1, \\ & e \leftarrow f, && e \leftarrow {\sim} d\}\end{aligned}$$

$$P_0(c_1) = 0.5 \quad P_0(c_2) = 0.3 \quad P_0(c_3) = 0.2$$
$$P_0(b_1) = 0.9 \quad P_0(b_2) = 0.1$$

Semantics of ICL

- There is a possible world for each selection of one element from each alternative.
- The logic program together with the selected atoms specifies what is true in each possible world.
- The elements of different alternatives are independent.

Meaningless Example: Semantics

$$\mathcal{F} = \{\; f \leftarrow c_1 \wedge b_1, \quad f \leftarrow c_3 \wedge b_2,$$
$$d \leftarrow c_1, \quad d \leftarrow \sim c_2 \wedge b_1,$$
$$e \leftarrow f, \quad e \leftarrow \sim d\}$$

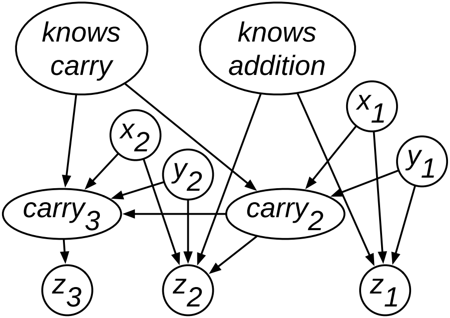
$$P_0(c_1) = 0.5 \quad P_0(c_2) = 0.3 \quad P_0(c_3) = 0.2$$
$$P_0(b_1) = 0.9 \quad P_0(b_2) = 0.1$$

$$\begin{array}{cccccccc} & & \overbrace{\quad\quad\quad\quad}^{\text{selection}} & & \overbrace{\quad\quad\quad\quad\quad\quad}^{\text{logic program}} & & & \\ w_1 & \models & c_1 & b_1 & f & d & e & P(w_1) = 0.45 \\ w_2 & \models & c_2 & b_1 & \sim f & \sim d & e & P(w_2) = 0.27 \\ w_3 & \models & c_3 & b_1 & \sim f & d & \sim e & P(w_3) = 0.18 \\ w_4 & \models & c_1 & b_2 & \sim f & d & \sim e & P(w_4) = 0.05 \\ w_5 & \models & c_2 & b_2 & \sim f & \sim d & e & P(w_5) = 0.03 \\ w_6 & \models & c_3 & b_2 & f & \sim d & e & P(w_6) = 0.02 \end{array}$$

$$P(e) = 0.45 + 0.27 + 0.03 + 0.02 = 0.77$$

Bayesian Networks

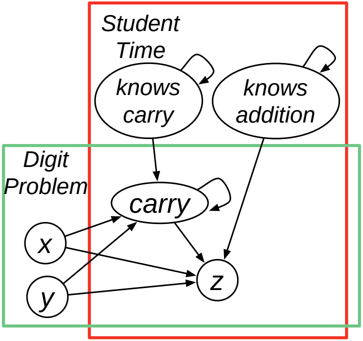
$$
\begin{array}{rrrr}
 & & x_2 & x_1 \\
+ & & y_2 & y_1 \\
\hline
 & z_3 & z_2 & z_1
\end{array}
$$



What if there were multiple digits, problems, students, times?
How can we build a model before we know the individuals?

Multi-digit addition with parametrized BNs / plates

$$
\begin{array}{ccccc}
 & x_{j_x} & \cdots & x_2 & x_1 \\
+ & y_{j_z} & \cdots & y_2 & y_1 \\
\hline
 & z_{j_z} & \cdots & z_2 & z_1
\end{array}
$$



Random Variables: $x(D,P)$, $y(D,P)$, $knowsCarry(S,T)$, $knowsAddition(S,T)$, $carry(D,P,S,T)$, $z(D,P,S,T)$
for each: digit D, problem P, student S, time T

☛ parametrized random variables

ICL rules for multi-digit addition

$$
\begin{aligned}
z(D,P,S,T) = V \leftarrow \\
&x(D,P) = Vx \land \\
&y(D,P) = Vy \land \\
&carry(D,P,S,T) = Vc \land \\
&knowsAddition(S,T) \land \\
&\neg mistake(D,P,S,T) \land \\
&V \text{ is } (Vx + Vy + Vc) \text{ div } 10.
\end{aligned}
$$

$$
\begin{aligned}
z(D,P,S,T) = V \leftarrow \\
&knowsAddition(S,T) \land \\
&mistake(D,P,S,T) \land \\
&selectDig(D,P,S,T) = V.
\end{aligned}
$$

$$
\begin{aligned}
z(D,P,S,T) = V \leftarrow \\
&\neg knowsAddition(S,T) \land \\
&selectDig(D,P,S,T) = V.
\end{aligned}
$$

