Learning Robot Soccer from Demonstration

Dan Grollman
CS 148 Guest Lecture

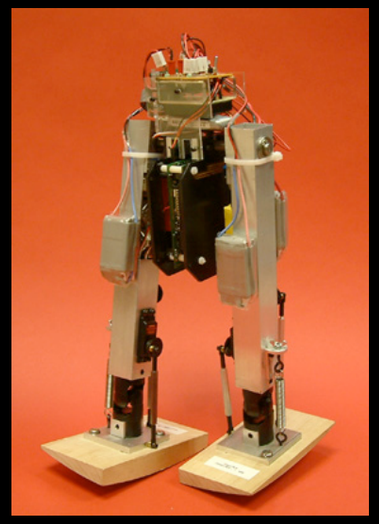
Robotics is awesome

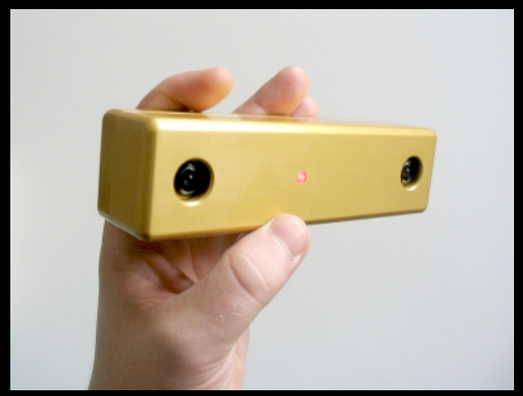
- Computer Science
  - Programming
- Electrical Engineering
  - Circuits
- Mechanical Engineering
  - Physical bodies
- Cognitive Science
  - How they think
- Psychology
  - Human interaction
- Biology
  - Inspiration
- Physics
  - Sensors
- Chemistry
  - Eating slugs

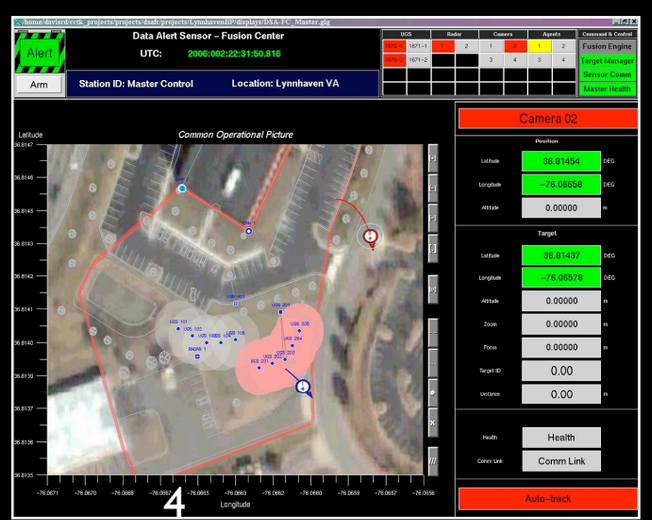
Slugs = Power

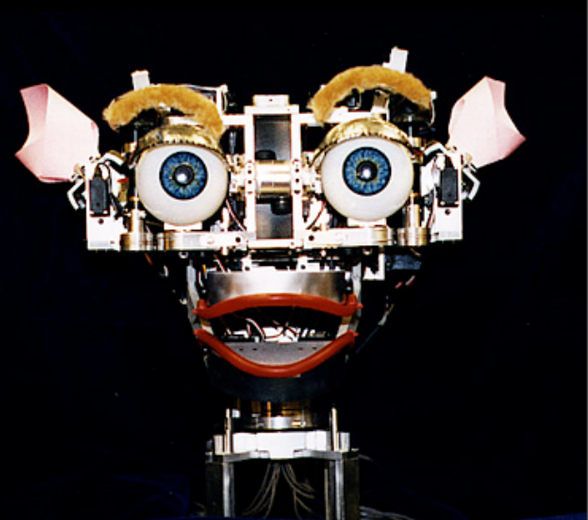
Things to do in Robotics

- Make Robots move
  - Locomotion
  - Grasping, manipulation
- Make Robots sense
  - Camera vision
  - Laser scanners
  - Touch sensors
- Human-Robot Interaction
  - Interface design
  - Socialability









What's interesting?

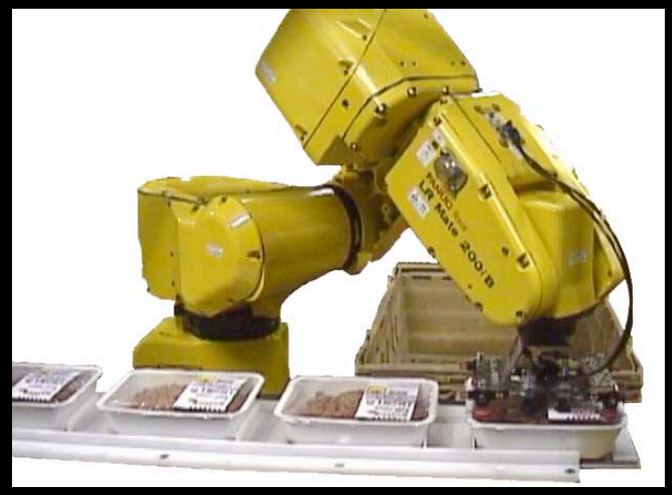
Doing stuff

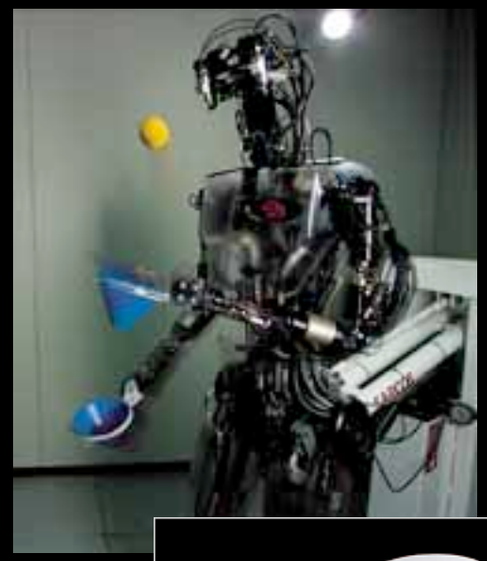
- Making a robot do a task
  - Hard - Not everyone is a programmer
    - Many programming approaches
      - What matters is doing it!
- What would you want a robot to do?

