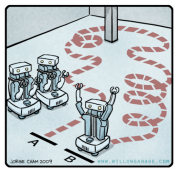
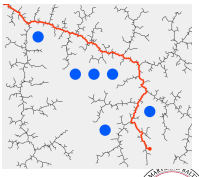
Cognition, Planning, HRI, Ethics, Questions

Bookkeeping

2

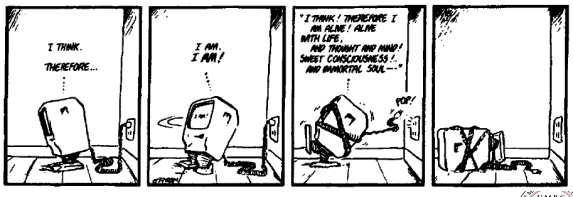
- Project milestones out
  - The evaluation section may be completed *after* the 15th
  - In that case, please turn in the whole document
- HW 4 out (finally)
  - Adjustments:
    - Due much later
    - Can be done in group
    - Hopefully, this will generate study sessions
  - *However*, come talk to me if the timing is impossible

Schedules

3

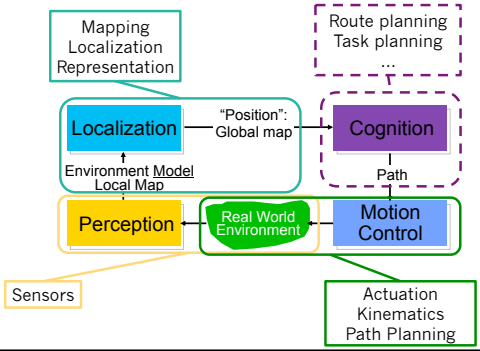
- This class: Planning; cognition and control; HRI; ethics; any other remaining topics you have questions about
- Next class: bring robots, computers, mazes (?), etc.
  - You'll have grades to ask questions about
- Dates from here on out
  - May 14: Project code, video, and writeup due
  - May 15: Robots in mazes (ENGR atrium)
  - May 17: Optional updated writeup due
  - May 20: HW 4 due
  - May 22: Final exam

Cognition



Cognition and Mobility

5



What Is Autonomy?

6

- Autonomous robots…
  - Perform their tasks in the world by themselves
  - Do not require human control / intervention
  - Learn about their environment and tasks
  - Avoid damage (to themselves, people, property)
  - Adapt to changing situations
  - Make and execute decisions
  - Possess some degree of self-sufficiency
- Intelligently and safely perform tasks
- Without direct human control

All of these?

A More Functional Definition

8

- Intelligently, self-sufficiently, and safely **perform tasks**
- Without human control / intervention
- Learn about environment and tasks
- Adapt to changing situations
- Make and execute decisions

How?

Artificial Intelligenc

9

- Key types
  - Strong AI: mental/thought capabilities equal to (or better than) human
  - Weak (bounded) AI: intelligent actions or reasoning in some limited situations
- These are problematic
  - How do we measure it?
  - What's an 'intelligent action'?
    - In practice, 'previously human only'
  - Is there something ineffable missing?

- How does it change when it's a robot?

Intelligent Action…

11

- Physical tasks
  - Or physical-related
    - Sensing
    - Path planning
- Understanding / modeling:
  - The **robot**
  - The **environment**
  - The **task**
- Autonomy
- Subtasks:
  - Knowledge Representation
  - Search
  - Planning
  - Learning
  - Inference
  - ~~Coping with uncertainty~~

Intelligent Action Needs…

12

- Knowledge Representation
- Search
- Planning
- Learning
- Inference

DARPA Grand Challenge 1

13

DGC 1 Challenges

14

- Localization
  - But not mapping
- Sensor management
  - What sensors?
  - Where's the road?
- Narrow pass
  - Switchbacks, turns
  - Tunnels
- Actuator management

✓ Knowledge Representation
✓ Search
✓ Planning
✗ Learning
✗ Inference

DARPA Grand Challenge 2

15

DGC 2 Challenges

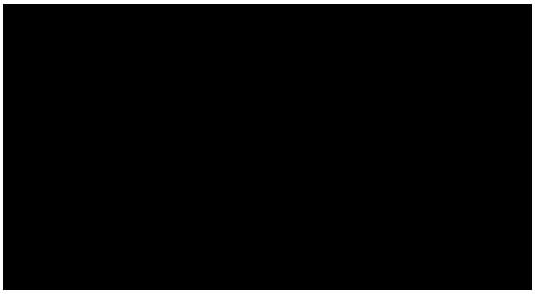
16

- All of the above, plus…
- Visual parsing of traffic elements
  - Sensing + Knowledge
- Awareness of other cars
  - Sensing + Planning
- Non-3D-guided tasks
- Faster speeds
- Safety

