

A Bayesian Model of Imitation in Infants and Robots

Rajesh Rao, Aaron Shon and Andrew Meltzoff
(2004)

Presented by Micha Elsner

How we gain new skills

- Maturation (more neurons, muscle power &c)
- Reinforcement learning (“trial and error”)
 - Behaviorists (Skinner, Watson)
- Independent invention and discovery
 - Piaget's theory: children are “little philosophers” who learn abstract principles from experience
- Imitation
 - More flexible than maturation
 - More efficient than discovery

Piaget's learning

- Relies on two processes:
 - assimilation: applies a known behavior (*schema*) in a new way... “grab the rattle” --> “grab the watch”
 - accomodation: adapts a known behavior to new circumstances... applies when assimilation fails
“grab the rattle” --> “grab the beach ball” -->
“squeeze the beach ball”
- Stages of cognitive development
 - Some dominated by assimilation, some by accomodation

Constructivism

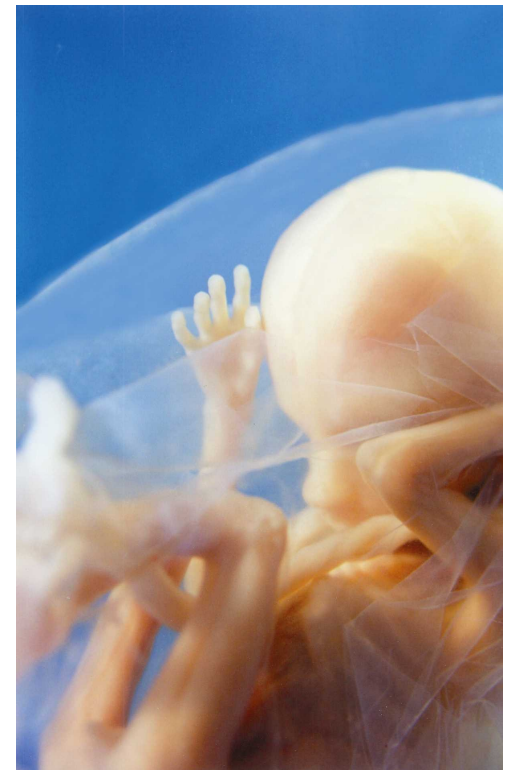
- Knowledge is 'constructed' from a combination of experience and innate principles.
 - Representations of the world are iteratively improved as they become inadequate
 - For instance, children have to learn 'conservation of mass' and 'object permanence'
- Something like our usual unsupervised learning
 - Clustering
 - Rule inference / Latent-variable modeling

How humans learn to imitate

- Famous four-stage model due to Meltzoff
- Body babbling
- Imitating body movements
- Imitating actions with objects
- Imitating intentional actions

Body babbling

- Repetitive, undirected movements
- Learn a mapping between nerve impulses and body state
- Begins *in utero*
- Relies on innate proprioception (ability to know where one's body parts are)
- But the mapping isn't innate!



