Learning What the Internal State Means, Through Action

Yoonsuck Choe
Department of Computer Science
Texas A&M University

† Joint work with S. Kumar Bhamidipati, Daniel Eng, Navendu Misra, Stuart B. Heinrich, Noah H. Smith, and Huei-Fang Yang

The Question

• What do these green lights mean (following slides)?
What Do Those Green Lights Represent?

- It is hard to get any idea at all.
- Actually, this is how it might be like looking at the brain's activity from the inside of the brain.
They Are Visual Cortical Responses to Oriented Lines

http://faculty.cs.tamu.edu/choe
They Are Visual Cortical Responses to Oriented Lines
They Are Visual Cortical Responses to Oriented Lines

The Main Question

How can we understand what the pattern of activity in the brain means? (cf. Freeman 2003)

1. How can scientists understand the pattern?
2. How does the brain itself make sense of its own activity?

Scientist vs. the Brain

- External observer (e.g., a neuroscientist) can figure out how $S$ relates to $I$ (transformation $f : I \rightarrow S$).
- Internal observer cannot: But the brain does this all the time, so this does not seem right!

Example: The Visual Cortex

- But even before that, and with access to only $S$, humans had no problem perceiving orientation.
Theories on RF Formation

Well-developed understanding on how RFs form:


However, how is the resulting code to be used remains a question.

A Metaphor of the Problem

• Imagine sitting in a room, looking at blinking lights, without knowledge of the sensors nor the RFs.
• The lights may be due to any other sensory modality (as in vision-audition rewiring Sur et al. 1999).
• Similar to the Chinese Room (Searle 1980): Problem of “Symbol Grounding” (Harnad 1990).

The Sensory Organ Can (Possibly) Give Us a Clue

• It could have been caused by a visual input.

But, Equally Likely Is ...

• It could have been caused by an auditory input.
• Sur et al., Rewiring cortex, Journal of Physiology, 41:33–43, 1999
Rewiring Vision to Auditory Area

von Melchner et al. (2000); Sharma et al. (2000); Sur et al. (1999)

- Rewired auditory cortex develops visual cortex-like organization.
- Question: does it see light or hear light?

Possible Solution: Through Action

- A major problem in the metaphor is the passiveness of the whole situation.
- Adding action can help solve the problem.
- But why and how?

Experimental Evidence

Held and Hein (1963)

- Active animal developed normal vision.
- Passive animal did not.
- Suggests the importance of action in vision.

Rewiring: Behavioral Results

von Melchner et al. (2000); Sharma et al. (2000)

- Ferret trained to behave differently for visual vs. auditory stimuli: Behavior suggests that the ferret is actually seeing light with its auditory cortex!
**Experimental Evidence**

- Vibrotactile array linked to a video camera.
- Passive viewing results in tactile sensation.
- Moving the camera results in a vision-like sensation.
- Sensation as related to voluntary/intentional action may be the key!

Bach y Rita (1972; 1983)

**Theoretical Insights**

- Philipona et al. (2003) showed that properties of ambient space (such as the dimensionality) can be inferred based on internal sensory input alone.
- The key concept is about the compensability between ego-motion and the change in the environmental input conveyed to exteroceptors.

**Approach: A Sensorimotor Agent**

- A simple visuomotor agent.
- How can it learn about the visual world?
- What should be the objective (or goal) of learning?

Choe and Bhamidipati (2003)

**Action for Unchanging Internal State**

- Diagonal motion causes the internal state to remain unchanging over time.
- Property of such a movement exactly reflects the property of the input $I$: Semantics figured out through action.
Action for Unchanging Internal State

- Diagonal motion causes the *internal state* to **remain unchanging** over time.
- Property of such a movement **exactly reflects** the property of the input $I$: Semantics figured out through action.

Action for Internal Invariance

- Agent can **move** its visual field.
- Movement in a certain direction (diagonal) causes the *sensory array* to **stay invariant** over time.
- Property of such a movement **exactly reflects** the property of the input $I$.

Outline of Experimental Methods

- Input preparation.
- Orientation response calculation.
- Learning algorithm and policy generation.
Methods: Input Preparation

- Convolve with Difference-of-Gaussian (DoG) filter $(15 \times 15)$.
- Then, sample a $31 \times 31$ region.

