ECE 4960

Prof. Kirstin Hagelskjær Petersen
kirstin@cornell.edu

Vivek Thangavelu
vs353@cornell.edu

Fast Robots

Lecture Outline

- Probability in Robotics:
  - The Why
  - The How
- Robot Environment Model
- Markov Processes
- Bayes Filter

Probability Recap

Random Variable

- A Random variable is described informally as a variable whose values depend on the outcomes of a random phenomenon.
- It is usually denoted by capital letters. It is a function

$$X : \Omega \rightarrow \mathbb{R}$$

- Here, Ω is the sample space i.e. the set of all possible outcomes and $\mathbb{R}$ is the set of all real numbers

Axioms of Probability

- Let X denote a random variable and x denote a specific event that X might take on
- To denote the probability that the random variable X has value x

$$P(X = x) \; or \; p(x)$$

- Probabilities sum to one

$$\sum_x P(X = x) = 1$$

- Probabilities are always non-negative, that is,

$$P(X = x) \geq 0$$

Joint and Conditional Probability

- The joint distribution of two random variables X and Y is given by

$$p(x,y) \; = P(X = x \; and \; Y \; = y)$$

This expression describes the probability of the event that the random variable X takes on the value x and that Y takes on the value y

- Suppose we already know that Y 's value is y, and we would like to know the probability that X's value is x conditioned on that fact.

- Such a probability is known as the conditional probability **and is** given by

$$p(x \mid y) = \frac{p(x,y)}{p(y)}$$

[Also known as the **general product rule of probability]**

Marginal Probability

- Following the axioms of probability and conditional probability, we have *the theorem of total probability that gives:*

$$p(x) = \sum_y p(x \mid y)\, p(y)$$

- Such a probability is known as the marginal probability

Independence

- Independence is a fundamental notion in probability theory
- Two events are independent if the occurrence of one does not affect the probability of occurrence of the other
- Let random variables X and Y take on the value x and y, respectively
- If X and Y are independent, then

$$p(x, y) = p(x)\, p(y)$$

$$p(x|y) = p(x) \text{ and } p(y|x) = p(y)$$

Conditional Independence

- If X and Y are conditional independent given $\mathrm{Z} = \mathrm{z}$, then

$$p(x, y|z) = p(x|z)p(y|z)$$

$$p(x|z, y) = p(x|z) \; and \; p(y|z, x) = p(y|z)$$

Bayes Theorem

likelihood prior

$$p(x|y) = \frac{p(y|x)\; p(x)}{p(y)}$$

posterior

marginal likelihood
(constant)

Bayes Theorem

- If x is a quantity that we would like to infer from y, then the probability $p(x)$ will be referred to as prior probability distribution, and y is called the data (e.g., a sensor measurement)
- $p(x)$ summarizes the knowledge we have regarding X prior to incorporating the data y
- The probability $p(x|y)$, computed after incorporating the data, is called the posterior probability distribution over X
- **Bayes Theorem** provides a convenient way to compute the posterior probability $p(x|y)$ using the "inverse" conditional probability $p(y|x)$ and the prior probability $p(x)$

$$\underbrace{\boxed{p(x|y)}}_{\text{posterior}} = \frac{\overbrace{\boxed{p(y|x)}}^{\text{likelihood}}\ \overbrace{\boxed{p(x)}}^{\text{prior}}}{p(y)}$$

