Reasoning under Uncertainty: Intro to Probability

Computer Science cpsc322, Lecture 24

(Textbook Chpt 6.1, 6.1.1)

Nov, 2, 2012

Tracing Datalog proofs in AIspace

- You can trace the example from the last slide in the AIspace Deduction Applet at http://aispace.org/deduction/ using file *ex-Datalog* available in course schedule

- Question 4 of assignment 3 asks you to use this applet

Datalog: queries with variables

```
in(alan, r123).
part_of(r123,cs_building).
in(X,Y) ← part_of(Z,Y) & in(X,Z).
```

Query: in(alan, X1).

yes(X1) ← in(alan, X1).

What would the answer(s) be?

Datalog: queries with variables

in(alan, r123).
part_of(r123,cs_building).
in(X,Y) ← part_of(Z,Y) & in(X,Z).

Query: in(alan, X1).

yes(X1) ← in(alan, X1).

What would the answer(s) be?

yes(r123).
yes(cs_building).

Again, you can trace the SLD derivation for this query
in the AIspace Deduction Applet

To complete your Learning about Logics

Review textbook and inked slides

Practice Exercise 12.B

Assignment 3

- It will be out on Mon. It is due on the 19th. Make sure you start working on it soon.
- One question requires you to use Datalog (with TopDown proof) in the AIspace.
- To become familiar with this applet download and play with the simple examples we saw in class (available at course webPage).

Paper just published in AI journal from Oxford

Towards more expressive ontology languages: The query answering problem ☆

Andrea Calì[c, b], Georg Gottlob[a, b], Andreas Pieris[a]

[a] Department of Computer Science, University of Oxford, UK

[b] Oxford-Man Institute of Quantitative Finance, University of Oxford, UK

[c] Department of Computer Science and Information Systems, Birkbeck University of London, UK

Abstract

…… query answering amounts to computing the answers to the query that are **entailed by the extensional database EDB and the ontology**. …..In particular, our new classes belong to the recently introduced family **of Datalog-based languages**, called Datalog$^{\pm}$. The basic Datalog$^{\pm}$ rules are (function-free) **Horn rules** extended with existential quantification in the head, known as *tuple-generating dependencies* (TGDs). …… We establish complexity results for answering **conjunctive queries** under sticky sets of TGDs, showing, in particular, that queries can be compiled into domain independent first-order (and thus translatable into SQL) queries over the given EDB.

Lecture Overview

- Big Transition
- Intro to Probability
- ….