✓ Knowledge Representation
✓ Search
✓ Planning
✗ Learning
✓ Inference

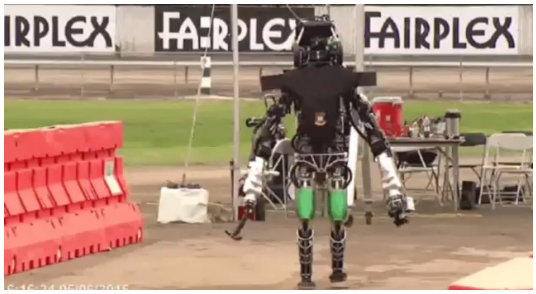
DARPA Robot Challenge 1

17



DRC 1: Robots Falling Over

18



DRC 1 Challenges

19

- All of the above, plus…
- Non-designed environment
  - Balancing
  - (Much) harder actuation
  - Bipedal motion
- Yet more sensor hassles
- Weight, power
- Manipulation
  - !!!

✓ Knowledge Representation
✓ Search
✓ Planning
✓ Learning
✓ Inference
✗ Autonomy

Self-Driving Cars

22

Car Challenges

23

- All of the above, plus…
- Sublimely oblivious drivers
- Full-speed actuator management
- Legal management
- Ethical management
- Human-robot Interaction

- ✓ Knowledge Representation
- ✓ Search
- ✓ Planning
- ? Learning
- ✓ Inference
- ✓ Autonomy

Other Autonomous Tasks

Intelligent Action…

25

- Physical tasks
  - Or physical-related
    - Sensing
    - Path planning
- Understanding / modeling:
  - The **robot**
  - The **environment**
  - The **task**

- Autonomy
- Subtasks:
  - Knowledge Representation
  - Search
  - Planning
  - Learning
  - Inference
  - ~~Coping with uncertainty~~

Robot Gangnam Style

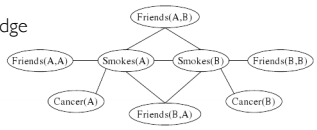
26

UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Knowledge Representation

27

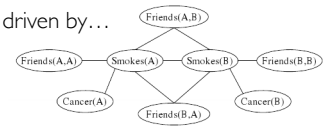
- What does a robot need to know?
- What would it be useful for a robot to know?
- Types of knowledge
  - Background Knowledge
  - Task-Specific Knowledge
  - Explicit vs. Implicit Knowledge
- Representation Choices
  - Probabilistic?
  - Human-understandable?



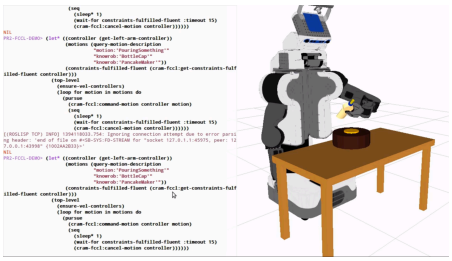
KR: Approaches

28

- Hand-coded knowledge
  - E.g., maps, object recognition, task descriptions, …
- Machine learning
  - Beforehand
  - On-the-Fly
- Representation choices driven by…
  - Efficiency
  - Requirements
  - Our limited abilities



Planning

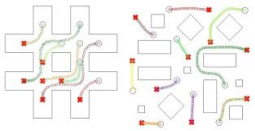


Knowledge-Based Specifications of Robot Motion. Tenorth, Bartels, Beetz, 2014.