Bayes Theorem

- If x is a quantity that we would like to infer from y, then the probability $p(x)$ will be referred to as prior probability distribution, and y is called the data (e.g., a sensor measurement)
- $p(x)$ summarizes the knowledge we have regarding X prior to incorporating the data y
- The probability $p(x|y)$, computed after incorporating the data, is called the posterior probability distribution over X
- **Bayes Theorem** provides a convenient way to compute the posterior probability $p(x|y)$ using the "inverse" conditional probability $p(y|x)$ and the prior probability $p(x)$

$$p(x|y) = \frac{p(y|x)\,p(x)}{p(y)}$$

$$p(x|y) = \frac{p(y|x)\,p(x)}{\sum_{x'} p(y|x')\,p(x')}$$

Bayes Theorem

- If x is a quantity that we would like to infer from y, then the probability $p(x)$ will be referred to as prior probability distribution, and y is called the data (e.g., a sensor measurement)
- $p(x)$ summarizes the knowledge we have regarding X prior to incorporating the data y
- The probability $p(x|y)$, computed after incorporating the data, is called the posterior probability distribution over X
- **Bayes Theorem** provides a convenient way to compute the posterior probability $p(x|y)$ using the "inverse" conditional probability $p(y|x)$ and the prior probability $p(x)$

$$p(x|y) = \frac{p(y|x)\, p(x)}{p(y)}$$

$$p(x|y) = \frac{p(y|x)\, p(x)}{\sum_{x'} p(y|x')\, p(x')}$$

$$\underbrace{\boxed{p(x|y)}}_{\textit{posterior}} = \eta \underbrace{\boxed{p(y|x)}}_{\textit{likelihood}} \underbrace{\boxed{p(x)}}_{\textit{prior}}$$

η *is the normalizing constant*

Probability for Robotics
THE WHY

Uncertainty is Inherent

- Uncertainty is inherent in the world. Resistance in futile!
- Five Major factors
  - Environment
  - Robot actions
  - Sensors
  - Models
  - Computation
- Traditionally, such uncertainty has mostly been ignored in robotics.
- As robots are moving away from factory floors into increasingly unstructured environments, the ability to cope with uncertainty is critical for building successful robots.

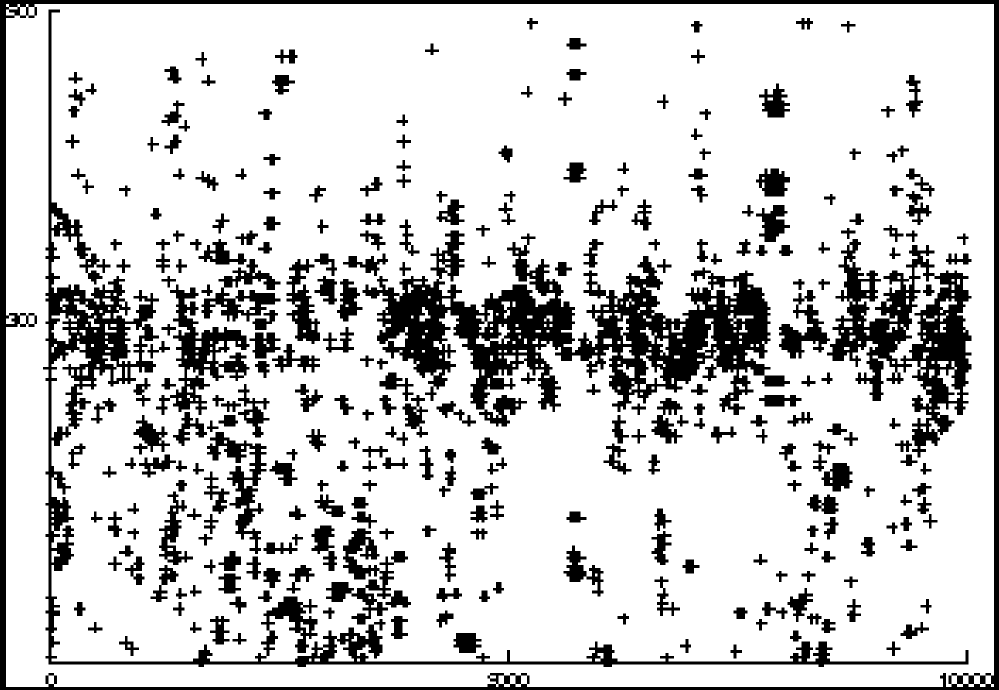
1. Environments

1. Environments

- Physical worlds are inherently unpredictable
- Uncertainty in well-structured environments such assembly lines is usually small
- Environments such as highways and private homes are highly dynamic and unpredictable.