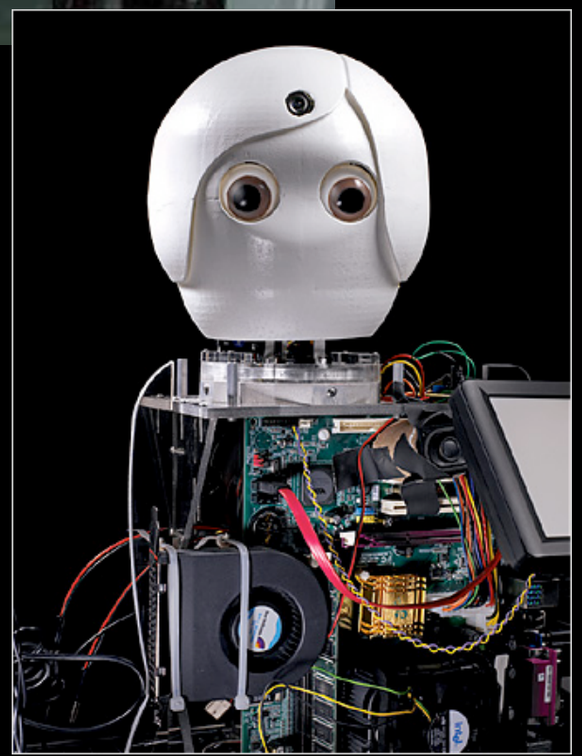
6

Things robots can do

- Pick up stuff



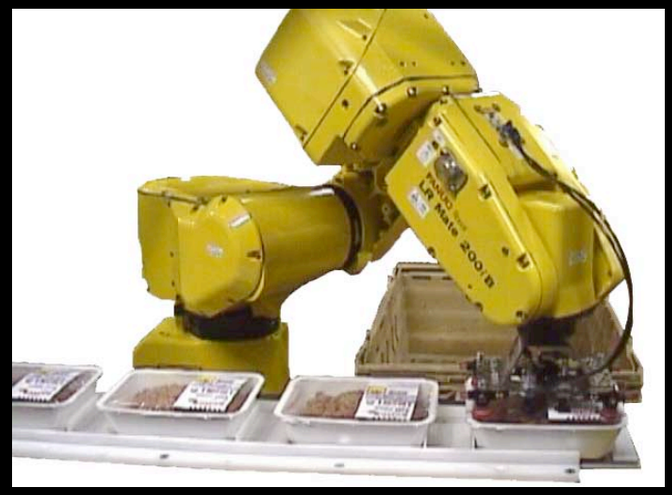


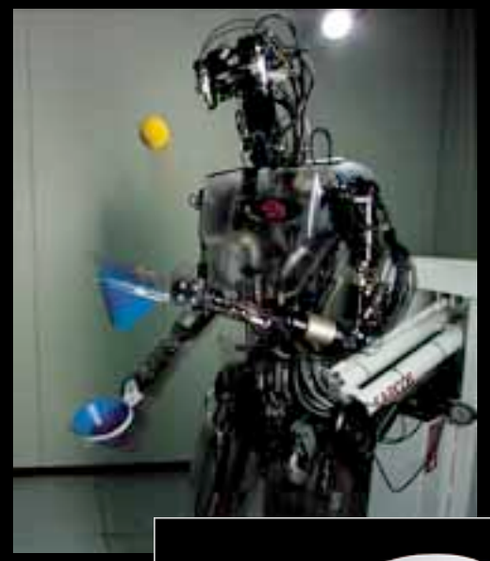


- juggle
- walk around
- help you lose weight

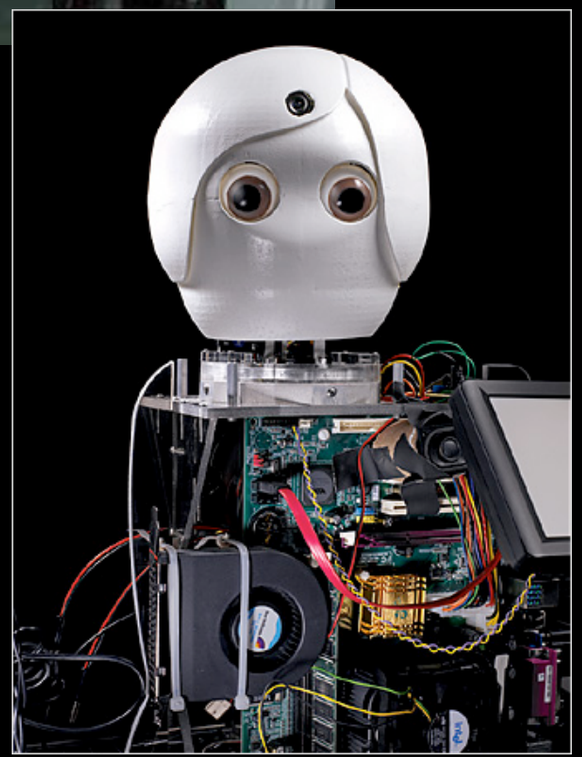
7

Things robots can do





- Pick up stuff
- juggle
- walk around
- help you lose weight



Each of these is single-task. And took lots of coding.

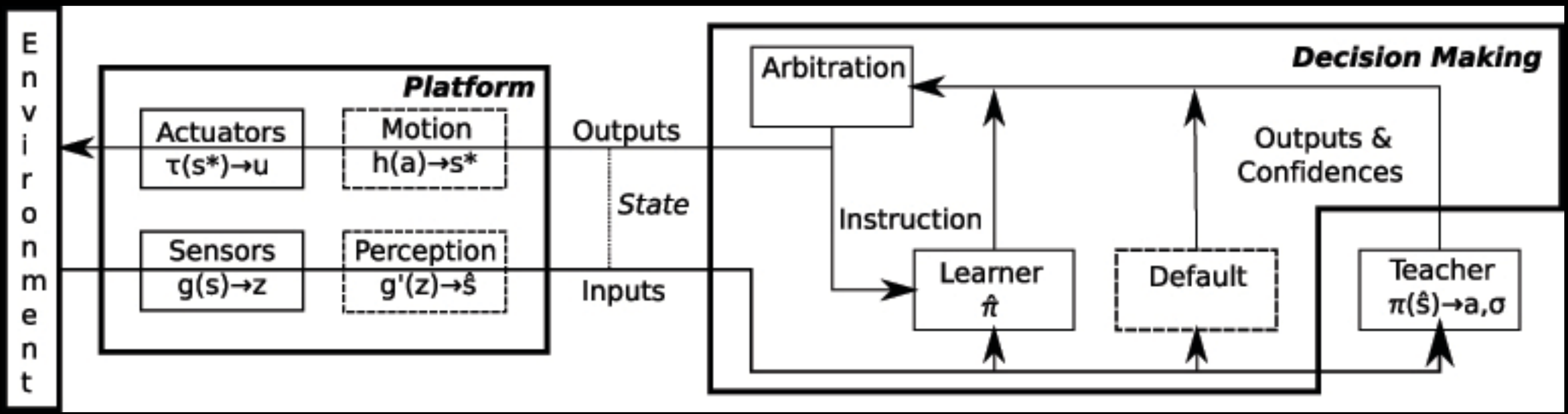
7

Learning from Demonstration

- If you can do it, you can teach it
- Simply show the robot, and it learns
- No need to learn to program
- Wider user base