Planning

30

- What does a robot need to plan?
- Motion
  - Mobility: where am I going?
  - Manipulation: how do I move myself?
- Tasks
  - Low-level: pick up this piece
  - High-level: win this chess game
- Steps to Goals
- Step Ordering



Planning: Approaches

31

- Explicit plan space
  - Rule-based
  - Probabilistic
  - Ordered (scripted)
- Implicit plan space
  - Learned task performance
    - Learning from demonstration
    - Reinforcement learning
  - Procedural planning
- Non-planned (rigid) behavior

Path and Motion Planning

32

- Path planning: Choosing a route in 2D space
  - Find "shortest" (actually, lowest cost) path between vertices through a connected graph
  - Cost can mean lots of things
    - Minimal backtracking, minimum distance, least movement through dangerous areas, …
- Motion planning: Figuring out a set of movements to accomplish a pose in 3D space
  - Find lowest-cost way to get a manipulator, drone, or other complex device into a specified state
- Planning is done in **configuration space**

Path Planning: How?

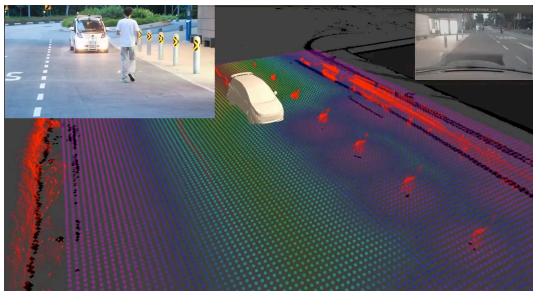
33

- Dijkstra's algorithm or A*
- D* search: Re-plan around obstacles, keep most of plan
- Tree-building algorithms
  - Rapidly exploring random trees (RRT):
  - Randomly build a tree that fills ("grows through") a space

http://correll.cs.colorado.edu/?p=965
https://www.youtube.com/watch?v=rPgZyq15Z-Q

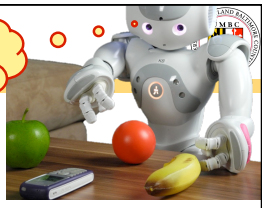
Self-Driving Cars

34



Learning

35

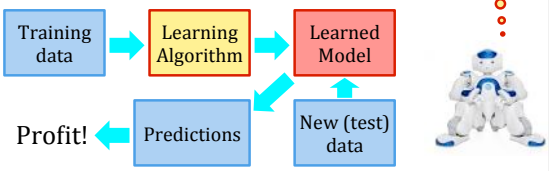


- Why do learning?
  - Hard to program tasks
  - More effective performance
  - Flexibility and autonomy
- What can be learned?
  - Previously unknown environment, objects, etc.
  - Previously unknown tasks
  - Background knowledge
- Machine Learning Approaches

Learning: Approaches

36

- Machine Learning
  - Approaches: Learning from demonstration, reinforcement learning, real-time search, statistical model-building, feature extraction, ensemble learning, active learning, lifelong learning, reading-based learning, learning to read, …
- Fundamental concept: data-driven learning



POMDPs

37

- Partially Observable Markov Decision Process
- A Markovian model for choosing next action when **result of actions** are uncertain
  - Must choose best action: What are immediate results? Long-term results? Under uncertainty?
- Robot cannot directly observe underlying state
  - But **some** MDP still describes world
- Maintain a probability distribution over **set of possible states**, based on observations, observation probabilities, and underlying world model (MDP)

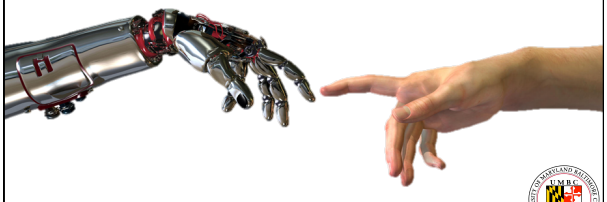
en.wikipedia.org/wiki/Partially_observable_Markov_decision_process

Search

39

- Knowledge Representation
  - Which bit of knowledge?
- Planning
  - What rules to apply?
  - Of many steps / paths / subgoals, which is best? In what order?
  - What is the goal?
- Inference
  - What rules to apply?
  - What form to apply?
  - Truth maintenance
- Learning
  - Usually NP-complete
  - Algorithms and learning methods

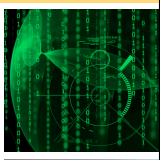
Human-Robot Interaction: HRI



Human-Robot Interaction

41

- What is an interaction with a robot?
  - What is a robot?
    - (We talked about this one a lot)
  - What counts as interaction?
- Robot(ic)s, for our purposes, is where computation meets the physical world
- Any real-world human-robot interaction
  - Physical interactions with humans
    - Speaking, hearing, gaze contact, holding hands…
  - Taking commans (spoken, typed…)