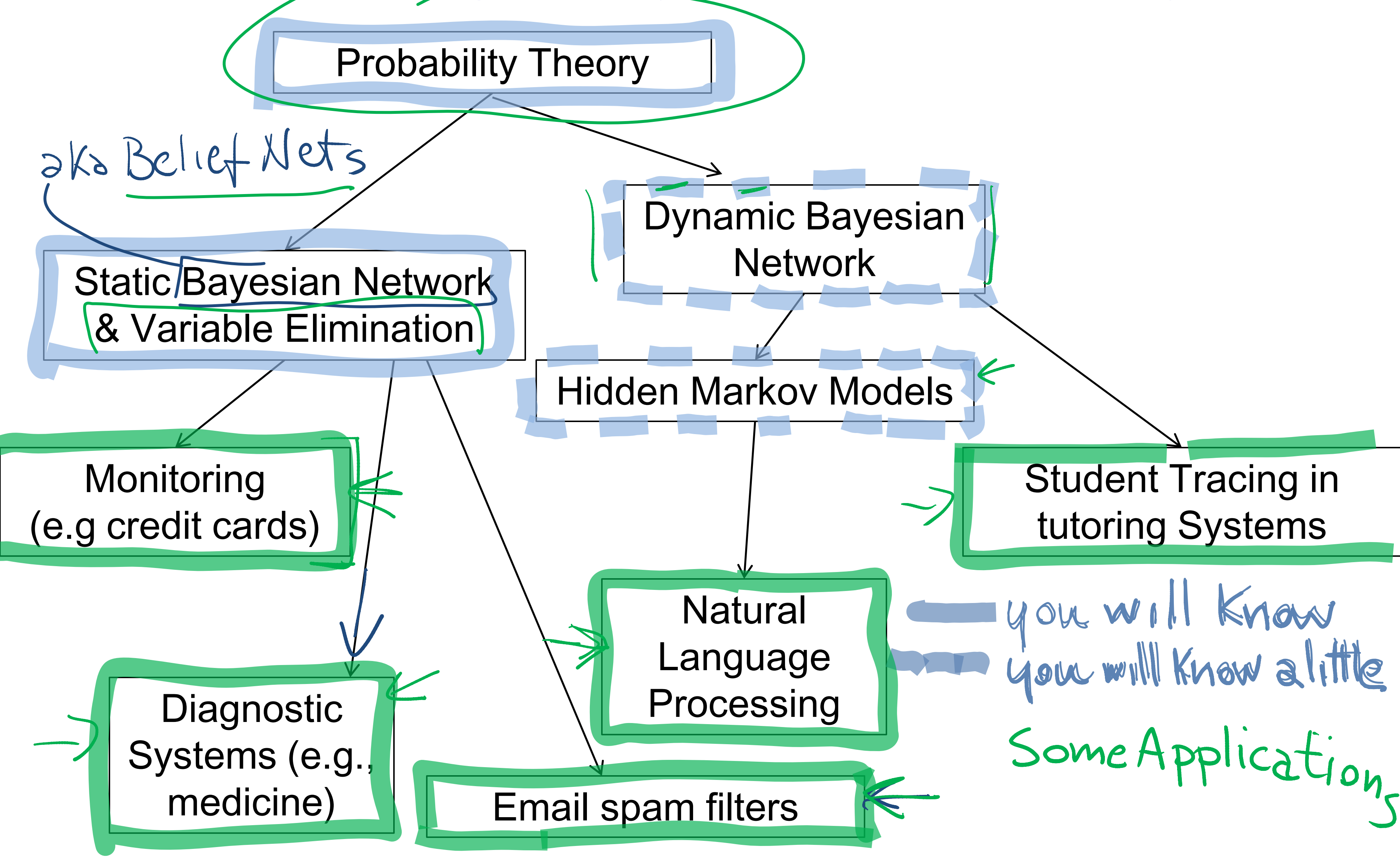
Big Picture: R&R systems

Problem		Environment: Deterministic	Environment: Stochastic
Static	Constraint Satisfaction	*Vars + Constraints* Arc Consistency Search → for CSP SLS	
Static	Query	*Logics* → CSP Search → for Inference	*Belief Nets* Var. Elimination
Sequential	Planning	*STRIPS* → CSP Search → for complex planning	*Decision Nets* Var. Elimination *Markov Processes* Value Iteration

Representation

Reasoning Technique

Answering Query under Uncertainty



Intro to Probability (Motivation)

- *Will it rain in 10 days? Was it raining 198 days ago?*
- *Right now, how many people are in this room? in this building (DMP)? At UBC? ….Yesterday?*

- AI agents (and humans ☹) are not omniscient (know everything)

  they are ignorant

- And the problem is not only predicting the future or "remembering" the past

  also current state

Intro to Probability (Key points)

- Are agents all ignorant/uncertain to the same degree? No
  it is subjective
- Should an agent act only when it is certain about relevant knowledge?
- (not acting usually has implications)
- So agents need to **represent and reason about their ignorance/ uncertainty**

Probability as a formal measure of uncertainty/ignorance

- Belief in a proposition *f* (e.g., *it is raining outside, there are 31 people in this room*) can be measured in terms of a number between 0 and 1 – this is the probability of *f*
  - The probability *f* is 0 means that *f* is believed to be
    definitely false
  - The probability *f* is 1 means that *f* is believed to be
    definitely true
  - Using 0 and 1 is purely a convention.

Random Variables

- A random variable is a **variable** like the ones we have seen in **CSP** and **Planning**, but the agent can be **uncertain about its value**.
- As usual
  - The domain of a random variable X, written $dom(X)$, is the set of values X can take
  - values are mutually exclusive and exhaustive

Examples (Boolean and discrete)

outside Raining

T F

#-of-people-rm

$[0, 10^3]$

Random Variables (cont')

- A tuple of random variables $<X_1, \ldots., X_n>$ is a complex random variable with domain..