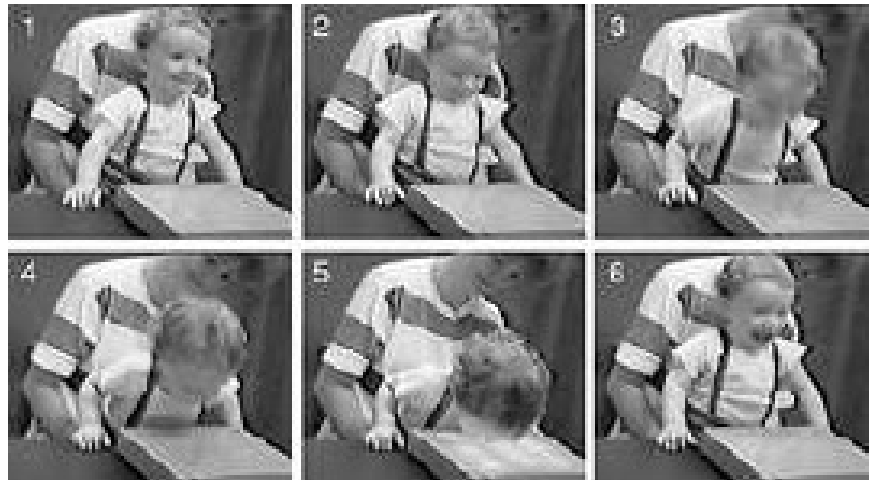
Imitating body movements

- Infants begin to do this right after birth
 - Uses the mapping between nerve impulses and body state learned from babbling.
 - Also requires map from visual system (observed state) to proprioception (own state)



Imitation with objects

- More complex dynamics
 - Using an object, not just the body
 - Starts at about 1 year old.
 - Not really modeled in this paper!
- A famous experiment
 - Adult 'teacher' presses a button with his forehead
 - Infants imitate him



Intentionality

- Full 'imitation' is not just mimicry
 - Learner may have different actions than teacher
 - Has to reach the same goal in a different way
 - (cf. pendulum upswing, Atkeson & Schaal)
- Starts at about 18 months.
 - Understand that humans have intentions
 - Learn from a demonstrator who makes 'mistakes'

Learning framework

- Uses an MDP-like structure.
- What we won't cover:
 - perception (inferring our state from observations; proprioception)
 - correspondence (inferring someone else's state from our observations)
 - discretization (clustering states and actions)
 - intention recognition (learning a useful prior over goal states)

Markov representation

- Discrete states s , actions a and time t .
- Define 'imitation' as 'following a memorized trajectory' $s_1 \rightarrow s_2 \rightarrow s_3 \dots s_g$
 - Isn't this just mimicry?
- We need a way (inverse model) to get us from state s_t to goal s_{t+1} .
- Optimal action selection is deterministic (MAP solution).
 - Humans sometimes use probability matching.

