CSC421 Intro to Artificial Intelligence

UNIT 20:
Uncertainty

Midterm Review

- What is confusing:
  - Greece <---isFrom-- George ---isWearing->Shirt
- Write in PL
  - George and Aristotle are from Greece
  - George and Aristotle are wearing a red shirt
- Prolog (why infinite loop ?)
  - parent(Alice, Bob).
  - parent(Bob, Charles).
  - ancestor(X,Z) :- ancestor(X,Y), parent(Y,Z).
  - ancestor(X,Z) :- parent(X,Z).
- Prove ancestor(Alice, Charles)
  - Forward Chaining, Backward Chaining

Making decisions under uncertainty

- Suppose I believe the following:
  - P(30min gets me there on time) = 0.04
  - P(1 hr gets me there on time) = 0.70
  - P(24 hrs gets me there on time) = 0.98
- Which action to choose ?
- Depends on my preferences
- Utility theory = represent preferences
- Decision theory = utility theory + probability theory

Probability Basics

- Set Ω – the sample space
- $\omega \in \Omega$ is a sample point/atomic event, outcome, possible world
- Probability space/model is a sample space with an assignment $P(\Omega)$ for every point
  - $0 <= P(\omega) <= 1$
  - $\sum P(\omega) = 1$
- An event is a subset of Ω

Axiomatic Foundation of Probability

- Russian mathematician Kolmogorov (born 1903 died 1987)
- Axioms (1925)
  - $0 <= P(a) <= 1$
  - $P(true) = 1$, $P(false) = 0$
  - $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Random Variables

- A random variable is a function from sample points to some range e.g the reals or booleans
  - e.g. Odd(1) = true
- P induces a probability distribution for any r.v X
- P(Odd = true) = P(1) + P(3) + P(5) = ½
- How can you rewrite the above example for the general case ?

Propositions

- Think of a proposition as the event (set of sample points) where the proposition is true
- Given boolen random variables A
  - Event a = set of sample points where $A(\omega) = 1$
- Often in AI sample points are defined by the values of a set of random variables i.e., the sample space is the Cartesian product of the ranges of the variables
- With boolean variables sample point = propositional logic model
  - A = true, B = false

Example

- $(a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$
- $P(a \vee b) = ?$

Example

- $(a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$
- $P(a \vee b) = P(\neg a \wedge b) + P(a \wedge \neg b) + P(a \wedge b)$

Syntax for propositions

- Propositional or Boolean variables
  - e.g. Student
  - Student = true is a proposition sometimes written student
- Discrete RV (finite or countable)
  - e.g. Lecture is one of <boring, ok, interesting>
  - Lecture is boring is a proposition
- Continuous (bounded or unbounded)
  - Temp = 21.6 also allow Temp < 22.0
- Arbitrary boolean combinations of basic propositions

Prior Probability

- Prior or unconditional probabilities
  - P(Student=true) = 0.1 and P(Class=boring) = 0.7
- Probability distribution gives values for all possible assignments
  - P(Lecture) = <0.72, 0.2, 0.08>
- Joint probability distribution for a set of r.v.s gives the probability of every atomic event on those r.v.s (i.e. Every sample point)
- P(Lecture, Student) = 3 * 2 matrix of values

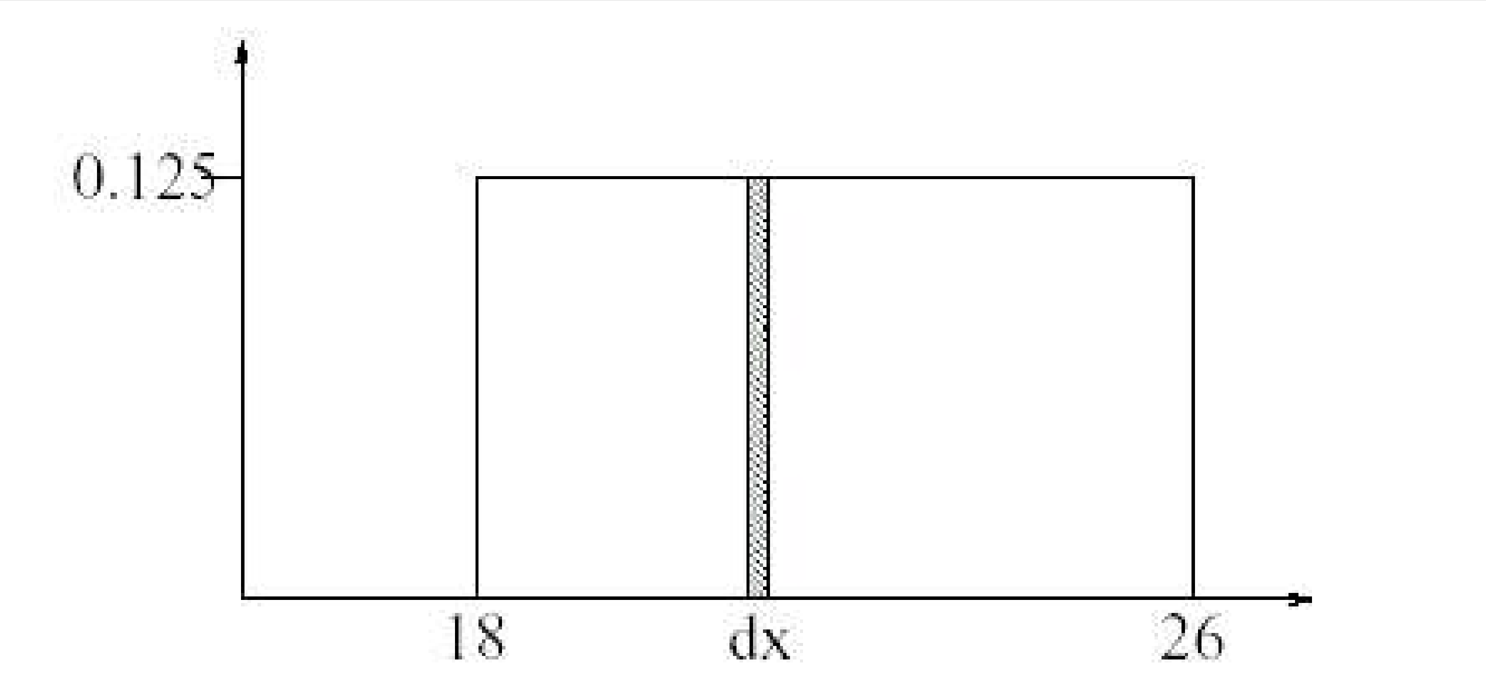
Joint Probability

Lecture =	boring	ok	interesting
Student = true	0.144	0.04	0.016
Student = false	0.576	0.16	0.064

- Every question about a domain can be answered by the joint probability because every event is a set of sample points (probability is the sum)
- Inference using full joint distributions
  - Marginalization
  - Conditioning

Probabilities for continuous variables

- Express distribution as a parametrized function of value: P(X = x) = U[18, 26](x) = uniform density between 18, 26
- P density (integrates to 1)



Continuous rvs

- P(X = 20.5) = 0.125 really means
  - $\lim_{dx \to 0} P(20.5 <= X <= 20.5 + dx) / dx = 0.125$
- Gaussian Density