$dom(X_1) \times \ldots \times dom(X_n)$

- Assignment $X=x$ means X has value x

outsideRaining = T

- A proposition is a Boolean formula made from assignments of values to variables

Examples

$\vee$ OR

outsideRaining = T $\wedge$ #people_rm = 47

AND

Possible Worlds

- A possible world specifies an assignment to each random variable

E.g., if we model only two Boolean variables *Cavity* and *Toothache*, then there are 4 distinct possible worlds:

w_1 $Cavity = T \wedge Toothache = T$
w_2 $Cavity = T \wedge Toothache = F$
w_3 $Cavity = F \wedge Toothache = T$
w_4 $Cavity = T \wedge Toothache = T$

cavity	*toothache*
T	T
T	F
F	T
F	F

As usual, possible worlds are mutually exclusive and exhaustive

$\boxed{w \models X=x}$ means variable X is assigned value x in world w

$w_3 \models Cavity\ F$ $w_4 \not\models Toothache = F$

Semantics of Probability

- The belief of being in each possible world *w* can be expressed as a probability $\boxed{\mu(w)}$
- For sure, I must be in one of them……so

set of all possible worlds

$$\sum_{w \in W} \mu(w) = 1$$

$\mu(w)$ for possible worlds generated by three Boolean variables: ***cavity***, ***toothache***, ***catch*** *(the probe caches in the tooth)*

cavity	***toothache***	***catch***	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
F	F	T	.144
F	F	F	.576

$\sum = 1$

Probability of proposition

equivalent, only different notation

- What is the **probability of a proposition f**?

cavity	toothache	catch	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
F	F	T	.144
F	F	F	.576

	toothache		¬ toothache	
	catch	¬ catch	catch	¬ catch
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

P(toothache=F)=.8

For any f, sum the prob. of the worlds where it is true:

$$P(f)=\Sigma_{w \models f}\ \mu(w)$$

Ex: P(*toothache = T*) = .2

Probability of proposition

- What is the **probability of a proposition** *f*?

cavity	*toothache*	*catch*	*µ(w)*
~~T~~	~~T~~	~~T~~	.108
~~T~~	~~T~~	~~F~~	.012
T	F	T	.072
T	F	F	.008
~~F~~	~~T~~	~~T~~	.016
~~F~~	~~T~~	~~F~~	.064
~~F~~	~~F~~	~~T~~	.144
~~F~~	~~F~~	~~F~~	.576

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

For any *f*, sum the prob. of the worlds where it is true:

$$P(f)=\Sigma_{w \models f}\ \mu(w)$$

P(cavity=T and toothache=F) = .08

Probability of proposition

- What is the **probability of a proposition** ***f*** ?

cavity	*toothache*	*catch*	*μ(w)*
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
~~F~~	~~F~~	~~T~~	~~.144~~
~~F~~	~~F~~	~~F~~	~~.576~~

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

For any *f*, sum the prob. of the worlds where it is true:

$$P(f)=\sum_{w \models f} \mu(w)$$

P(cavity or *toothache*) = 0.108 + 0.012 + 0.016 + 0.064 +
+ 0.072+0.08 = 0.28

= T   = T

$= 1 - (.144 + .576)$

One more example

- Weather, with domain {sunny, cloudy)
- Temperature, with domain {hot, mild, cold}
- There are now 6 possible worlds:
- What's the probability of it being cloudy or cold?

1 0.6 0.3 0.7

Weather	*Temperature*	$\mu(w)$
sunny	hot	0.10
sunny	mild	0.20
sunny	cold	0.10
cloudy	hot	0.05
cloudy	mild	0.35
cloudy	cold	0.20

- Remember
  - The probability of proposition f is defined by: $P(f) = \Sigma_{w \models f} \mu(w)$
  - sum of the probabilities of the worlds w in which f is true

One more example

- Weather, with domain {sunny, cloudy)
- Temperature, with domain {hot, mild, cold}
  - There are now 6 possible worlds:
  - What's the probability of it being cloudy or cold?
  - μ(w3) + μ(w4) + μ(w5) + μ(w6) = 0.7

	Weather	*Temperature*	*μ(w)*
w1	sunny	hot	0.10
w2	sunny	mild	0.20
w3	sunny	cold	0.10
w4	cloudy	hot	0.05
w5	cloudy	mild	0.35
w6	cloudy	cold	0.20

- Remember
  - The probability of proposition f is defined by: $P(f) = \Sigma_{w \models f} \mu(w)$
  - sum of the probabilities of the worlds w in which f is true

Probability Distributions

- A probability distribution **P** on a random variable **X** is a function *dom(X)* - > *[0,1]* such that

$$\boxed{x \rightarrow P(X=x)}$$

dom(cavity) = [T, F]

cavity? X

T → .2 P(cavity=T)

F → .8 P(cavity=F)

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

.2

.8

cavity	*toothache*	*catch*	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
~~F~~	~~T~~	~~T~~	~~.016~~
~~F~~	~~T~~	~~F~~	~~.064~~
~~F~~	~~F~~	~~T~~	~~.144~~
~~F~~	~~F~~	~~F~~	~~.576~~

Probability distribution (non binary)

