Local Probabilistic Models: Continuous Variable CPDs

Sargur Srihari
srihari@cedar.buffalo.edu

Topics

1. Simple discretizing loses continuity
2. Continuous Variable CPDs
3. Linear Gaussian Model
   – Example of car movement
4. Case Study: Robot Motion and Sensors
5. Hybrid Models
6. CLG Network
7. Thermostat

Continuous Variables

- Some variables best modeled as continuous
  - Ex: position, velocity, temperature, pressure
- We cannot use a table representation
- How about circumventing the issue by discretizing all continuous variables?
  - Discretization will not do
  - To get accurate model need fine discretization
    - Applying PGM to robot navigation task
      - with $15cm$ granularity in x and y coordinates results in each variable having a thousand values and more than *one million* discretized values for robot position
      - CPDs of this magnitude are outside range of most systems

Discretization loses Continuity

- When we discretize a variable we lose much of the structure that characterizes it
  - Million values of robot position cannot always be associated with arbitrary probability
- Basic continuity assumptions imply relationships between probabilities of nearby discretized values
  - Such constraints are hard to capture in a discrete distribution where there is no notion that two values are close to each other

CPD as a distribution

- Nothing in formulation of Bayesian network requires restricting attention to discrete variables
- Only requirement is that the CPD $P(X|Pa_X)$ represent for every assignment of values Pa_X, a distribution over X
- In this case X is continuous
- We might also have some of X's parents are continuous

Purely Continuous Case

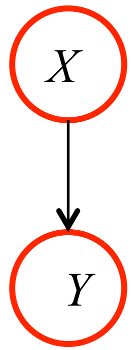
- Many possible models one could use
- Most commonly used parametric form for continuous density functions is the Gaussian
- We can see how it can be used in the context of a Bayesian network

Both Parent/Child are Continuous

- Representing the dependency of a continuous variable Y on a continuous parent X
- A simple solution:
  - Model Y as a Gaussian
  - Whose parameters depend on X
  - Common solution:
    - Decide that mean of Y is a linear function of X
    - For example:

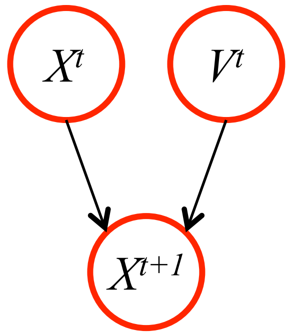
$$P(Y|x) \sim N(-2x + 0.9;\ 1)$$

    - This dependence is called a *linear Gaussian model*
    - It extends to multiple continuous parents



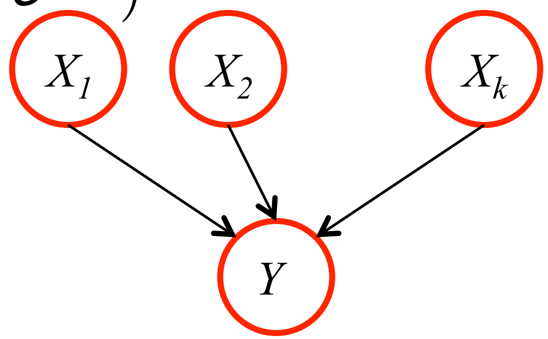
Linear Gaussian Example: Car motion

- Car moving along straight line
  - Position at t th second: X^t (at meter $\#510$)
  - Velocity at t th second: V^t (meters/sec)
  - Then $X^{t+1} = X^t + V^t$
  - If $V^t = 15$ meters/sec, then $X^{t+1} = \#525$
- If there is stochasticity in motion,
  - $X^{t+1} \sim N(525, 5)$ variance is 5 meters



Definition of Linear Gaussian Model

- Let Y be a continuous variable with continuous parents $X_1,.. X_k$
  - We say that Y has a *linear Gaussian model* with parameters $\beta_0,.. \beta_k$ and σ^2 such that
    $P(Y|x_1,..x_k) \sim N(\beta_0 + \beta_1 x_1 + .. \beta_k x_k; \sigma^2)$
    - In vector notation
      $P(Y|\boldsymbol{x}) \sim N(\beta_0 + \beta^T \boldsymbol{x}; \sigma^2)$

