Socially Beneficial Agents with Biologically Inspired Artificial Intelligence

ARAKAWA, Naoya, PhD

The Whole Brain Architecture Initiative

October 4, 2019, at the AI to Bridge the Borders Conference

Who am I?
Why am I here?

❖ **I have been working around AI.**

❖ **I am an ex-colleague of a Finnish researcher *Kristiina Jokinen*, who works in Tokyo.**

❖ My Backgrounds
  - ➢ Biophysical Engineering
  - ➢ Systems Engineering
  - ➢ **Natural Language Processing**
  - ➢ **Philosophy of Language (Ph.D)**
  - ➢ **Cognitive Science in general**

❖ AI, Agents, and Communication

❖ Creating AI by mimicking the brain

❖ *The Whole Brain Architecture Initiative*, an NPO

- AI is booming, and here we are 😃!

- It's the 3rd Boom 💥

  = *Machine Learning* Boom

- Previously knowledge for AI was hand-made

- This time, machines learn!

- Machines learn to recognize patterns: "Pattern Recognition"

- Now machines can see 👀 and hear 👂!

- Many promising applications ahead (Boom)!

- Pattern recognition is only part of AI.
- AI should be able to what humans can do, e.g.:
  - Manipulate things
  - Communicate with humans
  - In brief, work for us

  - Or, be a socially beneficial 'robot'…

  irasutoya.com

- Can the current AI do those things?

  - Not quite.
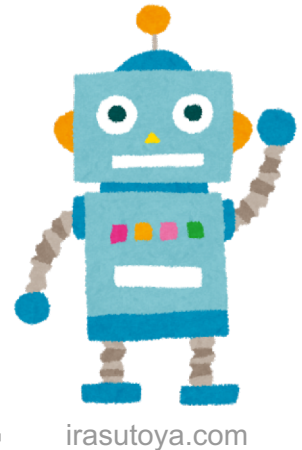
- Humans have "*General Intelligence*" in the sense:
  - we are able to cope with novel situations in ingenious ways.
- We may want AI to have such capability
  - to have them in our workplaces.
- Thus, we would like to have Artificial *General Intelligence*, or AGI.

- We haven't got it.

- Current AI=Machine Learning ≠ AGI

- They learn specific tasks,
  i.e., not general (narrow AI)

- But ML is part of AGI.

  - So the situation is better for AGI R&D.

- AI top researchers are well aware of AGI!
  e.g., @DeepMind, OpenAI, etc.

- We'll get AGI or human-like AI
  in 10 or 20 years?

- Agents perform tasks in the environment.
- Physical (embodied) agents ～ robots
- We'd like them to communicate with us to collaborate!
- In language and in other ways…
- Human-Agent communication
  - long studied topic…
- Communication: to **bridge the border**

Changing to the second topic:

❖ **Creating AI by mimicking the brain**

～ what I currently promote...

Looking back the previous topic:

- We'd like to have ***human-like*** agents. For:
   - the general intelligence capability

    - the ease of communication

        ⇒ to make them *socially beneficial*!
- So, it's natural to think about mimicking

    the human ***brain***…

Brain-*inspired* A(G)I research is pursued, e.g.:

- **Hassabis**, et al. **@DeepMind**
  "**Neuroscience-Inspired Artificial Intelligence**",
  Neuron (2017)

- **Cognitive Computational Neuroscience**
  Conferences (2018-)

- Recent **Brain Informatics** Conferences

- A Cold Spring Harbor Conference (2020)
  "From Neuroscience to Artificially Intelligent Systems"
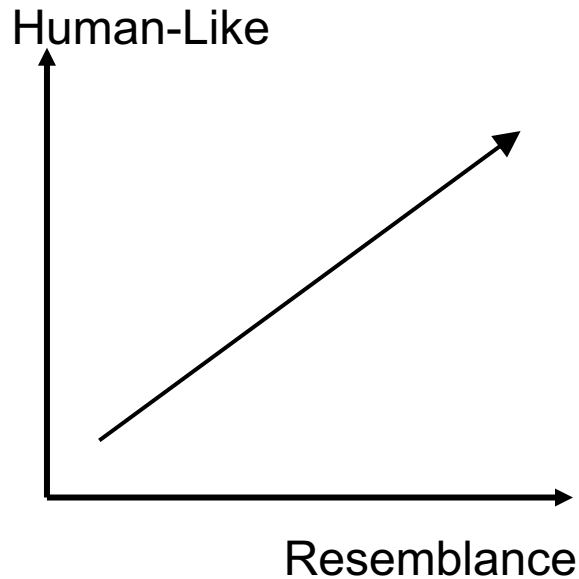
- ...

Human-Like

Resemblance
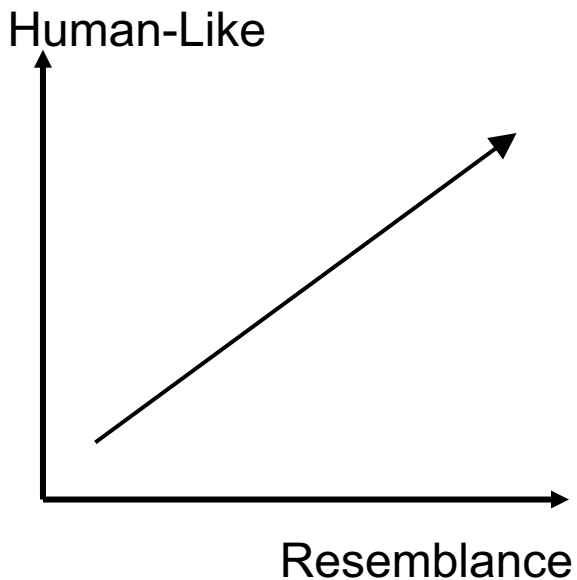
The more the system resembles the human brain, the more it gets human-like.