2. Sensors

- Sensors are inherently limited in what they can perceive
- Limitations arise from two primary factors.
    - Range and resolution of a sensor is subject to physical laws.
        - For example, cameras can't see through walls, and even within the perceptual range the spatial resolution of camera images is limited.
    - Noise
        - perturbs sensor measurements in unpredictable ways
        - limits the information that can be extracted from sensor measurements.



Typical data obtained from a laser-range sensor in an office environment for a "true" range of 300 cm and a maximum range of 500 cm

3. Robots

- Robot actuation involves motors that are, at least to some extent, unpredictable, due effects like control noise and wear-and-tear
- Low-cost mobile robots, can be extremely inaccurate
- Some actuators, such as heavy-duty industrial robot arms, are quite accurate, but much more expensive

4. Models

- Models are inherently inaccurate
- Models are abstractions of the real world
- They only partially model the underlying physical processes of the robot and its environment
- Model errors are a source of uncertainty that has largely been ignored in robotics, despite the fact that most robotic models used in state-or-the-art robotics systems are rather crude

5. Computation

- Robots are real-time systems, which limits the amount of computation that can be carried out
- Many state-of-the-art algorithms are approximate, achieving timely response through sacrificing accuracy

Probabilistic Approaches

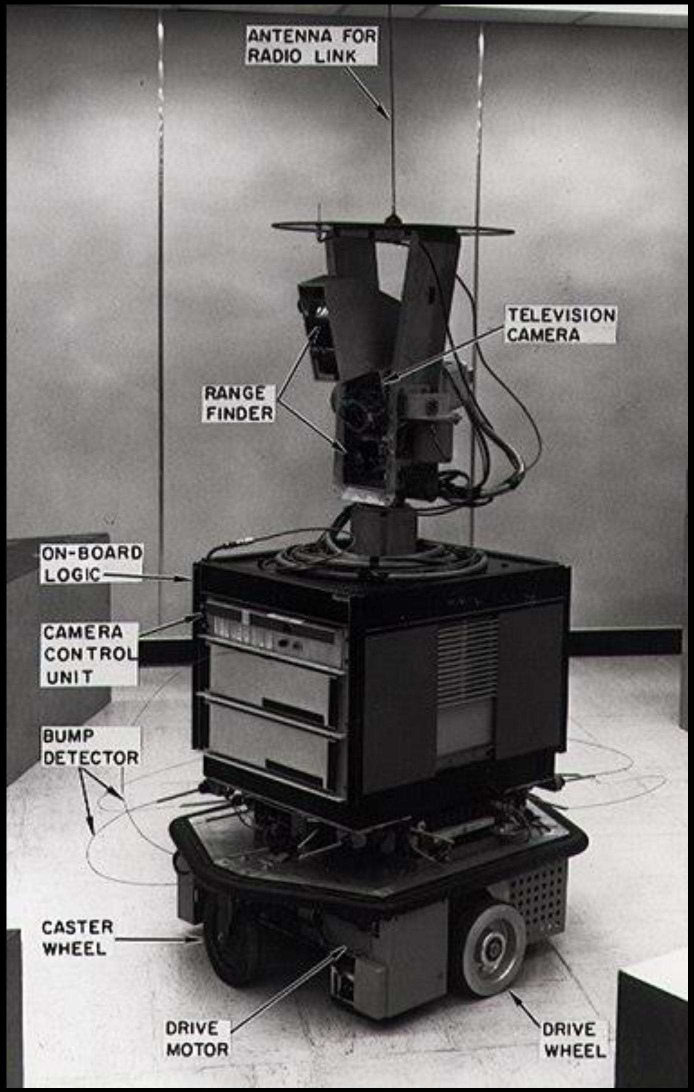
"A robot that carries a notion of its own uncertainty and that acts accordingly is superior to one that does not."