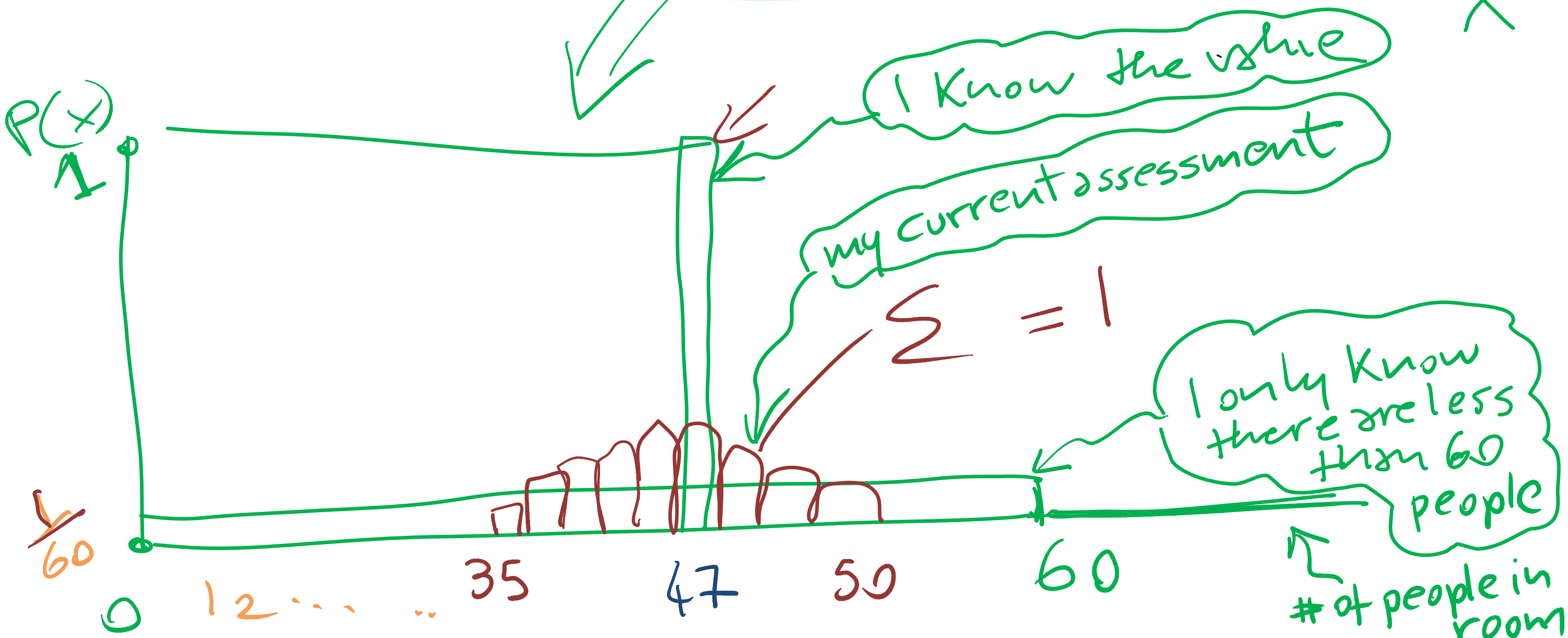
- A probability distribution P on a random variable X is a function *dom(X)* - > *[0,1]* such that

X

$$x \rightarrow P(X=x)$$

- Number of people in this room at this time

3 different distributions expressing 3 very different beliefs about X



Joint Probability Distributions

- When we have multiple random variables, their joint distribution is a probability distribution over the variable Cartesian product
  - E.g., P(<$X_1, \ldots, X_n$>)
  - Think of a joint distribution over n variables as an n-dimensional table
  - Each entry, indexed by $X_1 = x_1, \ldots, X_n = x_n$ corresponds to $P(X_1 = x_1 \wedge \ldots \wedge X_n = x_n)$
  - The sum of entries across the whole table is 1

for n Boolean vars

$\boxed{2^n}$

entries

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

Question

- If you have the joint of n variables. Can you compute the probability distribution for each variable?

yes you can compute the prob. of any proposition in

$X_1 \dots X_n$

Learning Goals for today's class

You can:

- Define and give examples of **random variables**, their domains and probability distributions.

- Calculate the **probability of a proposition f** given *μ(w)* for the set of possible worlds.

- Define a **joint probability distribution**

Next Class

More probability theory

- Marginalization
- Conditional Probability
- Chain Rule
- Bayes' Rule
- Independence

Assignment-3: Logics – out on Mon