Alternatives:

$\forall DPST\{noMistake(D,P,S,T), mistake(D,P,S,T)\}$

$\forall DPST\{selectDig(D,P,S,T) = V \mid V \in \{0..9\}\}$

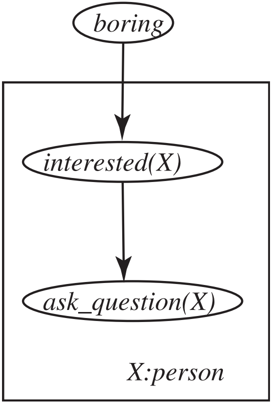
Outline

1. Semantic Science Overview
   - Ontologies
   - Data
   - Hypotheses and Theories
   - Models
2. Making Decisions
3. Relational Probabilistic Models
   - Lifted Inference
4. Existence and Identity Uncertainty

Lifted Inference

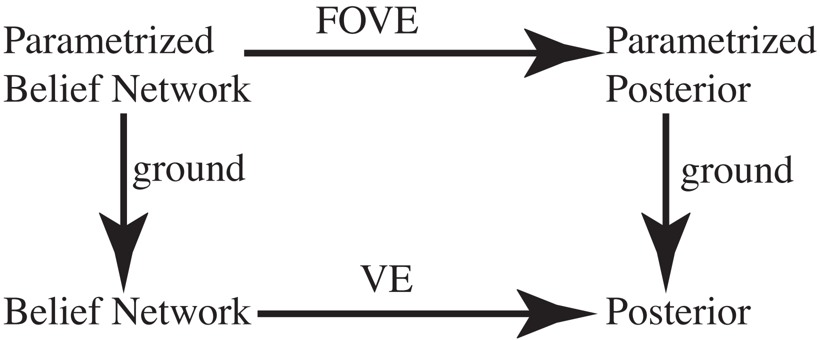
- Idea: treat those individuals about which you have the same information as a block; just count them.
- Use the ideas from lifted theorem proving - no need to ground.
- Relies on knowing the number of individuals (the population size).

Example parametrized belief network



$P(boring)$
$\forall X\ P(interested(X) | boring)$
$\forall X\ P(ask\_question(X) | interested(X))$

First-order probabilistic inference



Variable Elimination and Unification

- Multiplying parametrized factors:

$$\underbrace{[f(X,Z), p(X,a)] \quad \times \quad [p(b,Y), g(Y,W)]}_{[f(b,Z), p(b,a), g(a,W)]}$$

Variable Elimination and Unification

- Multiplying parametrized factors:

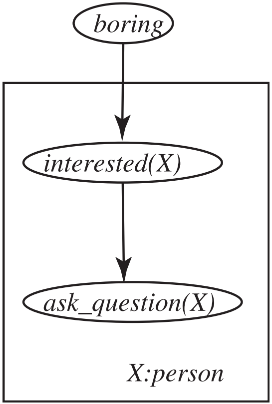
$$\underbrace{[f(X,Z),p(X,a)] \quad \times \quad [p(b,Y),g(Y,W)]}_{[f(b,Z),p(b,a),g(a,W)]}$$

  Doesn't quite work: $f(X,Z)$ can't now be used for $X=b$ but can be used when $X \neq b$.

- We split $[f(X,Z),p(X,a)]$ into

  $[f(b,Z),p(b,a)]$
  $[f(X,Z),p(X,a)]$ with constraint $X \neq b$,

Removing a parameter when summing



n people

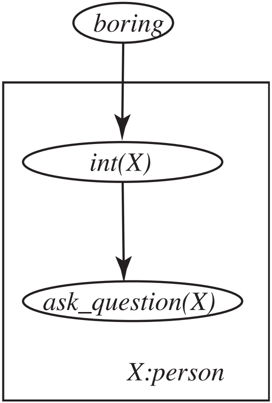
we observe no questions

Eliminate *interested*:

$$\langle \{\}, \{boring, interested(X)\}, t_1 \rangle$$
$$\langle \{\}, \{interested(X)\}, t_2 \rangle$$
$$\downarrow$$
$$\langle \{\}, \{boring\}, (t_1 \times t_2)^n \rangle$$

$(t_1 \times t_2)^n$ is computed pointwise; we can compute it in time $O(\log n)$.

Counting Elimination



$|people| = n$

[de Salvo Braz et al. 2007] and [Milch et al. 08]

Eliminate *boring*:
VE: factor on $\{int(p_1), \ldots, int(p_n)\}$
Size is $O(d^n)$ where d is size of range of interested.