- Probabilistic Robotics by Thrun, Burgard, Fox

- Probabilistic Approaches in contrast to traditional approaches such as model-based motion planning techniques or reactive behavior-based approaches:
  - tend to be more robust to sensor and model limitations
  - weaker requirements on the accuracy of the robot's models
- In fact, they are the only known working solutions to hard robotic estimation problems, such as the localization problem

Probabilistic Robotics

- Shakey the Robot was the first general-purpose mobile robot to be able to reason about its own actions
- Developed from approximately 1966-1972 and funded by DARPA
- It could
    - travel from one location to another
    - turn the light switches on and off
    - open and close the doors
    - climb up and down rigid objects
    - push movable objects around
- Interesting results: A* Search Algorithm and Hough Transform



Shakey the Robot
1972

The WHY

+ Explicitly represent the uncertainty using probability theory
+ Can accommodate inaccurate models
+ Can accommodate imperfect sensors
+ Robust in real-world applications
+ Best known approach to many hard robotics problems

- Computationally demanding
- Need to Approximate
- False assumptions

Poll

The dress is a photograph that became a viral internet sensation on 26 February 2015, when viewers disagreed over whether the dress pictured was colored black and royal blue, or white and gold.

Where does the uncertainty arise from?

a. Environment
b. Sensors
c. Models

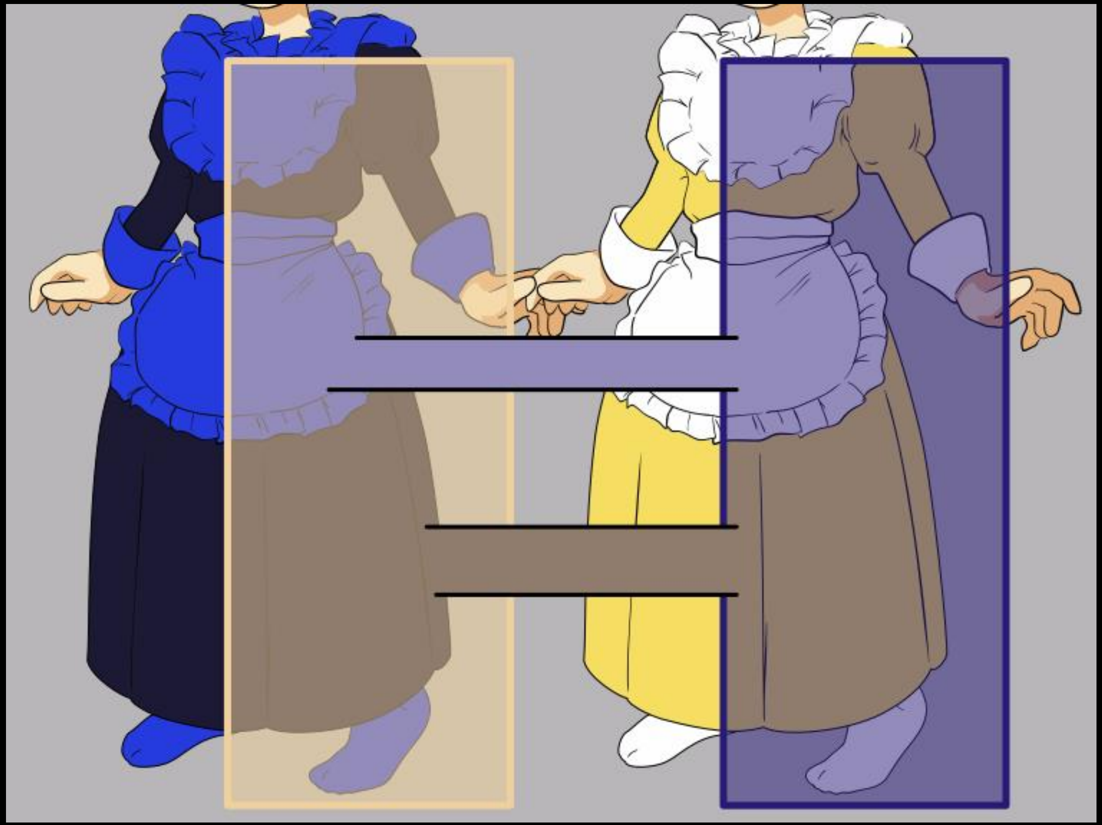
Poll

The dress is a photograph that became a viral internet sensation on 26 February 2015, when viewers disagreed over whether the dress pictured was colored black and royal blue, or white and gold.

