Symbol Emergence in Robotics and Unsupervised Machine Learning for Language Acquisition

Tadahiro Taniguchi
College of Information Science & Engineering
Ritsumeikan University
(Visiting associate professor, Imperial College London)

Japan-UK Robotics and Artificial Intelligence Seminar 2016
The Embassy of Japan, London
18th, February, 2016
Computational Understanding of Mental Development
From Behavioral Learning to Language Acquisition

- A human child acquires many physical skills, concepts, and knowledge, including language, through physical and social interaction with his/her environment.
- How do we become able to communicate via symbols?
- We’d like to obtain an understanding of the computational process of mental development and language acquisition.

Symbol Emergence in Robotics
Symbol Emergence in Robotics

Multimodal Communication

Concept Formation

Language Acquisition and Mental Development

Learning Motor Skills and Segmentation of Time-Series Information

Emergence of Communication

Learning Interaction Strategy

Emergent Symbol System

Symbolic System Emergence

organization

constraint

communication

Agents

interaction

Environment

Tadahiro Taniguchi, Takayuki Nagai, Tomoaki Nakamura, Naoto Iwahashi, Tetsuya Ogata, Hideki Asoh,
Unsupervised Machine Learning for Language Acquisition by a Robot

Without any pre-existing knowledge of phonemes and vocabularies. (like human infants)
Unsupervised Machine Learning for Language Acquisition by a Robot

Problems:
How can the robot come to
1. know a set of phonemes?
2. achieve accurate speech recognition performance?
3. find word segments in speech signals?
4. relate words to objects/events? (meanings / correspondence)

“A. Ringo

“Kore Ga Ringo Dayo”
= “This is an apple”

This is a GODAYA!!!!
The problems are mutually dependent.

- How can the robot come to 1. know a set of phonemes?
- 2. achieve accurate speech recognition performance?
- 3. find word segments in speech signals?
- 4. relate words to objects/events? (meanings / correspondence)

A language learner, i.e., an infant or a robot, has to solve the problems simultaneously.

1. NPB-DAA
   [Taniguchi et al. 2015]
Double articulation structure in semiotic data

- Semiotic time-series data often has double articulation
  - Speech signal is a continuous and high-dimensional time-series.
  - Spoken sentence is considered as a sequence of phonemes.
  - The phonemes are grouped into words, and people give them meanings.

Word

<table>
<thead>
<tr>
<th>How</th>
<th>much</th>
<th>is</th>
<th>this?</th>
</tr>
</thead>
<tbody>
<tr>
<td>[h a u ]</td>
<td>[m ã tʃ]</td>
<td>[i z ]</td>
<td>[ð í s]</td>
</tr>
</tbody>
</table>

Phoneme

| h a u | m ã tʃ | l | z | ð í s |

Speech signal

- Semantic (meaningful)
- Meaningless
- Unsegmented

1. NPB-DAA [Taniguchi et al. 2015]
Nonparametric Bayesian Double Articulation Analyzer (NPB-DAA) \[Taniguchi '15\]

- An integrative generative model (HDP-HLM) that combines language and acoustic models simultaneously.
- The model is applied to continuous artificial Japanese vowel speech signals.
- It outperformed baseline speech recognition system-based method.

\[
\begin{array}{ll}
\text{Method} & \text{Letter ARI} \quad \text{Word ARI} \quad \text{AM} \quad \text{LM} \\
\hline
\text{NPB-DAA (MAP)} & 0.599 \quad 0.497 \\
\text{NPB-DAA} & 0.574 \quad 0.385 & \checkmark \\
\text{Conventional DAA} & 0.584 \quad 0.072 & \checkmark \\
\text{Julius (phoneme dictionary} & 0.483 \quad 0.315 & \checkmark \\
\text{+ NPYLM) } & \\
\text{Julius (phoneme dictionary} & 0.524 \quad 0.426 & \checkmark \\
\text{+ latticelm) } & \\
\text{Julius (monophone} & 0.565 \quad 0.548 & \checkmark \checkmark \\
\text{+ word dictionary) } & \\
\text{Julius (triphone} & 0.516 \quad 0.636 & \checkmark \checkmark \\
\text{+ word dictionary) } & \\
\end{array}
\]

Summary & Open problems

✓ Symbol Emergence in Robotics is a constructive approach towards emergent symbol systems.
✓ Unsupervised machine learning for language acquisition by a robot is now becoming a (partially) solvable problem recently.
✓ Nonparametric Bayesian approach is effective for modeling language acquisition process. (Even though deep learning is booming these days.)

Open problems
- How can a robot learn syntax?
- How can a robot learn pragmatics?
- How can a robot understand metaphor?
- How can a robot use natural language in planning of their behavior?
- Is human-robot communication and collaboration possible using such a language learned in an unsupervised manner?
- and so on.
Information

Special Thanks
• Ritsumeikan University
  • R. Nakashima, S. Nagasaka, A. Taniguchi,
• UEC
  • T. Nagai, T. Nakamura, T. Araki, Y. Ando
• DENSO co.
  • T. Bando, K. Takenaka, K. Hitomi
• Okayama Pref. Univ.
  • N. Iwahashi
• Kyoto University
  • Tetsuo Sawaragi

Visit http://www.tanichu.com/

email: taniguchi@ci.ristumei.ac.jp

Twitter: @tanichu