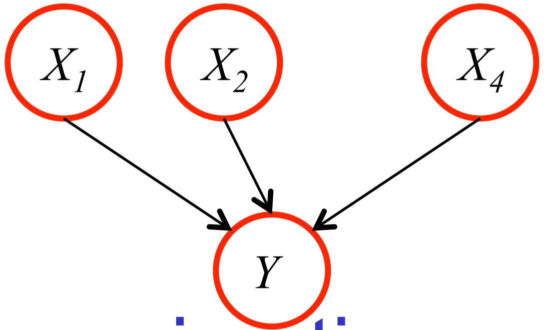


- Formulation says that Y is a
  - linear function of the variables $X_1,.. X_k$ with added Gaussian noise with mean 0 and variance σ^2:
  - $Y = \beta_0 + \beta_1 x_1 + .. \beta_k x_k + \varepsilon$
    - where ε is a Gaussian random variable
      - with mean 0 and variance σ^2

Properties of Linear Gaussian

- *Advantage*: Simple model that captures many dependencies
- *Disadvantage*: cannot capture dependence of variance of child on values of parents
  - Can be extended
    - E.g., mean of Y is $\sin(x_1)^{x2}$

      variance is $(x_3/x_4)^2$



- But linear Gaussian is a good approximation
- Provides an alternative for representing multivariate Gaussian distributions

Case Study: Robot Motion & Sensors

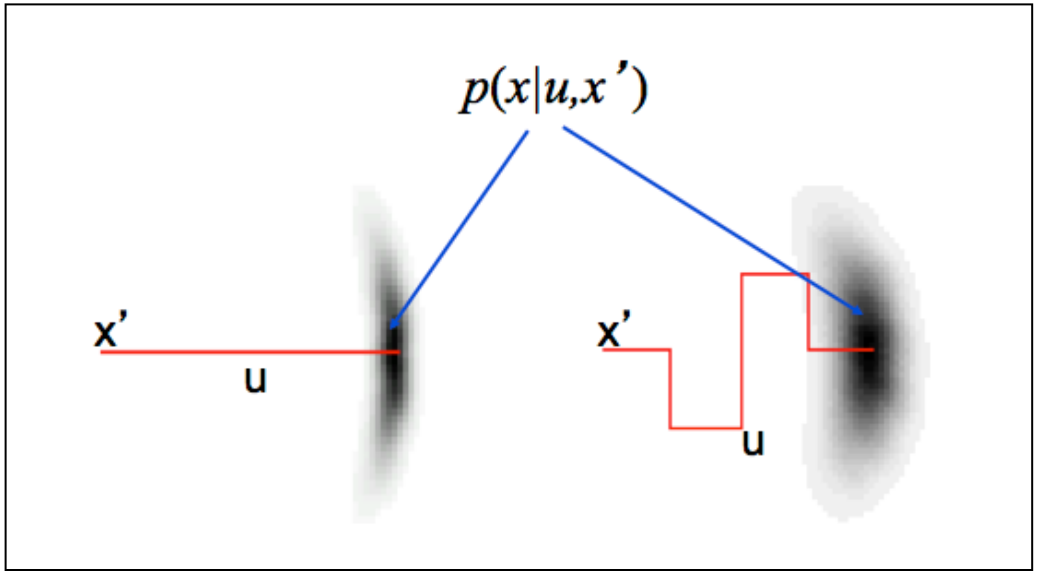
- Robot localization
  - Robot must keep track of its location as it moves in an environment
  - Obtains sensor readings as it moves

- Hybrid model: two local CPDs
  1. Robot dynamics (cont.): $P(L'|L,A)$
     - Specifies distribution of its position L' at next time step given its current position L and action taken A
  2. Robot sensor model (discrete) $P(D|L)$
     - Distribution over its observed sensor reading S at the current time given its current location L

Robot Dynamics Model: $P(L'|L,A)$

- Specifies distribution of position at time step L' given its current position L and action taken A
  - Robot location L is given by (X,Y,θ)
    - X,Y coordinates and angular orientation θ
  - Action A specifies distance to travel and rotation from current θ
- $P(L'|L,A)$ is product of two independent Gaussians $P(\theta'|\theta,A)$ and $P(X',Y'|X,Y,A)$

Ex: Robot Dynamics Model

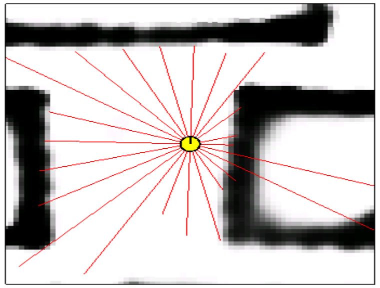


Different notation
From previous slide
$x' \rightarrow L$
$x \rightarrow L'$
u$\rightarrow A$