8

For Unknown tasks via Mixed-Initiative Control

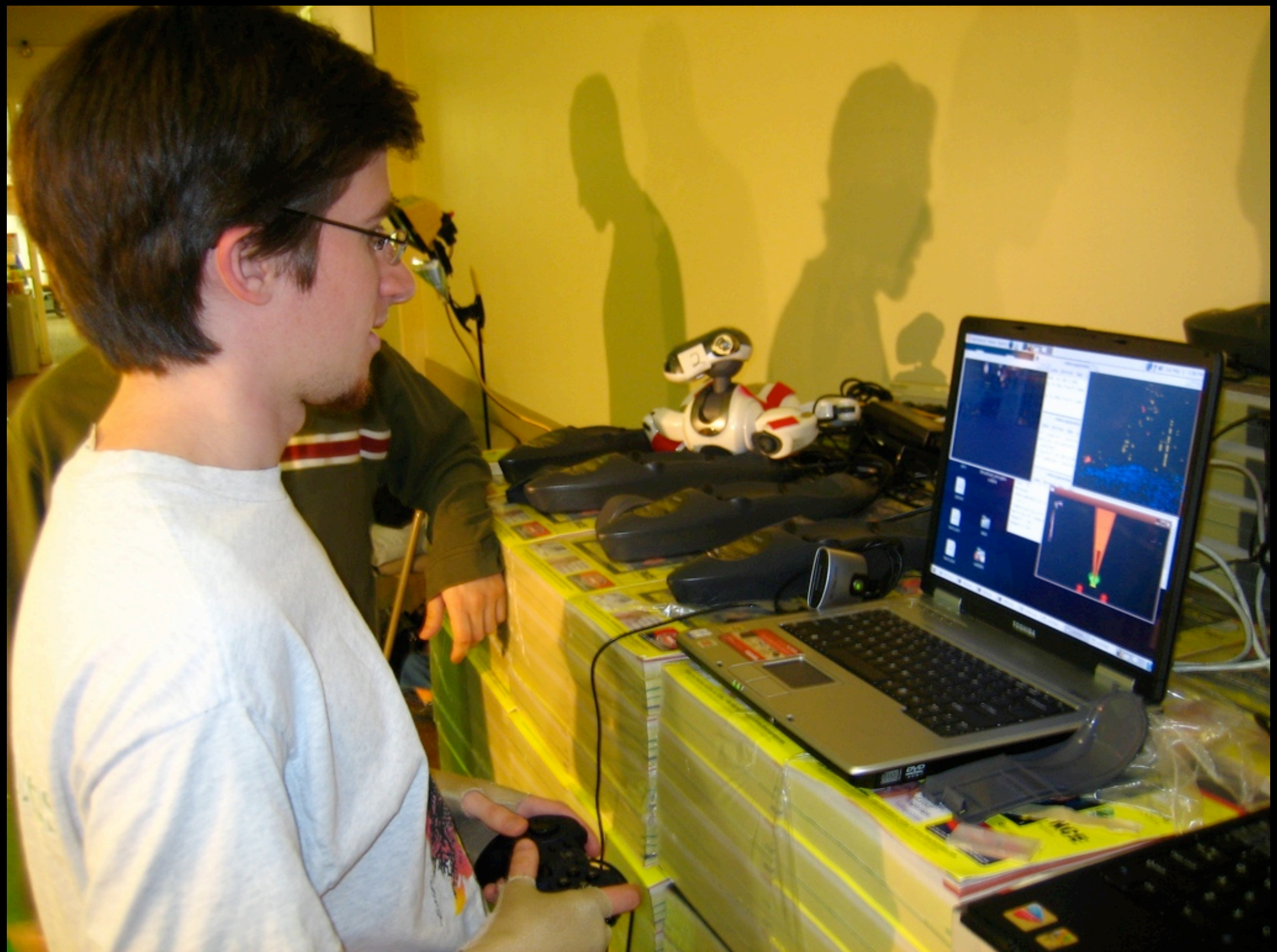


9

How Show?

- Robot watches a human
  - Body mapping problem
  - Viewpoint matching problem
- Human teleops robot
  - Task is proven robot-performable
- Human uses robot sensor data
  - Task is proven robot-detectable

