Overview

Bayesian model of imitation in infants and robots: by Rao et al. (2004).

Stages in imitation:

- Body babbling.
- Imitation of body movements.
- Imitation of actions on objects.
- Imitation based on inferring intentions of others.

Stages in Imitation (I)

Body babbling:

- Repetitive motion.
- Establishes mapping between movement and bodily configurations.
- Builds an “internal model”.

Stages in Imitation (II)

Imitation of body movements

- Main issue: correspondence problem.
- Example: tongue protrusion.
- Properties: deferred imitation, correction of imitative response without any feedback.

Stages in Imitation (III)

Imitation of actions on objects

- Interaction with object with a particular body part.
- Novel ways of interaction with an object are also mimicked.
- Imitation shown after 1 day, or even 1 week.

Stages in Imitation (IV)

Inferring intentions

- Imitates **unsuccessful** acts.
- Imitates with different body parts (means), to achieve the same (inferred) goal (ends): e.g., use of legs instead of hand to hold large object in place.
- Ignorance of inanimate device.

Bayesian Framework

Body Babbling: Learning Internal Models

- Forward model:
  Current state, action $\longrightarrow$ next state.
- Inverse model:
  Current state, desired state $\longrightarrow$ action.
- Hybrid approach: Estimate inverse model using forward model and constraints on actions (priors).

Bayesian Imitative Learning

- Perceptual input: $I_1, I_2, ..., I_N$.
- States of observed objects: $s_1, s_2, ..., s_N$.
  - $s_t \in \{S_1, S_2, ..., S_M\}$, at time t.
  - s_N is the **goal state**
  - **correspondence problem**: how does s_t relate to ones own body state?

Action Selection from Memorized States

- Given a state sequence, e.g., $S_7 \to S_1 \to ... \to S_{12}$,
- current state $s_t = S_i$, and
- goal state $s_G = S_k$,
- find action a_t to maximize the probability of $s_{t+1} = S_j$ from the memorized sequence (**inverse model**):

$$P(a_t = A_i | s_t = S_i, s_{t+1} = S_j, s_G = S_k)$$

- Sensory consequence of action (**forward model**) is also probabilistic:

$$P(s_{t+1} = S_j | s_t = S_i, a_t = A_i)$$

(Forward model is determined by the environment alone, thus s_G is not needed.)

Let's Simplify the Notation (YC)

- A for action: $a_t = A_i$
- C for current state: $s_t = S_i$
- N for next state: $s_{t+1} = S_j$
- G for goal state: $s_G = S_k$
- Forward model:

$$P(N|C, A, G) = P(N|C, A)$$

- Inverse model:

$$P(A|C, N, G)$$

Estimating the Inverse Model

Certain probabilities that can be learned easily:

- Forward model (through body babbling):

$$P(N|C, A)$$

- Relationship between intermediate states and the goal (by observing the teacher):

$$P(N|C, G)$$

- Prior probabilities on actions (by observing the teacher):

$$P(A|C, G)$$

From these, the **inverse model** can be learned:

$$P(A|C, N, G)$$

Estimating the Inverse Model (II)

- Recall the extended (conditionalized) Bayes rule:

$$P(X|Y, E) = \frac{P(Y|X, E)P(X|E)}{P(Y|E)}$$

and that the forward model is independent of the goal G:

$$P(N|C, A, G) = P(N|C, A)$$

- From which we can calculate the inverse model as:

$$P(A|C, N, G) = cP(N|C, A)P(A|C, G)$$

where $c = 1/P(N|C, G)$.

Estimating the Inverse Model (III)

- Finally, $P(N|C,G)$ can be calculated by marginalizing over A:

$$\begin{aligned} P(N|C,G) &= \sum_m P(N|C,G,A_m)P(A_m|C,G) \\ &= \sum_m P(N|C,A_m)P(A_m|C,G) \end{aligned}$$

again using the independence of the forward model from the goal.

Inferring the Intent

- Where is the teacher eventually headed to?
- That can also be estimated, given the teacher's action, and current and next state.

$$\begin{aligned} P(G|A,C,N) &= k_1 P(N|C,A,G)P(G|C,A) \\ &= k_2 P(N|C,A,G)P(A|C,G)P(G|C) \\ &= k_3 \underbrace{P(N|C,A,G)}_{\text{forward model}} \underbrace{P(A|C,G)}_{\text{prior}} P(C|G)P(G) \end{aligned}$$

where k_i are normalization constants.