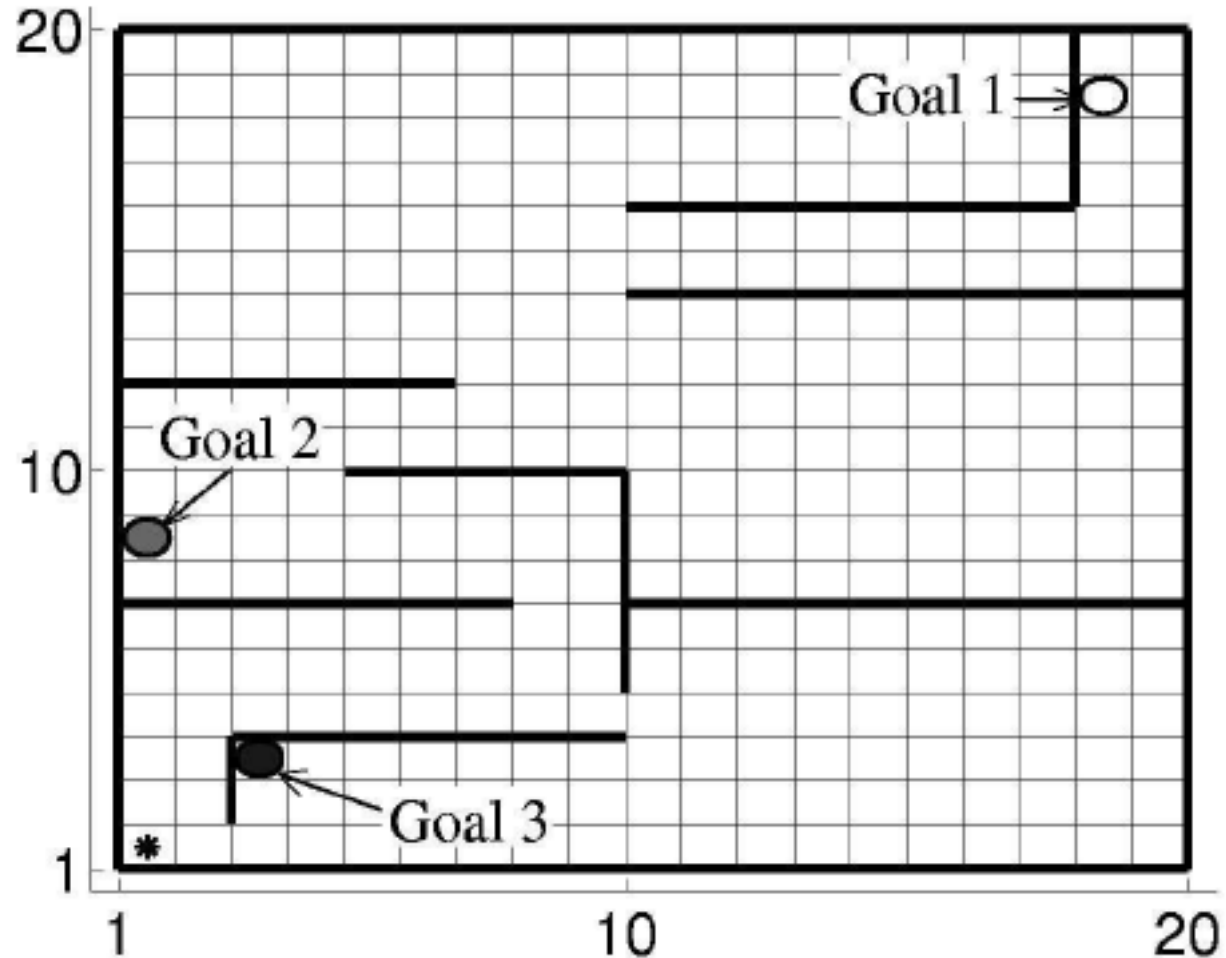
Forward model

- Maps state, action to next state:
 - $p(s_{t+1} | s_t, a_t)$
- Learned from exploring the state-space at random
 - Body babbling
 - Supervised process (assuming proprioception)

Inverse model

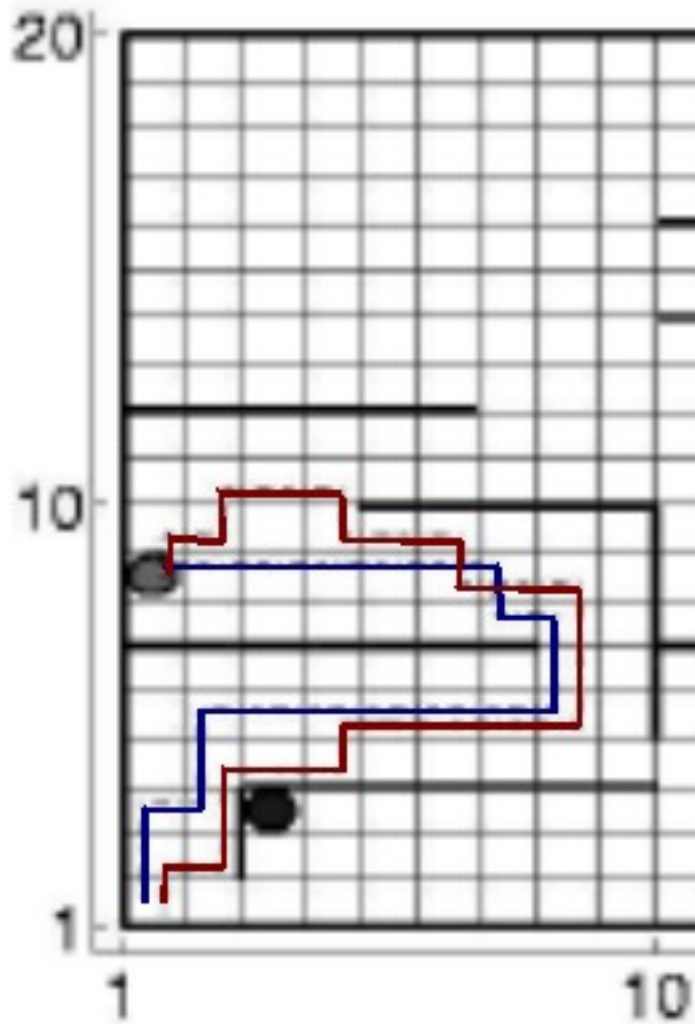
- $p(a_t | s_{t'}, s_{t+1}', s_g)$: probability that an action is chosen
 - given the desired next state, and the goal
- $p(a_t | s_{t'}, s_{t+1}', s_g) \propto p(s_{t+1}' | s_{t'}, a_t) * p(a_t | s_{t'}, s_g)$
 - assuming that the forward model is independent of the goal state
 - the prior over actions is learned (max likelihood or MAP) from observing the teacher