Digitally remastered

Poll

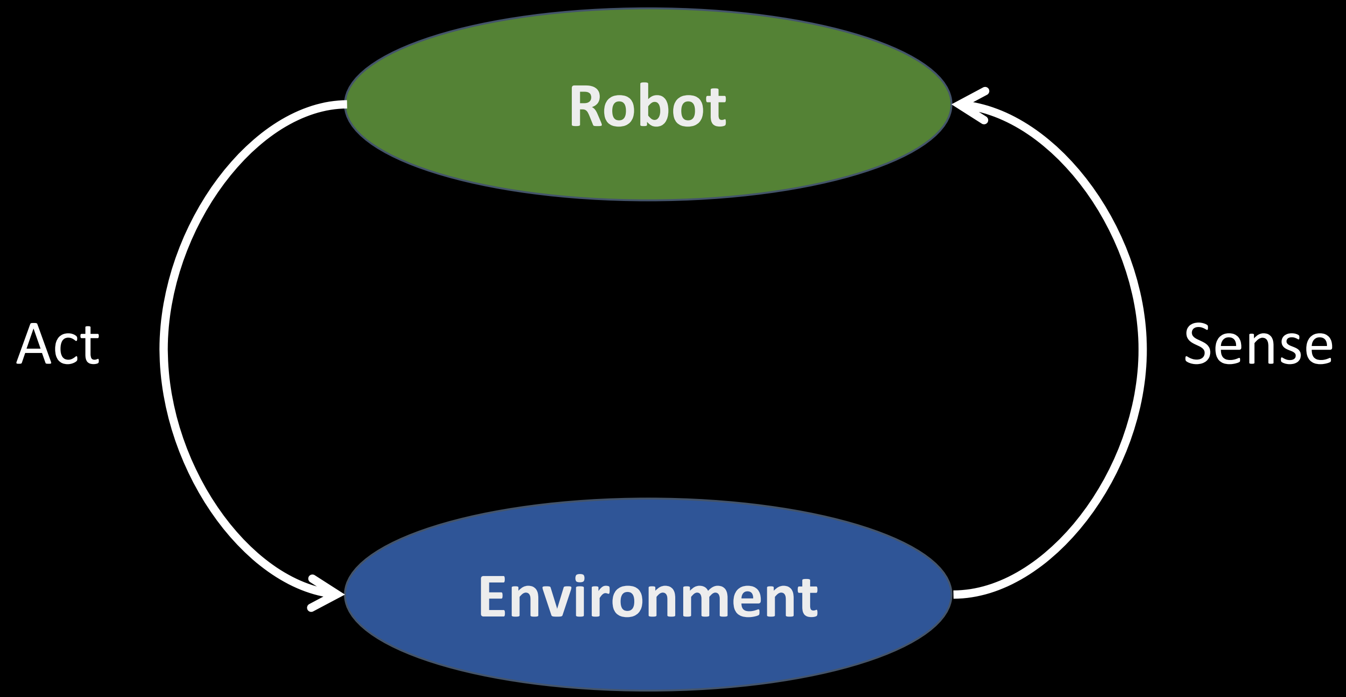
- Two ways in which the photograph of *The dress* may be perceived:
  - blue and black under a yellow-tinted illumination (left figure) or
  - white and gold under a blue-tinted illumination (right figure)