10



| |

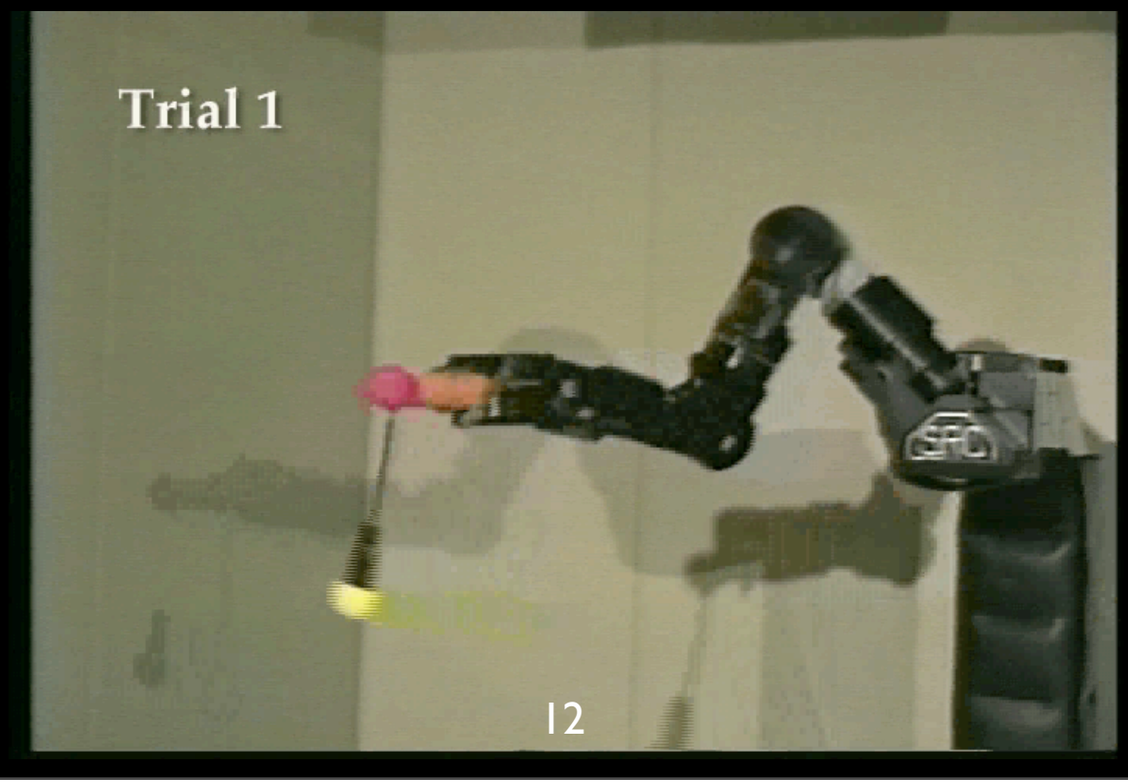
Critiquing

- Robot rarely gets it right the first time
- Known task
  - Can use state exploration and value function to improve (RL)

12

Critiquing

- Robot rarely gets it right the first time
- Known task
  - Can use state exploration and value function to improve (RL)



Unknown task

- Need more input from user
- One demonstration not enough
- Train-Observe-Train more
  - 2 phases
  - Interactive

13

Mixed-Initiative Control: Arbitration

- Both demonstrator and learner want control of robot.
- Who gets it?
    - CONFIDENCE!

14

How to learn?

- Have matched inputs and outputs
  - Robot's sensors (s) and robot's actuators (a)
  - Want to learn mapping, $\pi(s) \rightarrow a$
    - Can use regression

15

Regression

- Function approximation
- Functional $y = f(x) + \epsilon$
  - Assumes unimodal output

$$\epsilon \in N(0, \sigma^2)$$

  - Find f from data of from

$$\{X, Y\} = (x_1, y_1)...(x_n, y_n)$$

- Inference = discover the model
- Prediction = use model to predict output ($\hat{y}$) for new input (x_6')

Linear (LSQ)

- Assume f is of form y=ax+b
  - find a and b to minimize squared error
  - “Fitting a line to the data”

- Could fit other functions
  - Parametric regression

17

Nonparametrics

- Not "No-paramaters"
- Number of parameters is unfixed
  - Grows with data
  - infinite, countable
- Form of function is not predetermined
  - More flexible

18

K-NN

- $f(x') = \frac{1}{K} \sum_{k=1}^{K} y_k$
  - Ordered by distance from x'
- Parameters = All other points
- Hyperparameter = K, how many to count
- Given enough data, can approximate any function

19

Basis Functions

- $f(x') = \frac{\sum_{k=1}^{K} d(x_k, x') * y_k}{Z}$
  - Weighted by distance
  - Can let K= inf
  - Use a different distance function (kernel) to create ‘areas of influence’
    - RBF: $d = \exp(-0.5 * \frac{||x' - x_k||^2}{\sigma^2})$

20

Gaussian Process Regression

- Nonparametric regression with RBFs
  - 'Kernalized KNN'
- Can approximate any function, given enough data
- Gives a distribution over functions
  - KNN is point estimate

21

Time is of the essence

- Want to interact in ‘realtime’
  - ~30 hz
- Robot needs to observe new data, predict and update in < 33ms
- Nonparametric algorithms keep all the old data for comparative purposes

22

Sparsity

- Keep only a subset of the data of size P
- Limits storage
- Limits processing
- How to pick P?
  - realtime
- How to pick the points?

23

Incremental point picking

- Need to pick points as they arrive
- For each point:
  - Is this point necessary?
  - Yes = Keep
  - No = Toss
- No going back!

24

Sparse, Incremental regression algorithm

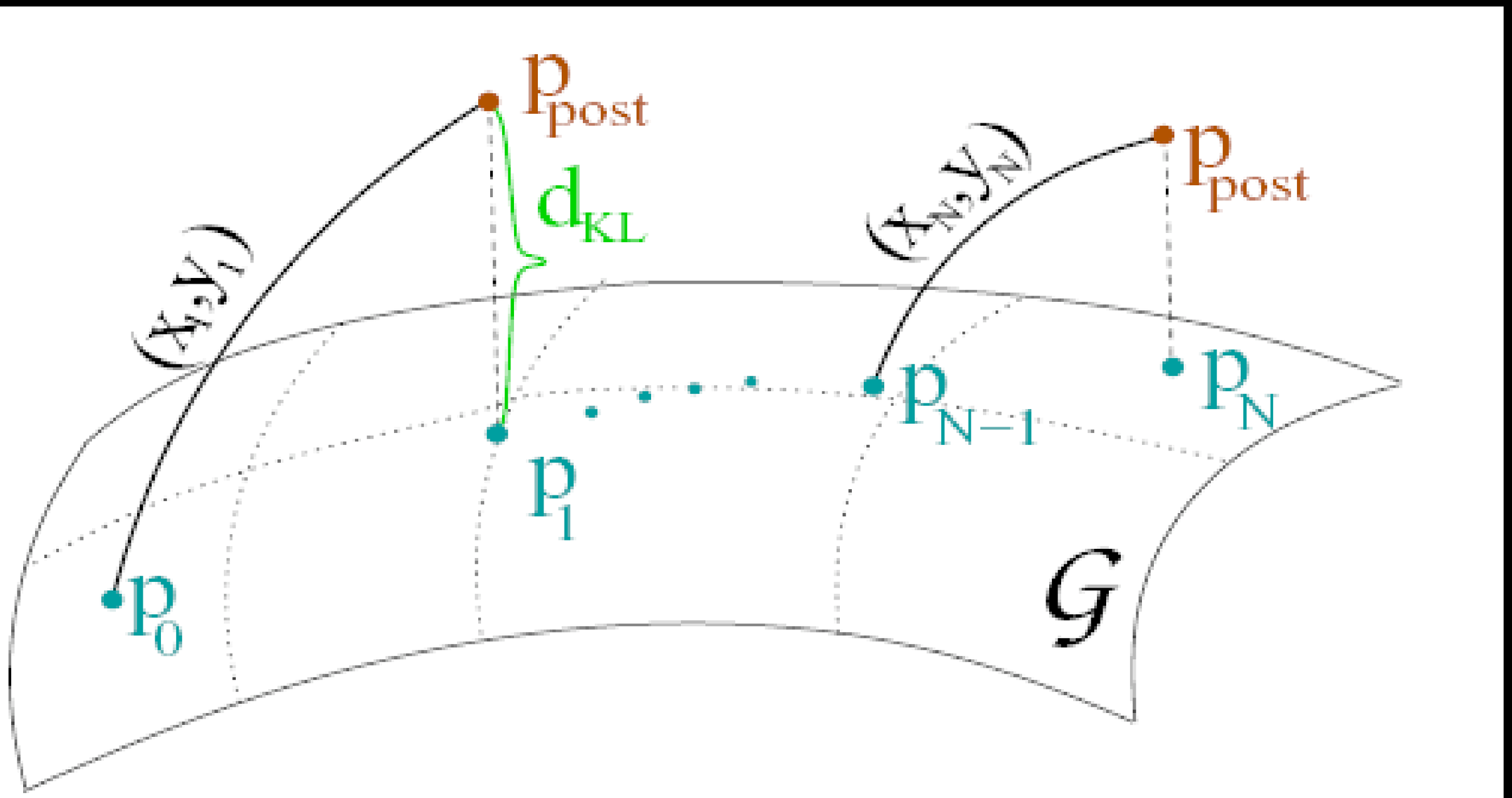
- One based on GPR: SOGP
- Another popular algorithm: LWPR