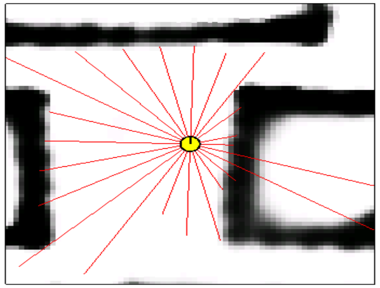
- Distributions for two different actions $\mathbf{u}$
  - $\mathbf{x}'$=previous location
  - $\mathbf{u}$=action
  - $\mathbf{x}$=current location
- Two banana-shaped distributions

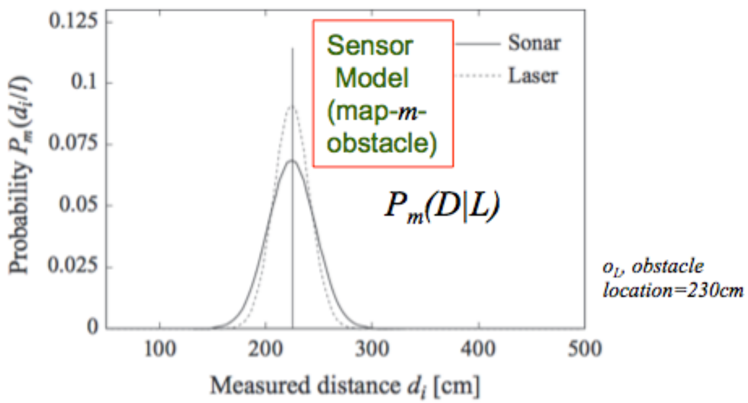
Robot Sensor model: $P(D|L)$

- Obtain sensor readings that depend on location
- Sensors
  - Contact sensors: Bumpers
  - Internal sensors
    - Accelerometers (spring-mounted masses)
    - Gyroscopes (spinning mass, laser light)
    - Compasses, inclinometers (earth magnetic field, gravity)
  - Proximity sensors
    - Sonar (time of flight)
    - Radar (phase and frequency)
    - Laser range-finders (triangulation, tof, phase)
    - Infrared (intensity)
  - Visual sensors: Cameras
  - Satellite-based sensors: GPS
  - D is distance between robot and nearest obstacle



Discrete Sensor Model





- D=distance between robot and obstacle along direction of sensor $D = \{d_1, d_2, \ldots, d_K\}$

  d_1=laser, d_2=sonar

- Individual measurements are independent given the robot location and map

$$P_m(D \mid L) = \prod_{i=1}^{K} P(d_k \mid L) \qquad m=map \quad L=location$$

$$P_m(D|L)=N(o_L;\sigma^2)$$

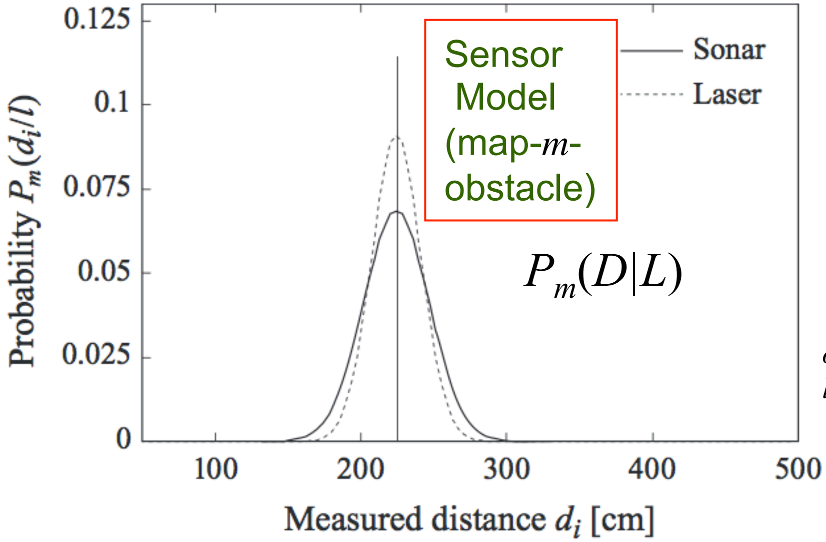
 15

Two Models: Dynamics and Sensors (2)