Probability for Robotics
THE HOW

Robot-Environment Model

Robot-Environment Interaction



- Two fundamental types of interaction between a robot and its environment:
  - Sensor Measurements/Observations
  - Control Actions

Robot-Environment Interaction

- The environment is a dynamical system that possesses its own internal state
- The robot can acquire information about its environment using its sensors
  - But sensors are noisy and cannot sense all aspects of the environment directly
- The robot maintains an internal belief with regards to the actual state of the environment
- The robot can influence its environment through its actions,
  - But actions are unpredictable
- Each control state affects the environment state and the robot's internal belief regards to this state

Robot-Environment Model

- The model helps us express a robot-environment interaction using probability
- Robot environment interactions are typically modeled as a discrete time system
- The state at time t will be denoted by as x_t
- A sensor measurement at time t will be denoted as z_t
- A control action will be denoted by u_t, which carries information about the change of the robot state in the time interval $(t\text{-}1{:}t]$ i.e. u_t indices a transition from state x_{t-1} to x_t

Conventions as per Siegwart, R., Nourbakhsh, I.R. and Scaramuzza, D., 2011. Introduction to autonomous mobile robots. MIT press.

Robot-Environment Model

- For no particular reason, we assume the robot executes a control action $\mathrm{u_t}$ first and then takes a measurement $\mathrm{z_t}$
- There is only one control action per time step t, and include as legal action "*do-nothing*"
- There is only one measurement per time step t
- Shorthand Notation: $\mathrm{x_{t1:t2} = x_{t1}, x_{t1+1}, x_{t1+2}, \ldots, x_{t2}}$

Robot State

- The state of the robot is the collection of all aspects of the robot and its environment that can impact the future
- The state includes:
  - Robot Specific:
    - Pose, Velocity, Sensor status(whether they are working or not), etc.
  - Environment Specific:
    - Variables that are static, such as the location of walls in (most) buildings **(static state)**
    - Variable that are dynamic, such as the whereabouts of people in the vicinity of the robot **(dynamic state)**
- The robot state will be denoted as x, although the specific variables included in x is context-specific
- The state at time t will be denoted by as x_t

Robot State (Examples)

Some typical state variables used:

- Robot pose, for a mobile robot, which comprises of its location and orientation relative to a global coordinate system
- Robot configuration, for a robot manipulator, which comprises of the configuration of the robot's actuators
- Robot velocity and velocities of its joints
- Location and features of surrounding objects in the environment: An object may be a tree, a wall, or a pixel within a larger surface. Features of such objects may be their visual appearance such as color, texture , etc.

Sensor Measurements/Observations

- The robot perceives its environment through sensors and the result of a perceptual interaction will be termed as measurement/observation
- A measurement at time t will be denoted as z_t
- Provides information about the environment's state, and hence tends to increase the robot's knowledge

Control Actions

- They change the state of the world.
  - Ex: robot motion, manipulation, etc.
- A control action will be denoted by u_t , which carries information about the change of the robot state in the time interval $(t\text{-}1{:}t]$
- Tends to induce a loss of knowledge due to inherent noise in robot actuation and stochasticity of robot environments

Probabilistic Generative Laws

- The evolution of state and measurements is governed by probabilistic laws
- We are interested in generative laws concerning the evolution of the:
  - State: How is x_t generated stochastically?
  - Measurements: How is z_t generated stochastically?

State Generation

- x_t is generated stochastically from the state x_{t-1}
- x_t depends on $x_{0:t-1}$, $z_{1:t-1}$ and $u_{1:t}$

$$p(x_t \mid x_{0:t-1}, z_{1:t-1}, u_{1:t})$$

Markov Assumption

Markov Assumption

The Markov assumption postulates that past and future data are independent if one knows the current state

- It is assumed that future states depend only on the current state, not on the events that occurred before it
- A stochastic model/process that obeys the Markov assumption is a Markov model
- Plays a fundamental role in probabilistic robotics

Andrey Markov (1856–1922) was a Russian mathematician best known for his work on stochastic processes