Exchangeable: only the number of interested individuals matters.

Counting Formula:

#interested	Value
0	v_0
1	v_1
...	...
n	v_n

Complexity: $O(n^{d-1})$.

Potential of Lifted Inference

- Reduce complexity:

$$\textit{polynomial} \longrightarrow \textit{logarithmic}$$

$$\textit{exponential} \longrightarrow \textit{polynomial}$$

- We need a representation for the intermediate (lifted) factors that is closed under multiplication and summing out (lifted) variables.
- Still an open research problem.

Outline

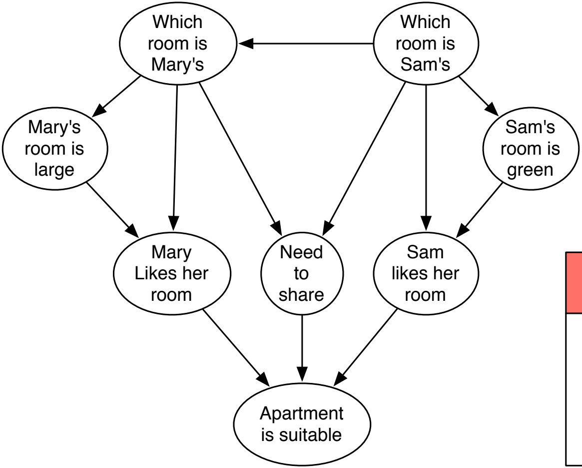
Role assignments

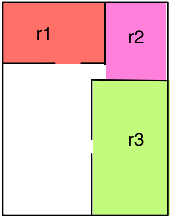
Hypothesis about what apartment Mary would like.

Whether Mary likes an apartment depends on:

- Whether there is a bedroom for daughter Sam
- Whether Sam's room is green
- Whether there is a bedroom for Mary
- Whether Mary's room is large
- Whether they share

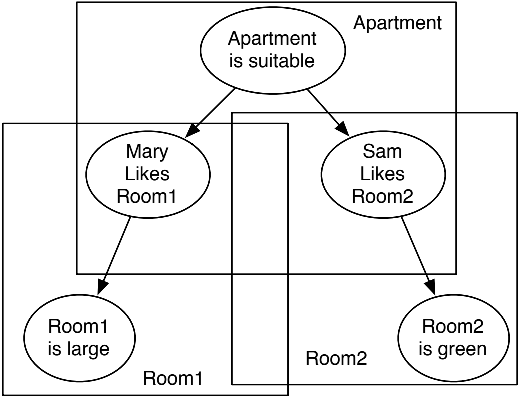
BN Representation

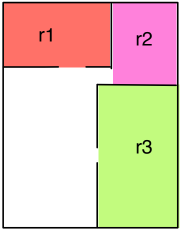




How can we condition on the observation of the apartment?

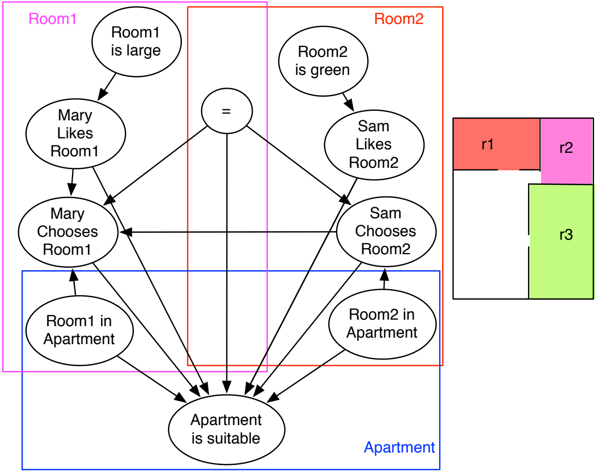
Naive Bayes representation





How do we specify that Mary chooses a room?
What about the case where they (have to) share?

Causal representation



How do we specify that Sam and Mary choose one room each, but they can like many rooms?

Conclusion

- To decide what to do an agent should take into account its uncertainty and it preferences (utility).
- Ontologies allow heterogeneous data sets to interact.
- The field of "statistical relational AI" looks at how to combine first-order logic and probabilistic reasoning.
- We need to combine many different research strands to build the World Wide Mind.

Challenges

- Representations that are heuristically and epistemologically adequate
- Condition on all of the (possibly) available evidence
- Interoperate with heterogenous data sets and allow multiple (persistent) predictions.
- Base practical actions on the best available evidence.

Research Interests