Robot Dynamics Model

$P(L'|L,A)$

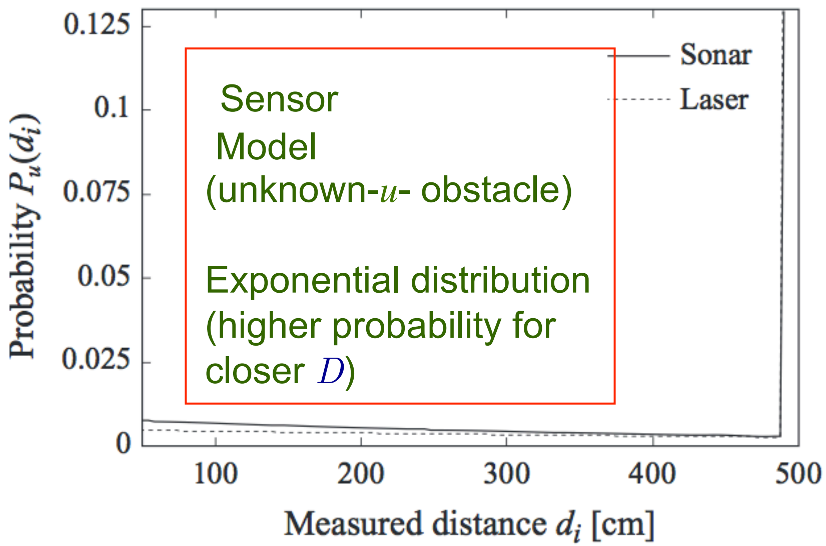
(a)



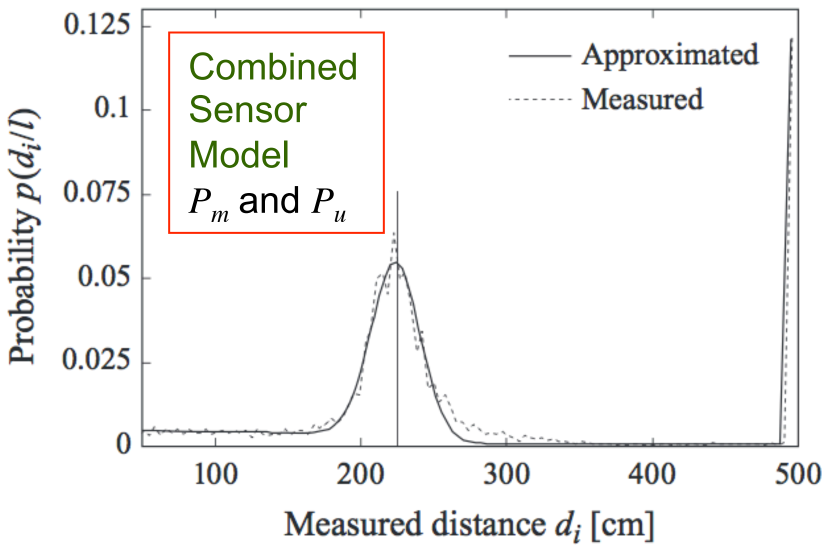
o_L, *obstacle location=230cm*

(b)

$P_u(D|L)$

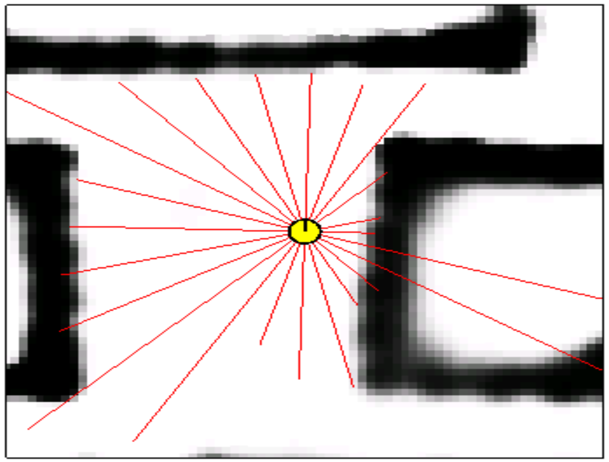


$P(D|L)$



Probabilistic Robotics

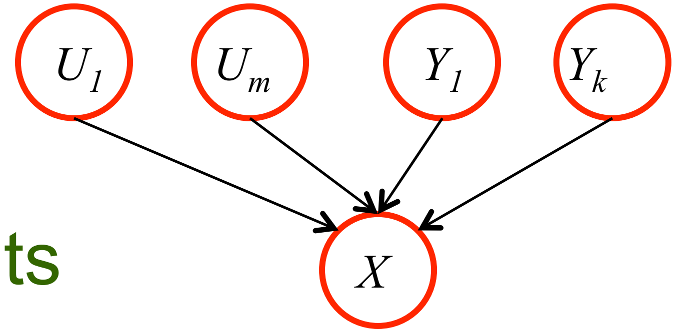
- Odometry
  - Internal sensors (compass, gyroscope)
  - Proximity sensors (sonar, radar)
- Odometry information is inherently noisy
- Called *probabilistic kinematics*