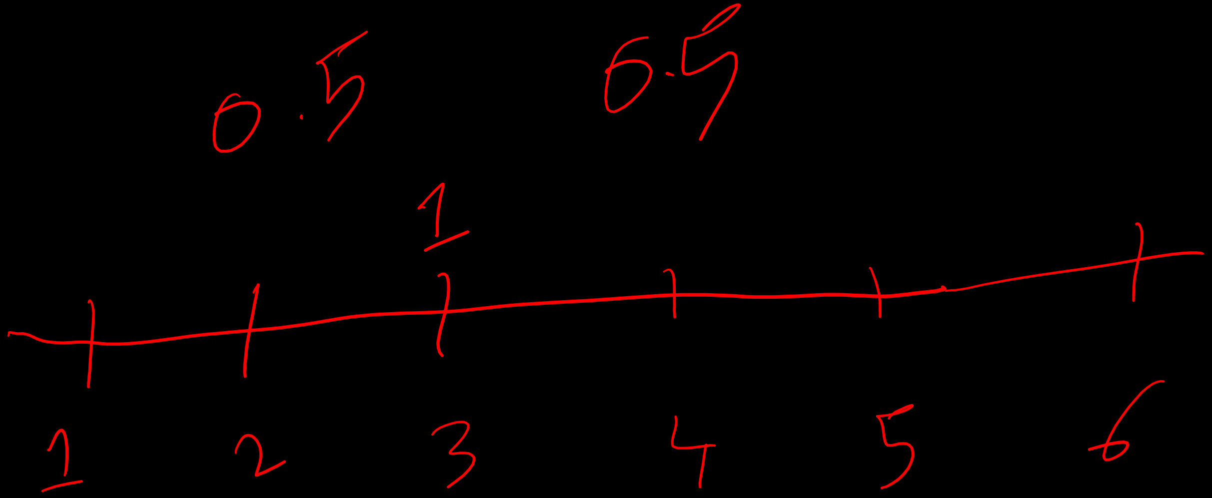
Markov Assumption

The Markov assumption postulates that past and future data are independent if one knows the current state

- The knowledge of past states, measurements or controls carry no additional information that would help us predict the future more accurately
- <u>This does not mean the future is a deterministic function of the current state</u>
- Hence, if we somehow model the state using the Markov assumption, we can recursively estimate the robot state x_t from the previous state x_{t-1} , measurement data z_t and control input u_t

Markov Process 1

- A famous Markov chain is the so-called "drunkard's walk“
- A random walk on the number line where, at each step, the position may change by +1 or −1 with equal probability
- From any position there are two possible transitions, to the next or previous integer
- The transition probabilities depend only on the current position, not on the manner in which the position was reached.



Markov Process 1

- A famous Markov chain is the so-called "drunkard's walk“
- A random walk on the number line where, at each step, the position may change by +1 or −1 with equal probability
- From any position there are two possible transitions, to the next or previous integer
- The transition probabilities depend only on the current position, not on the manner in which the position was reached.
- For example, the transition probabilities from 5 to 4 and 5 to 6 are both 0.5, and all other transition probabilities from 5 are 0.
- These probabilities are independent of whether the system was previously in 4 or 6.

Markov Process 2

- A coin purse contains five quarters (each worth 25¢), five dimes (each worth 10¢), and five nickels (each worth 5¢). Coins are randomly drawn, one by one, from the purse and are set on a table.
- If X_n represents the total value of the coins set on the table after n draws, then the sequence $\{ X_n : n \epsilon \mathbb{N} \}$ represents a stochastic process
- Suppose that in the first six draws, all 5 nickels and 1 quarter are drawn.

X_6 : \$0.50

DISCUSSION

1. Can you say, with a probability of 1, what the value of X_7 will be?
   a) For example, I can say $P(X_7 \geq 0.55) = 1$. Can you do better i.e with a higher value?
   b) Can you draw a nickel in the 7th draw?
2. Is this is a Markov process?
3. If not, how can you model this scenario into a Markov process by somehow changing the definition of X_n ?