25

SOGP

- Like GPR, but keep only a subset of the points
- Points chosen by ‘novelty’ to minimize KL-divergence
    - The distance between distributions

26

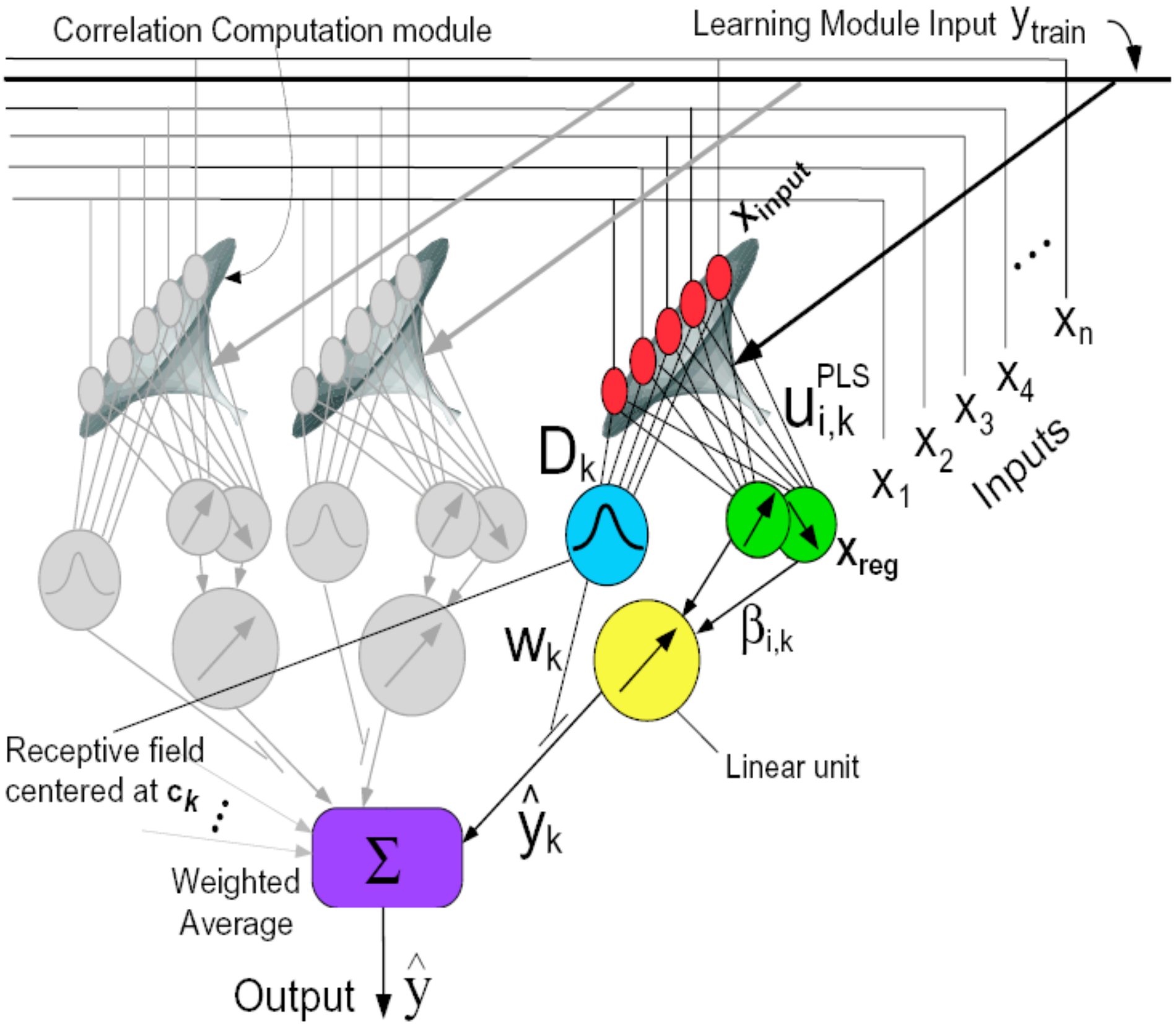


27

LWPR

- Defines receptive fields
  - Like areas of influence
- Each RF has a simple linear model
  - Fitting a line to the data
    - Actually, to a set of projections of the data

28


Correlation Computation module
Learning Module Input y_{train}
x_{input}
x_n
x_4
x_3
x_2
x_1
Inputs
$u_{i,k}^{PLS}$
D_k
x_{reg}
w_k
$\beta_{i,k}$
Linear unit
Receptive field
centered at c_k
$\hat{y}_k$
Σ
Weighted
Average
Output $\hat{y}$