Methods: Orientation Response

- Find the vectorized dot product of the $31 \times 31$ input $I$ and the $n$ Gabor filters $G_i (i = 1..n, \theta = [(i - 1)\pi/n])$:
  \[
  r_i = \sum_{x,y} G_i(x,y)I(x,y).
  \]
- The above results in a response vector $r$, and the orientation response $s$:
  \[
  s = \arg \max_{i=1..n} r_i
  \]

Methods: Reinforcement Learning

- Immediate reward is measured as the dot product of current and previous response vectors:
  \[
  \rho_{t+1} = r_t \cdot r_{t+1}
  \]
- The task the agent is to learn a state-to-action mapping so that it maximizes the reward $\rho$.
Methods: Policy $\pi$

Suppose we know the probability $P(a|s)$ (let us call this $R(s, a)$), where stochastically generating action given the state $s$ with this probability maximizes the reward.

1. Given the current state $s_t \in S$, randomly pick action $a_t \in A$.
2. If $a_t$ equals $\arg \max_{a \in A} R(s_t, a)$,
   (a) then perform action $a_t$,
   (b) else perform action $a_t$ with probability $R(s_t, a_t)$.
3. Repeat steps 1 to 3 until exactly one action is performed.

In practice, momentum was added so that $a_{t+1} = a_t$ with a 30% chance, and in step 2, if a random draw from $[0..1]$ was less than $cR(s_t, a_t)$, then the action was accepted.

Methods: Learning $R(s, a)$

- A simple update rule was used:
  $$R_{t+1}(s_t, a_t) = R_t(s_t, a_t) + \alpha \rho_t,$$
  where $\alpha = 0.002$ is the learning rate, and $\rho_t$ the immediate reward.
- $R_{t+1}(s_t, a)$ was then normalized by:
  $$R_{t+1}(s_t, a) := \frac{R_{t+1}(s_t, a)}{\sum_{a' \in A} R_{t+1}(s_t, a')}$$
  for all $a$.

Results: Overview

1. Synthetic input and natural image input.
2. Learned $R(s, a)$.
3. Error in $R(s, a)$ and average reward $\rho$ over time.
4. Distribution of reward $\rho$.
5. Gaze trajectory.

Reward Probability Table

<table>
<thead>
<tr>
<th>S: sensory state (orientation)</th>
<th>A: direction of motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5 0 0 0 0.5 0 0 0</td>
<td>0 0 0 0 0.5 0 0 0</td>
</tr>
<tr>
<td>0 0.5 0 0 0 0.5 0 0</td>
<td>0 0 R(s, a) 0 0 0.5 0</td>
</tr>
<tr>
<td>0 0 0 0 0.5 0 0 0</td>
<td>0 0 0 0.5 0 0 0 0.5</td>
</tr>
</tbody>
</table>

- Reward probability $R(s, a)$ can be tabulated.
- In an ideal case (world consists of straight lines only), we expect to see two diagonal matrices (shaded gray, above).
Results: Learned $R(s, a)$ for Synthetic Input

- Learned $R(s, a)$ close to ideal.

Results: Learned $R(s, a)$ for Natural Images

- Learned $R(s, a)$ close to ideal even for natural image inputs.

Results: Error in $R(s, a)$

- Left: Root-mean-squared error in $R(s, a)$ compared to the ideal case.
- Right: running average of immediate reward $\rho$:
  
  \[ \mu_t = (1 - \alpha)r_t + \alpha \mu_{t-1}, (\mu_1 = \rho_1, \alpha = 0.999). \]
Results: Average $\rho$

- $\rho$ values for different categories:
  - Flowers: Initial and Final plots showing reward distribution.
  - Ducks: Initial and Final plots showing reward distribution.
  - Plant: Initial and Final plots showing reward distribution.
  - Oleander: Initial and Final plots showing reward distribution.

Results: Distribution of $\rho$

- Synthetic Input:
  - Initially, two peaks: near negative min and positive max $\rho$.
  - Near the end, only one peak: near positive max $\rho$.

Results: Gaze Traj. for Synth. Input

- Gaze trajectory reflects orientation represented by internal state.
Results: Gaze Traj. for Nat. Input

(a) Flowers  (b) Ducks

Initial

Final

(c) Plant  (d) Oleander

Results: Demo

Work in Progress: Q-Learning

(a) Initial  (b) Learned

Trajectories from Q-Learning sessions (Choe and Smith 2006).
Interpretation of the Results

• Using invariance as the only criterion, particular action pattern that has the same property as the input that triggered the sensors was learned.

• Question: Can this approach be extended to learning complex stimulus concepts?

Learning About Complex Objects

• For complex objects, a history of sensory activity may be needed (i.e., some form of memory).