Markov Process 2

- A coin purse contains five quarters (each worth 25¢), five dimes (each worth 10¢), and five nickels (each worth 5¢). Coins are randomly drawn , one by one, from the purse and are set on a table.
- If X_n represents the total value of the coins set on the table after n draws, then the sequence $\{ X_n : n \epsilon \mathbb{N} \}$ is not a **Markov process**.
- Suppose that in the first six draws, all 5 nickels and 1 quarter are drawn.

$$X_6 : \$0.50$$

- If we know not just X_6, but the earlier values as well, then we can determine which coins have been drawn, and we know that the next coin will not be a nickel; so we can determine that $X_7 \geq 0.60$ with probability 1
- But if we do not know the earlier values, then based only on the value X_6 we might guess that we had drawn four dimes and two nickels, in which case it would certainly be possible to draw another nickel next. Thus, our guesses about X_7 are impacted by our knowledge of values prior to X_6.

Markov Process 2

- How can I model this scenario into a Markov process by somehow changing the definition of X_n?
- We could define X_n to represent the count of the various coin types on the table.
- For instance, $X_n = [1, 0, 5]$ could be defined to represent the state where there is one quarter, zero dimes, and five nickels on the table after 6 one-by-one draws.
- Suppose that the first draw results in state $X_1 = [0, 1, 0]$. The probability of achieving X_{2} now depends on X_1; for example, the state $X_2 = [1, 0, 1]$ is not possible.
- After the second draw, the third draw depends on which coins have so far been drawn, but no longer only on the coins that were drawn for the first state
- In this way, the likelihood of the state X_n depends exclusively on the outcome of X_{n-1}
- This new model would be represented by 216 possible states (6x6x6 states, since each of the three coin types could have zero to five coins on the table by the end of the 6 draws)

Robot-Environment Model

State Generative Model

- x_t is generated stochastically from the state x_{t-1}
- x_t depends on $x_{0:t-1}$, $z_{1:t-1}$ and $u_{1:t}$

$$p(x_t \mid x_{0:t-1}, z_{1:t-1}, u_{1:t})$$

- If state x_t is modeled under the **Markov Assumption**, then

$$p(x_t \mid x_{0:t-1}, z_{1:t-1}, u_{1:t}) = p(x_t \mid x_{t-1}, u_t)$$

(conditional independence)

Knowledge of only the previous state x_{t-1} and control u_t is sufficient to predict x_t

Measurement Generative Model

- Similarly, the process by which measurements are generated are of importance

$$p(z_t|x_{0:t}, z_{1:t-1}, u_{1:t})$$

- If x_t conforms to the **Markov Assumption**, then

$$p(z_t|x_{0:t}, z_{1:t-1}, u_{1:t}) = p(z_t|x_t)$$

(conditional independence)

- The state x_t is sufficient to predict the (potentially noisy) measurements

Knowledge of any other variable, such as past measurements, controls, or even past states, is irrelevant under the Markov Assumption

Robot-Environment Model

+

Markov Assumption

+

Bayes Theorem

=

Bayes Filter

References

1. Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. Probabilistic robotics. MIT press, 2005.
2. http://gki.informatik.uni-freiburg.de/teaching/ws0607/advanced/recordings/BayesFiltering.pdf
3. http://home.deib.polimi.it/restelli/MyWebSite/pdf/E05_h.pdf
4. https://en.wikipedia.org/wiki/Markov_chain
5. https://en.wikipedia.org/wiki/Markov_model
6. Image Sources:
   a. https://blog.robotiq.com/a-brief-history-of-robots-in-manufacturing
   b. https://northwesturbanist.files.wordpress.com/2015/01/20151104_082523.jpg
   c. https://expresswriters.com/is-the-dress-blue-black-white-or-gold-how-it-went-viral/
   d. https://en.wikipedia.org/wiki/Andrey_Markov