Hybrid Models

- Incorporate both discrete and continuous variables
- Case 1: Continuous child X
- If we ignore discrete parents, CPD of X can be represented as a linear Gaussian of continuous parents
- Simplest way of making continuous variable X depend on discrete variable U is to define a different set of parameters for every value of the discrete parent

Conditional Linear Gaussian (CLG)



- CLG CPD: X is continuous
  - $U=\{U_1,..U_m\}$ are discrete parents
  - $Y=\{Y_1,..,Y_k\}$ are continuous parents
  - For every u we have $k+1$ coefficients $a_{u,i}$ such that

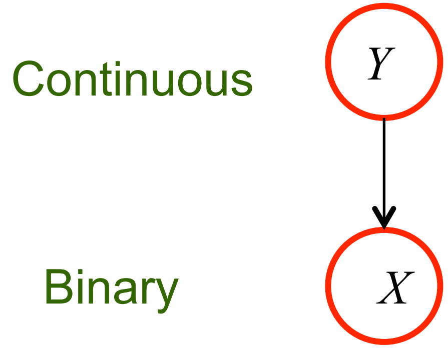
$$\boxed{p(X \mid \mathrm{u}, \mathrm{y}) = N\left(a_{\mathrm{u},0} + \sum_{i=1}^{k} a_{\mathrm{u},i} y_i ; \sigma_{\mathrm{u}}^2\right)}$$

- CLG network: every discrete variable has only discrete parents and every continuous variable has a CLG CPD
  - A continuous variable cannot have discrete children
  - Distribution is a mixture: weighted average of Gaussians

Discrete Child with a Continuous Parent

- Threshold model

$$P(x^1) = \begin{cases} 0.9 & y \le 65 \\ 0.05 & otherwise \end{cases}$$



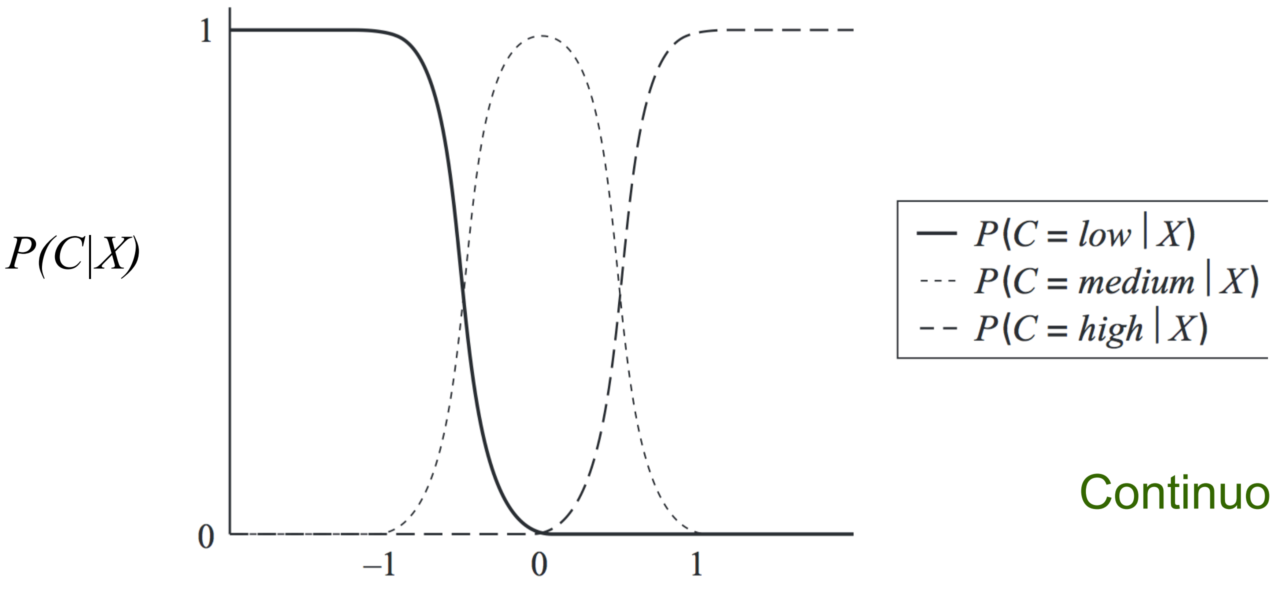
$P(x^0)=1 - P(x^1)$

- Y is the temperature in Fahrenheit and X is the thermostat turning the heater on
- Threshold model has abrupt change in probability with Y, i.e., from *0.9* to *0.05*
  - Can use the logistic model to fix this

$P(x^1 | Y) = sigmoid(w_0 + w_i Y)$

Generalized Linear Model for a Thermostat

- Sensor has three values: *low, medium, high*



X: Continuous parent