Putting Everything Together

Training:

1. Imitator learns the forward model through body babbling:

$$P(N|C,A)$$

2. Teacher shows a sequence of actions/states leading to one of the goals: $s_1, s_2, ..., s_G$
3. Imitator learns the prior from the teacher's trajectory:

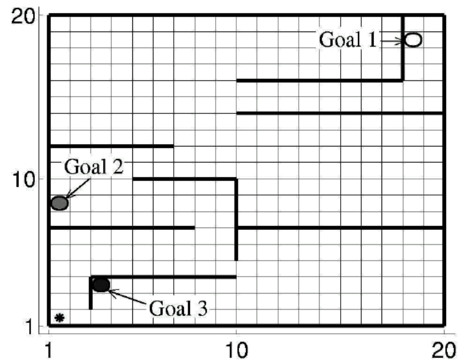
$$P(A|C,G)$$

Putting Everything Together

Testing:

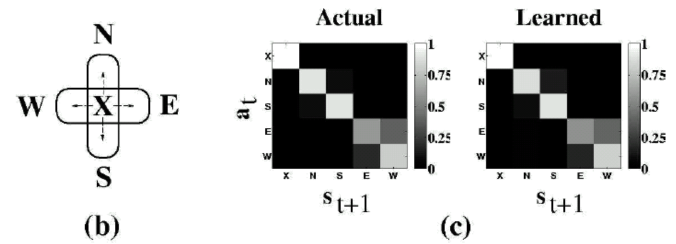
1. Teacher starts moving, and is still far from the goal.
2. Observing only A, C, and N, the imitator can infer the teacher's G.
3. The imitator can also generate A based on maximum posterior probability (or stochastically).

Example Domain: Maze



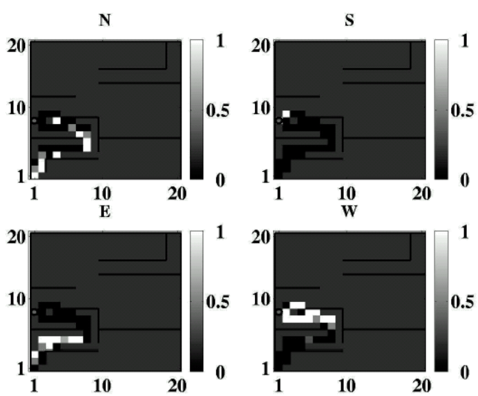
- Grid maze.
- C, N, G are (x, y) locations in the grid.
- Several goal locations.
- $A \in \{\mathrm{N}, \mathrm{S}, \mathrm{W}, \mathrm{E}, \mathrm{X}\}$, where X is to stay at that location.

Forward Model



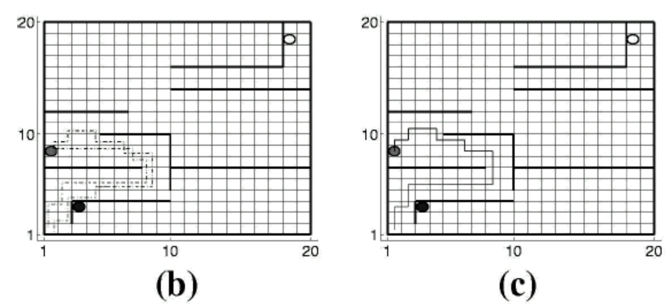
- Given C and N, we can say what direction N is, relative to C: N. S, W, E, or X (this is what is shown in the column index in (c) above as s_{t+1}).
- Action A (shown as a_t) is the action taken at time t.

Prior



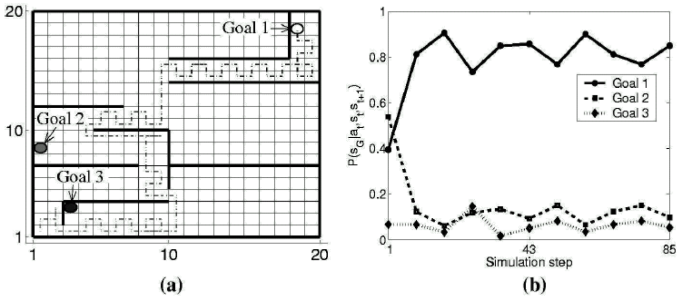
- Given a particular goal G and the current location, we can learn $P(A|C, G)$.
- The prior can be learned by counting how many times a particular action was taken when at C, while the end goal was G.

Imitation Run



- Teacher's run (left) and the imitator's run (right) are shown.
- The goal in this case was $(1, 9)$.

Learning the Intent



- The teacher starts from $(1, 1)$, while moving along to goal 1.
- The imitator can infer the intended goal of the teacher before the teacher reaches the goal (shown on the right).

Conclusion

- A Bayesian framework for imitation learning.
- Can deal with noisy and uncertain environments.
- Can help understand imitation learning in humans.

Discussion (YC)

- Does the algorithm assume that the imitator know A of the teacher?
- How does A differ from action inferred from just observing C and N? Is the real A knowable at all?
- Correspondence problem is not resolved (as the authors state up front).
- The problem of “Goal”.
- Are the conditional probabilities easy to estimate?: $P(N|C, A), P(A|C, G),$ etc.?
- How are mirror neurons relevant to these discussions, particularly that of **intention**?

References

Rao, R. P. N., Shon, A. P., and Meltzoff, A. N. (2004). A bayesian model of imitation in infants and robots. Cambridge, UK: Cambridge University Press. In press.