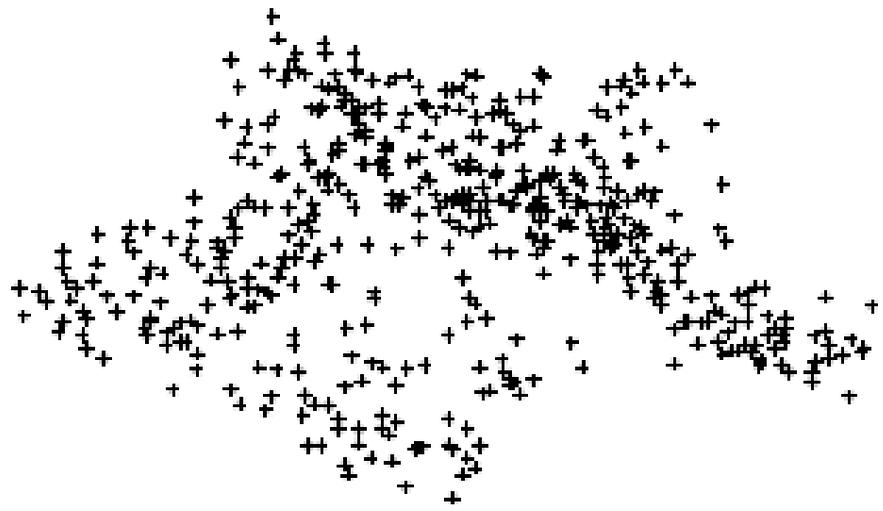
29

Comparison

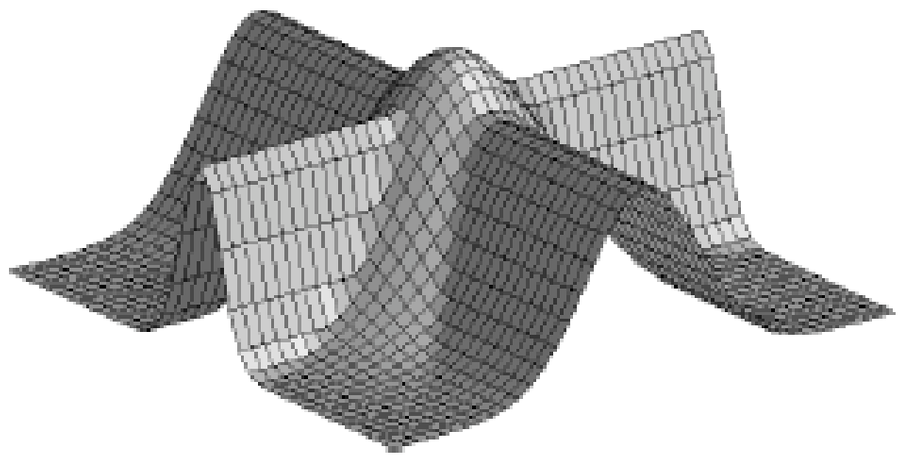
- Both are sparse
  - SOGP only stores up to a capacity
  - LWPR only stores RFs
- Both are incremental
  - No need to recalculate everything
- How do they regress?

