

# Reasoning Under Uncertainty: Introduction to Probability

CPSC 322 Lecture 23

March 12, 2007

Textbook §9

# Lecture Overview

- 1 Recap
- 2 Probability Introduction
- 3 Syntax and Semantics of Probability
- 4 Probability Distributions

# Syntax of Datalog

## Definition (variable)

A **variable** starts with upper-case letter.

## Definition (constant)

A **constant** starts with lower-case letter or is a sequence of digits.

## Definition (term)

A **term** is either a variable or a constant.

## Definition (predicate symbol)

A **predicate symbol** starts with lower-case letter.

# Syntax of Datalog (cont)

## Definition (atom)

An **atomic symbol** (atom) is of the form  $p$  or  $p(t_1, \dots, t_n)$  where  $p$  is a predicate symbol and  $t_i$  are terms.

## Definition (definite clause)

A **definite clause** is either an atomic symbol (a fact) or of the form:

$$\underbrace{a}_{\text{head}} \leftarrow \underbrace{b_1 \wedge \dots \wedge b_m}_{\text{body}}$$

where  $a$  and  $b_i$  are atomic symbols.

## Definition (knowledge base)

A **knowledge base** is a set of definite clauses.

# Formal Semantics

## Definition (interpretation)

An **interpretation** is a triple  $I = \langle D, \phi, \pi \rangle$ , where

- $D$ , the **domain**, is a nonempty set. Elements of  $D$  are **individuals**.
- $\phi$  is a mapping that assigns to each constant an element of  $D$ . Constant  $c$  **denotes** individual  $\phi(c)$ .
- $\pi$  is a mapping that assigns to each  $n$ -ary predicate symbol a relation: a function from  $D^n$  into  $\{TRUE, FALSE\}$ .

# Variables

## Definition (variable assignment)

A **variable assignment** is a function from variables into the domain.

- Given an interpretation and a variable assignment, each term denotes an individual and each clause is either true or false.
- A clause containing variables is true in an interpretation if it is true **for all** variable assignments.
  - Variables are **universally quantified** in the scope of a clause.
- Now we can use our previous definition of logical entailment.

# Lecture Overview

- 1 Recap
- 2 Probability Introduction**
- 3 Syntax and Semantics of Probability
- 4 Probability Distributions

# Using Uncertain Knowledge

- Agents don't have complete knowledge about the world.
- Agents need to make decisions based on their uncertainty.
- It isn't enough to assume what the world is like.  
**Example:** wearing a seat belt.
- An agent needs to reason about its uncertainty.
- When an agent makes an action under uncertainty, it is gambling  $\implies$  probability.



# Probability

- Probability is formal measure of uncertainty. There are two camps:
- **Frequentists**: believe that probability represents something *objective*, and compute probabilities by counting the frequencies of different events
- **Bayesians**: believe that probability represents something *subjective*, and understand probabilities as degrees of belief.
  - They compute probabilities by starting with **prior beliefs**, and then **updating** beliefs when they get new data.
  - **Example**: Your degree of belief that a bird can fly is your measure of belief in the flying ability of an individual based only on the knowledge that the individual is a bird.
  - Other agents may have different probabilities, as they may have had different experiences with birds or different knowledge about this particular bird.
  - An agent's belief in a bird's flying ability is affected by what the agent knows about that bird.

# Numerical Measures of Belief

- Belief in proposition,  $f$ , 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.
- $f$  has a probability between 0 and 1, doesn't mean  $f$  is true to some degree, but means you are ignorant of its truth value. Probability is a measure of your ignorance.

# Random Variables

## Definition (random variable)

A **random variable** is a variable that is randomly assigned one of a number of different values.

## Definition (domain)

The **domain** of a random variable  $X$ , written  $dom(X)$ , is the set of values  $X$  can take.

- A tuple of random variables  $\langle X_1, \dots, X_n \rangle$  is a complex random variable with domain  $dom(X_1) \times \dots \times dom(X_n)$ .
  - Often the tuple is written as  $X_1, \dots, X_n$ .
- Assignment  $X = x$  means variable  $X$  has value  $x$ .
- A **proposition** is a Boolean formula made from assignments of values to variables.

# Lecture Overview

- 1 Recap
- 2 Probability Introduction
- 3 Syntax and Semantics of Probability**
- 4 Probability Distributions

# Possible World Semantics

## Definition (possible world)

A **possible world** specifies an assignment of one value to each random variable.

- $w \models X = x$  means variable  $X$  is assigned value  $x$  in world  $w$ .
- Logical connectives have their standard meanings:

$$w \models \alpha \wedge \beta \text{ if } w \models \alpha \text{ and } w \models \beta$$

$$w \models \alpha \vee \beta \text{ if } w \models \alpha \text{ or } w \models \beta$$

$$w \models \neg\alpha \text{ if } w \not\models \alpha$$

- Let  $\Omega$  be the set of all possible worlds.

# Semantics of Probability: finite case

When there are a finite number of possible worlds:

- Define a nonnegative measure  $\mu(w)$  to each world  $w$  so that  $\sum_w \mu(w) = 1$ 
  - The measures of the possible worlds sum to 1.
  - The measure specifies how much you think the world  $w$  is like the real world.

## Definition (probability)

The **probability** of proposition  $f$  is defined by:

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

# Axioms of Probability: finite case

Four axioms define what follows from a set of probabilities:

- 1 **Axiom 1**  $P(f) = P(g)$  if  $f \leftrightarrow g$  is a tautology.
  - That is, logically equivalent formulae have the same probability.
- 2 **Axiom 2**  $0 \leq P(f)$  for any formula  $f$ .
- 3 **Axiom 3**  $P(\tau) = 1$  if  $\tau$  is a tautology.
- 4 **Axiom 4**  $P(f \vee g) = P(f) + P(g)$  if  $\neg(f \wedge g)$  is a tautology.

These axioms are sound and complete with respect to the semantics.

- if you obey these axioms, there will exist some  $\mu$  which is consistent with your  $P$
- there exists some  $P$  which obeys these axioms for any given  $\mu$

# Semantics of Probability: general case

In the general case (possibly infinite set of possible worlds) we have a measure on sets of possible worlds, satisfying:

- $\mu(S) \geq 0$  for  $S \subseteq \Omega$
- $\mu(\Omega) = 1$
- $\mu(S_1 \cup S_2) = \mu(S_1) + \mu(S_2)$  if  $S_1 \cap S_2 = \{\}$ .



# Lecture Overview

- 1 Recap
- 2 Probability Introduction
- 3 Syntax and Semantics of Probability
- 4 Probability Distributions**

# Probability Distributions

## Definition (probability distribution)

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

$$x \mapsto P(X = x).$$

- When  $dom(X)$  is infinite we need a **probability density** function.

# Joint Distribution and Marginalization

- When there are multiple random variables, their **joint distribution** is a probability distribution over the variables' Cartesian product
  - E.g.,  $P(X, Y, Z)$  means  $P(\langle X, Y, Z \rangle)$ .
  - Think of a joint distribution over  $n$  variables as an  $n$ -dimensional table
  - Each entry, indexed by  $X_1 = x_1, \dots, X_n = x_n$ , corresponds to  $P(X_1 = x_1 \wedge \dots \wedge X_n = x_n)$ .
  - The sum of entries across the whole table is 1.
- Given the joint distribution, we can compute distributions over smaller sets of variables through **marginalization**:
  - E.g.,  $P(X, Y) = \sum_{z \in \text{dom}(Z)} P(X, Y, Z = z)$ .
  - This corresponds to summing out a dimension in the table.
  - The new table still sums to 1.