Some High-Level Topics

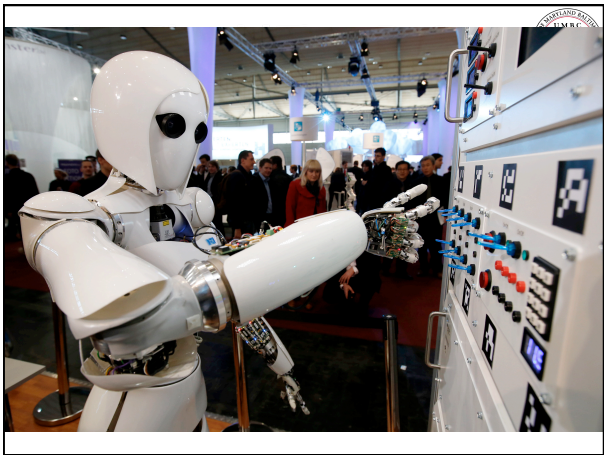
42

- Where will we begin seeing robots?
- What is a "Human-centric environment"? Task?"
  - And how is it relevant to robots?
- Social robots
  - How can robots be social beings? When do we want them to?
  - How can robots express emotion – and (when) should they?
- Human-robot collaboration – how and when?
- Assistive robots – helping people with special needs
- Robot ethics – what should and shouldn't robots do?

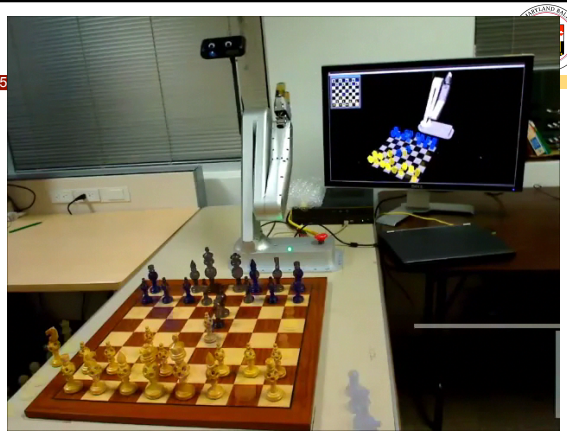
HRI Includes…

43

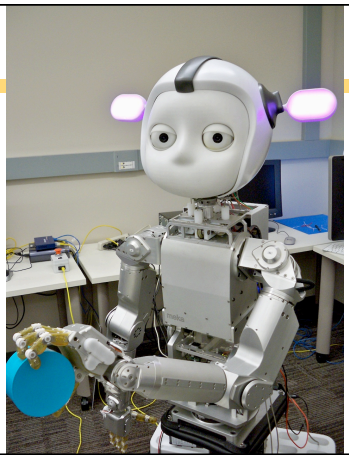
- HRI taxonomy and metrics
- HRI methods
- Assistive robots
- Socially-assistive robots
- Human-robot collaboration
- Collaborative manipulation, human-robot handover
- Natural-language interactions with robots
- Remote tele-operation
- Expressive robot motion
- Navigation around humans
- Perception of humans
- Social learning
- Non-verbal communication: Gestures
- Human-robot dialog
- Remote presence



