Predictive Coding for Robot Cognition

Yukie Nagai
NICT / Bielefeld University
Take Home Messages

• **Predictive coding/learning** is a powerful unified theory for designing *cognitive mechanisms* in robots.
  – Various dynamics of cognitive development can be reproduced based on the theory.

• **Consciousness** emerge as a *byproduct of the development of cognition functions*.
  – Cognitive behaviors require internal models which produce consciousness.
14-month-old Infants Cannot Recognize Self

(Discovery Channel “The Baby Human,” 2004)
... But, They Can Help Others
Underlying Mechanism?

**Self cognition** (from 24 mo)
[Amsterdam, 1972; Povinelli et al., 1996]

**Helping others** (from 14 mo)
[Warneken & Tomasello, 2006]

**Joint attention** (from 12 mo)
[Butterworth & Jarrett, 1991]
[Moore et al., 1996]
[Brooks & Meltzoff, 2002]

**Imitation** (from 0 mo)
[Meltzoff & Moore, 1977]
[Heyes, 2001]

**Reading intention** (from 6 mo)
[Woodward, 1998]
[Gergely et al., 1995]

**Expressing emotion** (from 6 mo)
[Bridges, 1930; Lewis, 2007]
Our Hypothesis: Predictive Learning
[Nagai & Asada, IROS-WS 2015]

Minimization of prediction error $e(t+1)$ leads to cognitive development during early infancy.

Prediction error

$e(t+1) = s(t+1) - \hat{s}(t+1)$

(Modified from [Blakemore et al., 1999])
Our Hypothesis: Predictive Learning
[Nagai & Asada, IROS-WS 2015]

Minimization of prediction error $e(t+1)$ leads to
cognitive development during early infancy.

(a) Update the predictor through
sensorimotor experiences
→ Self-other cognition, goal-directed action, etc.

(b) Execute a predicted action in response to others’ action
→ Imitation, helping action, etc.
Multimodal Predictive Learning for Action Production and Perception

- Predictive learning to integrate multimodal signals enables infants to recall own motor experiences while observing others’ action (i.e., mirror neuron system).
  - Motor learning: integrating visual $v$, tactile $u$, and motor $m$ signals

[Copete, Nagai, & Asada, ICDL-EpiRob 2016]
Multimodal Predictive Learning for Action Production and Perception

- Predictive learning to integrate multimodal signals enables infants to recall own motor experiences while observing others’ action (i.e., mirror neuron system).
  - Motor learning: integrating visual $v$, tactile $u$, and motor $m$ signals
  - Action observation: generating imaginary $u$ and $m$ from actual $v \rightarrow$ better prediction of $v$

[Copete, Nagai, & Asada, ICDL-EpiRob 2016]
Results 1: Prediction of Others’ Goal

Input/output signals:
- Vision: camera image (30)
- Tactile: on/off (3)
- Motor: joint angles of shoulder and elbow (4)

<table>
<thead>
<tr>
<th>Predicted image</th>
<th>Classifications of prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image]</td>
<td>Correct goal</td>
</tr>
<tr>
<td>![Image]</td>
<td>Incorrect goal</td>
</tr>
<tr>
<td>![Image]</td>
<td>No goal</td>
</tr>
</tbody>
</table>

[Copete, Nagai, & Asada, ICDL-EpiRob 2016]
Result 2: Motor Experience Improved Accuracy of Prediction

**W/ motor experience**

**W/o motor experience (only observation)**

- **Correct goal**
- **Incorrect goal**
- **No goal**

Reaching for left
Reaching for center
Reaching for right

[Copete, Nagai, & Asada, ICDL-EpiRob 2016]
Multimodal Predictive Learning for Emotion Imitation

• Emotion is perceived through inference of interoceptive and exteroceptive signals. [Seth et al., 2012]

• Predictive learning of multimodal signals enables infants to estimate and imitate others’ emotion by *putting themselves in others’ shoes* (i.e., *mirror neuron system*).

[Horii, Nagai, & Asada, Paladyn 2016]
Prediction of Sensory Signals Improves the Estimation of Emotion

Only auditory input is given. → Imaginary visual signals improved the estimation of emotion.

[Horii, Nagai, & Asada, Paladyn 2016]
Different Level of Consciousness in Autism Spectrum Disorder
What Letters Can You See?

- People with ASD recognize the local letter quicker than the global letter. [Behrmann et al., 2006]
Difficulty in Feeling Hunger in ASD

• Feeling of hunger is hard to be recognized and requires conscious process of selecting and integrating proper sensory signals. [Ayaya & Kumagaya, 2008]

1. Equally perceive multimodal sensations
2. Enhance hunger-relevant signals while diminishing hunger-irrelevant signals
3. Recognize hunger by integrating relevant signals

- : likely relevant to hunger
- : likely irrelevant to hunger
- : limited to hunger
- : psychological
Schizophrenia-like Behaviors Generated by Modifications in Neural Network [Yamashita & Tani, 2012]

- Multiple timescale recurrent neural network (MTRNN)
  - Lower layer (*fast* context): behavioral primitives
  - Higher layer (*slow* context): combinations of primitives

Use of the robot was made possible through cooperation with the SCNY Corporation.
Schizophrenia-like Behaviors Generated by Modifications in Neural Network [Yamashita & Tani, 2012]

- Multiple timescale recurrent neural network (MTRNN)
  - Lower layer (fast context): behavioral primitives
  - Higher layer (slow context): combinations of primitives
Atypical Visual Perception in ASD
Take Home Messages

• **Predictive coding/learning** is a powerful unified theory for designing cognitive mechanisms in robots.
  – Various dynamics of cognitive development can be reproduced based on the theory.

• **Consciousness** emerge as a byproduct of the development of cognition functions.
  – Cognitive behaviors require internal models which produce consciousness.

• Different levels of consciousness in ASD provide deeper insights into the roles of consciousness.
Thank You!

Osaka University
• Minoru Asada
• Jimmy Baraglia
• Takato Horii
• Shibo Qin
• Jorge L. Copete
• Konstantinos Theofilis
• Jyh-Jong Hsieh, et al.

University of Tokyo
• Shinichiro Kumagaya
• Satsuki Ayaya

yukie@nict.go.jp
http://developmental-robotics.jp