But

Simplicity

?

Resemblance
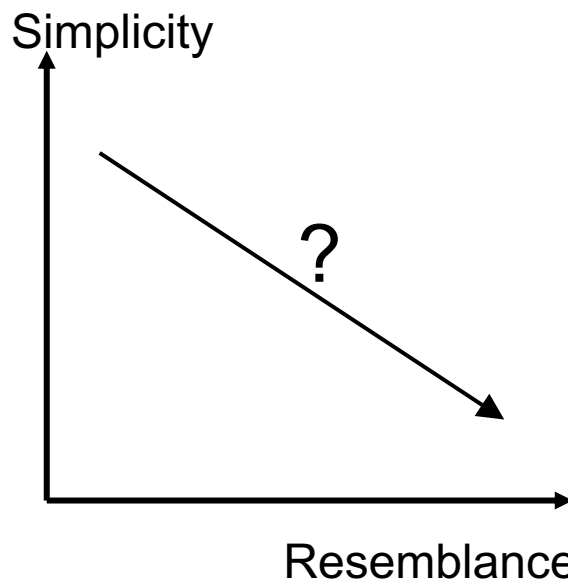
Many researchers may opt for just being inspired than mimicking.

Brain is too complex?

Human-Like

Resemblance

Human-Like
- easier to understand each other (?)
- easier to communicate (?)
  - similar 'constitution'
  ⇒ common background

Other motivation
- referring to the brain as common ground for collaborative development
  - to be mentioned later

❖ (Some) humans are **bad**, *really* **bad**…

  ➢ Through simulation, we can understand
    the badness or the human nature in general.
  ➢ Then, we can control the badness in AI...

❖ Humans can be mentally sick.

  ➢ Through simulation, we can understand mental
    disorders.
  ➢ Then, it may serve for cure…

❖ Besides, machine learning is generally black boxes.

  ➢ Simulated humans may cast light into the boxes.
    ⇒ *eXplainable AI (XAI)*

So, hopefully the approach will be socially beneficial!

Changing to the third & last topic:

❖ *The Whole Brain Architecture Initiative*, the NPO

～ where I currently work...

Whole Brain Architecture Initiative

An NPO established in 2015

Vision: to create a world in which AI exists in harmony with humanity (being *Beneficial*)

Mission: to promote the *open development* of human-like AGI by studying *the entire architecture of the brain (i.e., Whole Brain Architecture)*

through
- Education
- R&D

Education

❖ Seminars for local general audience on neuroscience, AI, & Cog. Sci.

❖ Hackathons - events for engineers

This year: we promote the ***Animal-AI Olympics***, ongoing AI software competition with agents in virtual environments.

http://animalaiolympics.com/

Image Credit Squidoodle

- Organized by people at Leverhulme Centre for the Future of Intelligence (CFI).
- Agents' cognitive capabilities to be tested are inspired by animals.
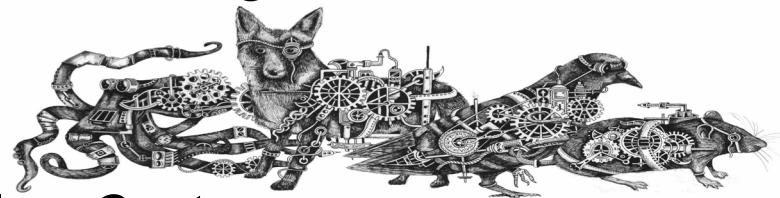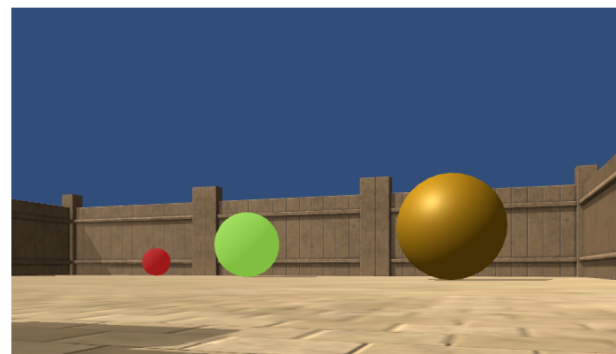- Prizes: total $32,000 WBA Prize ($4,000) for the most biologically plausible entry.

Basically, we do not conduct R&D, but promote/support R&D in the whole brain architecture approach.

R&D areas include:
- Software platforms
- Modeling (parts of) the brain
- ...

Current R&D Focus:

- To create & maintain

  the Whole Brain **Reference** Architecture

  - Circuit diagram of the brain
  - ***Design chart*** for brain-inspired AGI
  - to be a ***common ground*** for researchers

    N.B. human brain architecture is unique!
  - to ***bridge*** neuroscientists and engineers!
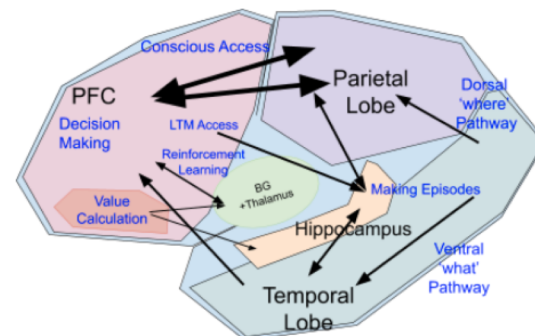
Bridging borders between cultures...

Currently:

- Partnership with Project AGI
  (Australia-based research entity)
- Collaborating with the Animal-AI Olympics

If you are interested, please check our HP
& contact us! https://wba-initiative.org/en/

Kiitos huomiostanne!

ご清聴ありがとうございました。