30

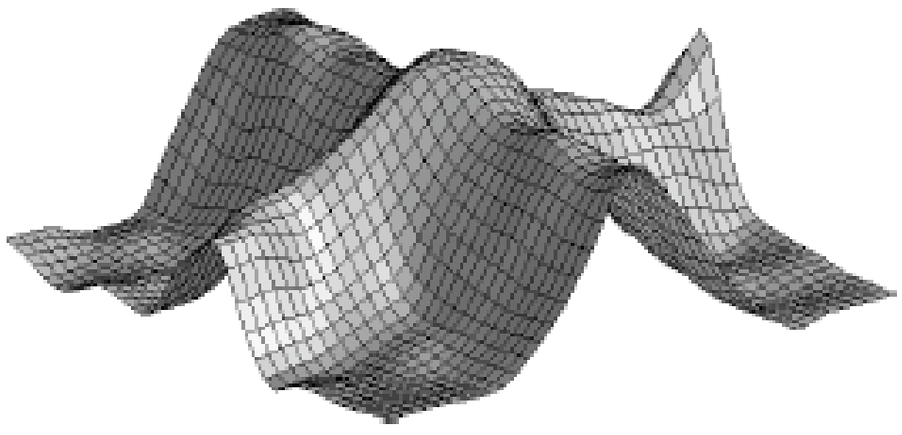
22 bases



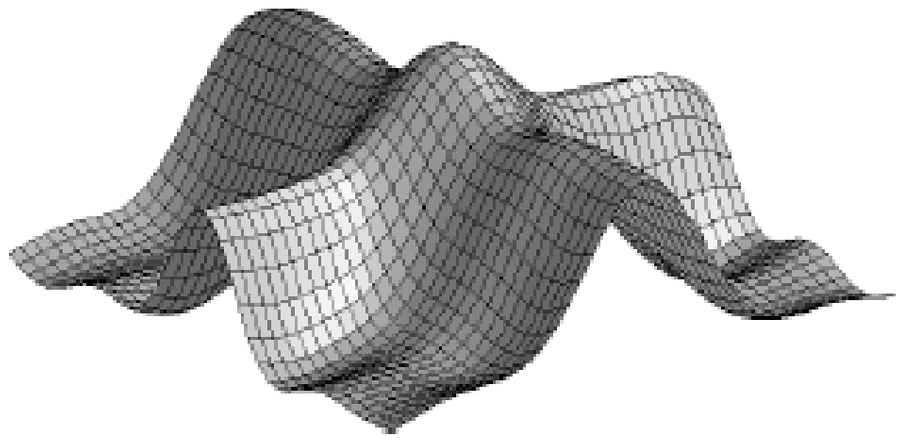
(a) Random Noisy Data



(b) Ground Truth



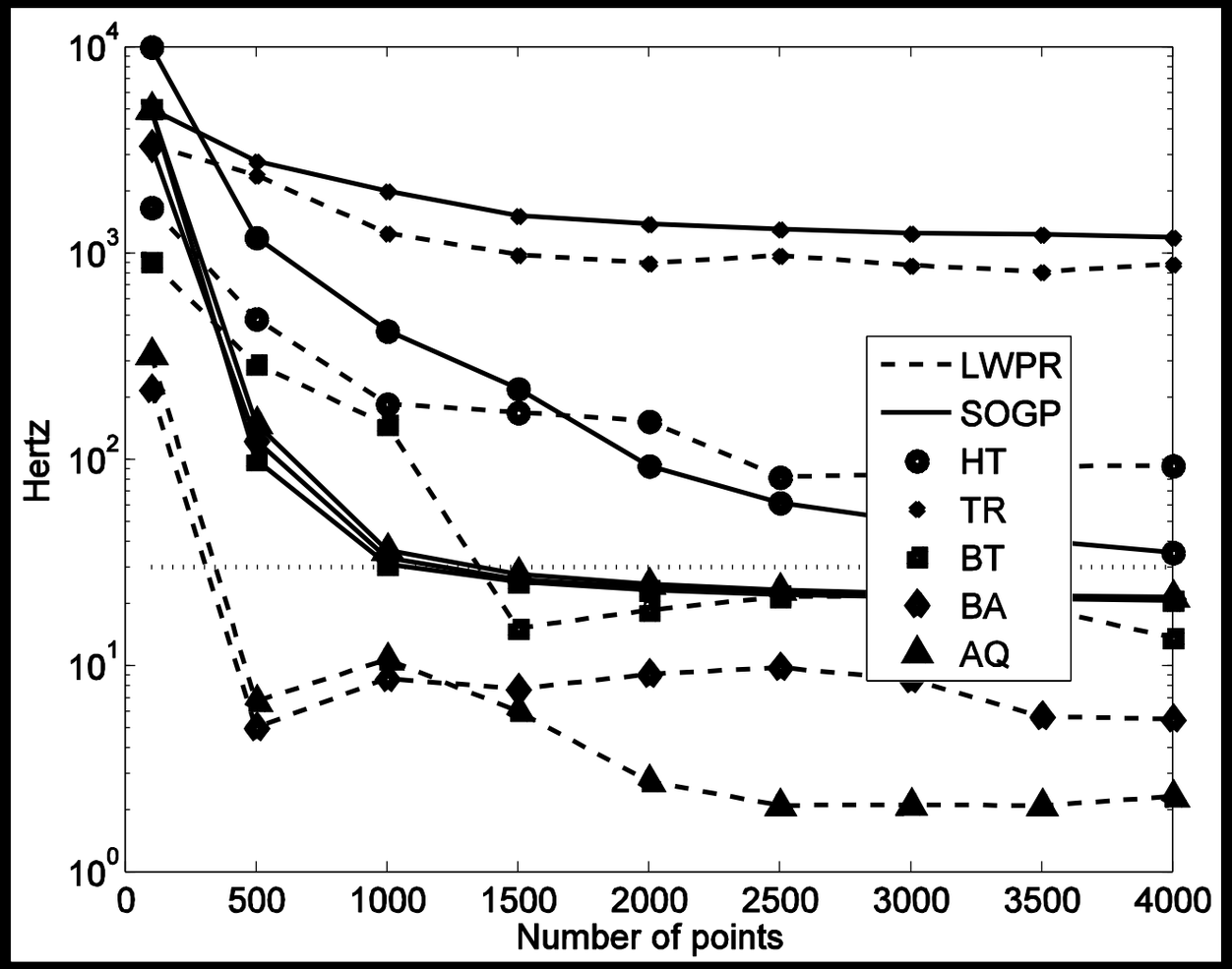
(c) LWPR, MSE = 0.0202



(d) SOGP, MSE = 0.0244

Timing!

- MUST run as fast as data



Model Selection

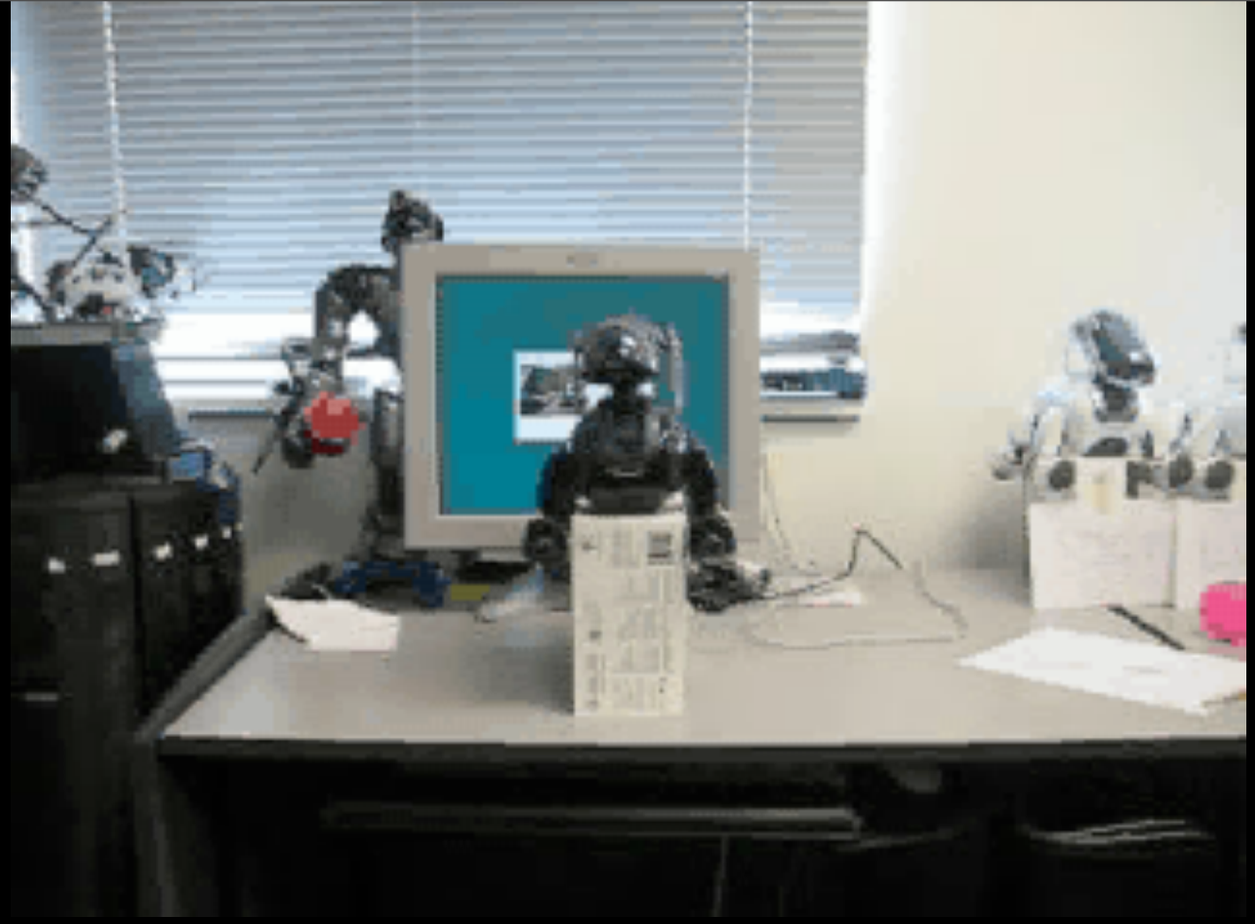
- Where the RFs / basis functions are located
- Wrong locations -> too many -> slow
  - Also wrong answers