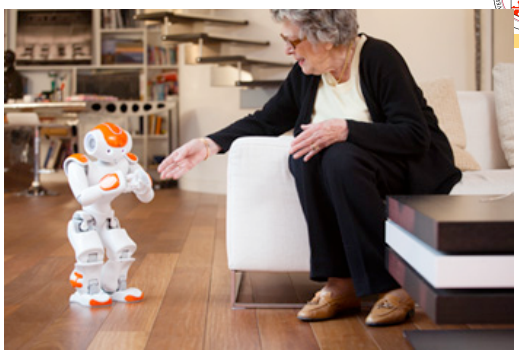
45

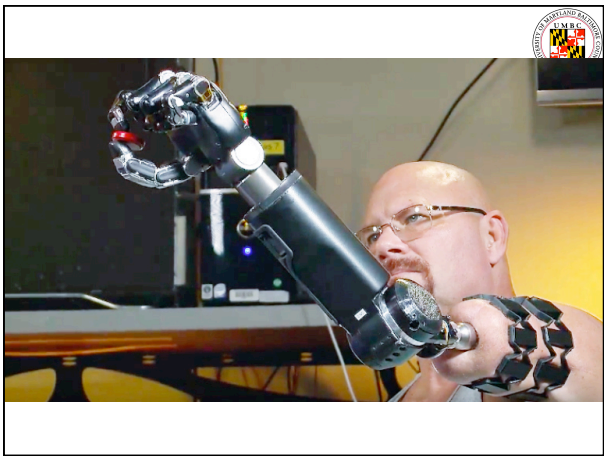


46



47





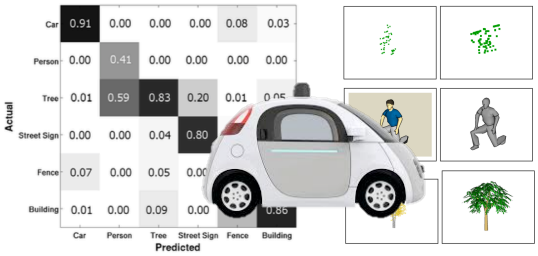
Ethics and Robotics

50

Robot Vision

54

- You're building a vision system:



Cybernetics

55

Categories of Harm

56

- Robots can physically harm people.
  - How?
  - Negligently
  - Deliberately
  - ...?

  This is a huge category – but is it the most likely?
- Robots can impinge on people's self-determination.
- Robots can fail to act as expected, causing…
  - Physical harm
  - Emotional distress
  - Disappointment

"victimless" ethics is rarely "applied"

Categories of Harm

57

- Robots can change our definition of "humanity."
- Robots can have rights that are impinged upon.
- Robots can discriminate.
- Robots can do environmental damage.
- Robots can increase the have/have-not gap.
  - For people
  - For nations

What Robots, and How?

58

- Military Robots
- Caretaker robots
  - Elderly
  - Children
- Assistive robots
- Exploration robots
- Factory robots
- Surgical robots
- Cybernetics
- Search-and-rescue
- Automated cars
- Bipeds and Quadrupeds
- Why is ATLAS Scary?

Big Questions

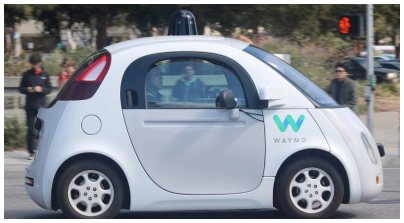
59

- Can computers "hurt" people? **Sure.**
- What about robots? **Even more so.**
- Can a machine be "unfair" (discriminatory)? **Sort of. There's a GIGO aspect.**
- Why do we, as roboticists, care? **Ethics and morals, legal liability**
- What are some immediate issues, right now?

Topics

60

- Drive the discussion an example: Self-driving cars



- And generalize from there

Self-Driving Cars

61

- Cars can hurt or kill people.
  - How many fatalities is acceptable?
  - Is it enough to not cause accidents?
- People cause accidents!
  - ~38,000 deaths per year in the U.S.
  - Lately it's been going up
  - How many of you text and drive?
- Do cars have to be perfect? Just better than humans? Somewhere in between?

Harder Questions

62

- What about naked self-driving cars?
  - No control mechanisms inside at all
- Should it be legal for a person to drive?
  - Even if cars are demonstrably better at it?
- Why?
  - Because I wanna?
  - Because we dislike giving up control?
- Even if you accept the risks, what about my rights?
- Who's legally liability? ← this is a big question that will affect the future

The Hardest One

63

- When an accident is inevitable…
  - Should the car occupants get hurt?
  - That is, the person who paid for it?
  - If it's not their fault?

- Would you buy a car that could hurt or kill you?
  - If it could be avoided by hurting or killing someone else?
- But consider:
  - Would you swerve to avoid a kid in the road?
  - What about a baby stroller?
- Who should be deciding these things? Uber?

← Correct answer: "oh no no no no"

Meta-Questions

64

- Questions we will **not** answer today:
  - What do "right" and "wrong" mean?
  - Who gets to decide what's right and wrong?
  - How do/should those decisions be made?
  - What should we do about things that are wrong?
- We'll use commonly understood ideas of wrong:
  - It's wrong to **harm** people
    - Physically, emotionally, financially…
  - It's wrong to **discriminate** against people
  - It's wrong to **steal** from people
  - It's wrong to invade people's **privacy**
  - It's wrong to be **unfair** to people

"Without extenuating circumstances," and understanding that sometimes there's no "right" alternative