• Invariance can be detected in the spatiotemporal pattern of sensor activity.

Supporting Evidence?

Yarbus (1967)

• When we look at objects, our gaze wanders around.

• Could such an interaction be necessary for object recognition?

Advantage of Motor-Based Memory

(Habit, or Skill)

• Sensor-based representations may be hard to learn and inefficient.

• Motor-based approaches may generalize better.

• Comparison: Make both into a 900-D vector and compare backpropagation learning performance.
Class Separability

- Comparison of PCA projection of 1,000 data points in the visual and motor memory representations.
- Motor memory is clearly separable.

Speed and Accuracy of Learning

- Motor-based memory resulted in faster and more accurate learning (10 trials).

Summary

- Internal observer can learn about the properties of the external environment – through action maximizing invariance in neural activity.
- Such actions closely reflect the property of the stimulus that triggered the sensory neuron to fire: Meaning of the spike recovered (through action)!
- Main contribution: The invariance criterion for autonomously learning the meaning of neural states.

Related Work (Selected)

- Piaget (1952): Sensorimotor period in child development
- Freeman (1999): Brain creates meaning through action and choices. Also see Kozma and Freeman (2003) for a KIV model of the emergence of goal-directed, intentional behavior.
- O’Regan and Noë (2001): Sensorimotor contingency theory
- Philipona et al. (2003): Inferring space through sensorimotor interaction
- Rizzolatti et al. (2001): Mirror neurons
- Gibson (1950): Direct perception of invariance and affordance
- Taylor (1999): Corollary discharge and awareness of attention movement prior to sensory awareness.
Discussion

• Why is knowing ones own action any easier than perceptual interpretation?: Knowledge of own action may be more immediate than perception (cf. Moore 1996, citing Bergson).

• What gives rise to voluntary, intentional action and why is it special? (Freeman 1999; Kozma and Freeman 2003; Taylor 1999).

• A different view of invariance: Not (only) something to be detected in the environment (cf. Gibson 1950), but something that we actively seek within.

Discussion (Cont’d)

• Why not just analyze the input directly?: The raw input is only available at the immediate sensory surface.

• What about other sensory modalities (such as touch, olfaction, or audition)?

• The learning scheme depends on structure in the environment: If the environment didn’t have structure, the agent can never learn.

Discussion (Cont’d)

• Relation to mirror neurons (Rizzolatti et al. 2001)?

• Role of attention (e.g. Rensink et al. 1997; Taylor 1999)?: Attention may be needed when ambiguities are present.

• Do motor primitives restrict the kind of sensory property that can be learned? What kinds of motor primitive do we have?

• What about meaning other than sensorimotor-like, such as reinforcement signals (Rolls 2001) or “feeling” (Harnad 2001)?

• Grounding on perception alone may not be sufficient: cf. Perceptual symbol system (Barsalou et al. 2003).

• What to make of the segregation in the dorsal–ventral pathway? (Goodale and Milner 1992).
Predictions

• Perceived orientation of a line can be altered by eye movement in the direction of incompatible orientation.
• Motor structures (cerebellum, basal ganglia) may be intimately involved in semantics.
• Geometrical understanding may be limited by the motor primitive repertoire.

Future Work (and Work in Progress)

• Learning receptive field structure based on SIDA.
• Lateral inhibition in sensory array.
• Crossmodal association through sensory invariance.
• Extending to more complex concepts.

Conclusions

• We must ask how the brain understands itself.
• Autonomous understanding of own internal state is non-trivial without direct access to the stimulus.
• Action can help solve the conundrum.
• Action that maintains invariance in internal state can recover meaning (the property of the stimulus).

Credits

• Kuncara A. Suksadadi helped in the early stages of the idea’s development.
• Thanks to Ricardo Gutierrez-Osuna, Ronnie Ward, Stevan Harnad, James Clark, and Ben Kuipers for helpful discussions.
• Thanks to Texas A&M Cognoscenti and NIL members for insightful comments.
• Partially supported by Texas Higher Education Coordinating Board (ATP 000512-0217-2001).
Why Do We Have a Brain?

- Tree (no Brain)
- Tunicate Free-floating (w/ Brain)
- Tunicate Settled (w/o Brain)


References


55  http://faculty.cs.tamu.edu/choe

56  http://faculty.cs.tamu.edu/choe

58  http://faculty.cs.tamu.edu/choe

60  http://faculty.cs.tamu.edu/choe


http://faculty.cs.tamu.edu/choe