33

Task learning

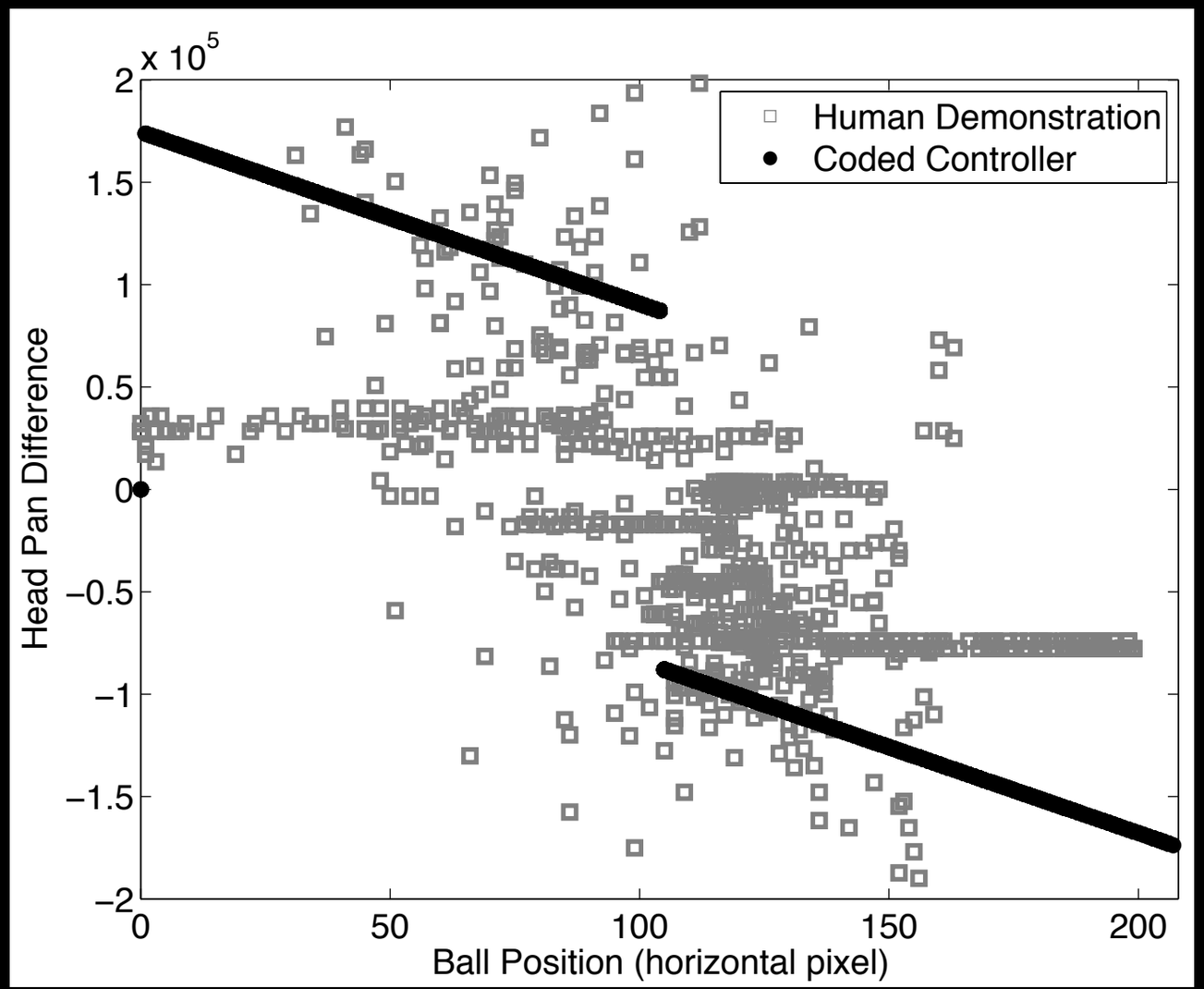
Task learning

34

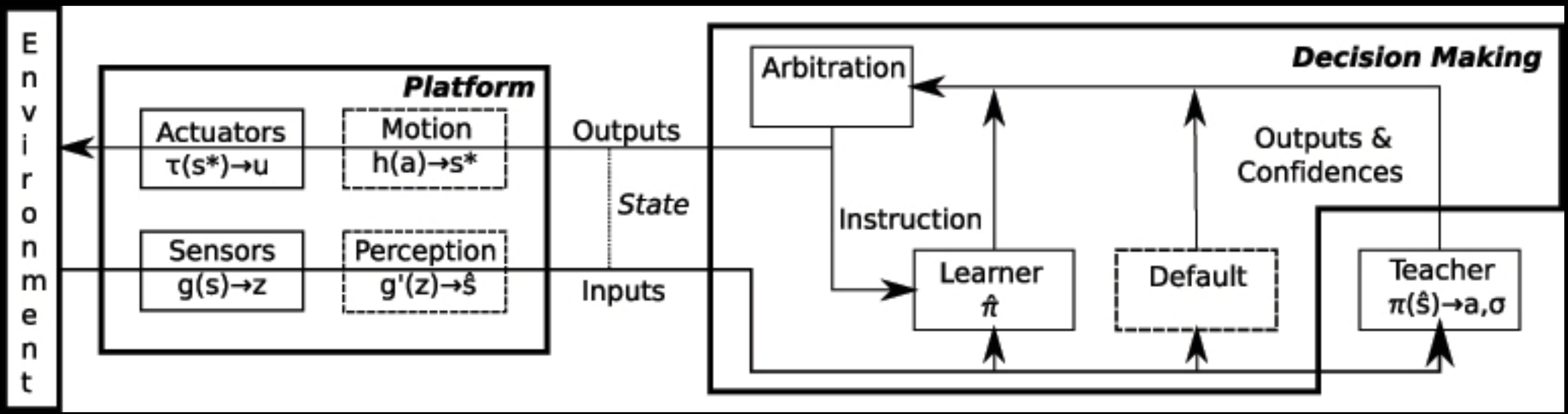


35

Human control



Recap


Environment
Platform
Actuators
τ(s*)→u
Motion
h(a)→s*
Outputs
State
Sensors
g(s)→z
Perception
g'(z)→ŝ
Inputs
Decision Making
Arbitration
Outputs &
Confidences
Instruction
Learner
π̂t
Default
Teacher
π(ŝ)→a,σ

37

Current Issues

- Functional assumption is often wrong
  - Turn left or Right vs Walk
  - State
- Overlapping mixtures of experts?

38

Data collection

- Need lots of human demonstration to learn unknown tasks
  - Video Game

39

Unknown Task?

- Learn the task goal
  - Reinforcement Learning

40

Questions?

- Pictures and videos copyright their respective owners

41