Welcome to mazeworld

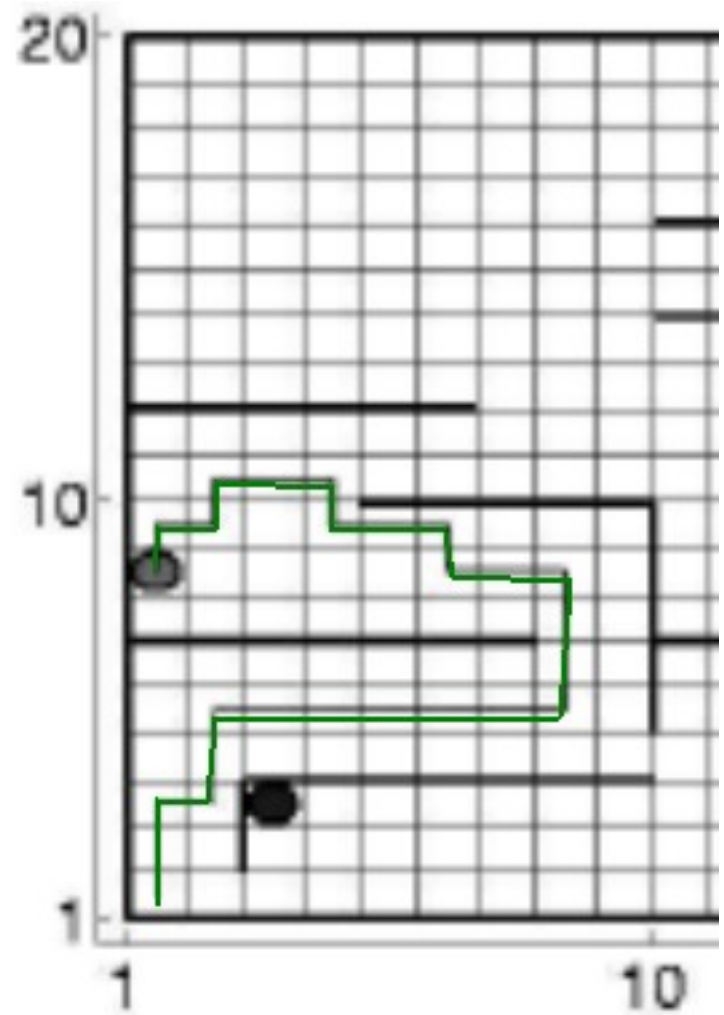


Simple stochastic transition model (easy to learn)

Learning the maze



Demonstration trajectories



Learner trajectory

Inferring intentions

- If we have a prior over goals, we can find the posterior over the teacher's intentions.
- No discussion of where this prior comes from.

Conclusions

- The authors propose to get some robots and test using real data.
 - I imagine they'll have problems.
 - It looks like this is just the simple part...
- But the human model has some interesting elements
 - Intention recognition would certainly help out robotics.

