The Enactive Pathway to AI

Rafik Hadfi
Department of Social Informatics, Kyoto University
rafik.hadfi@i.kyoto-u.ac.jp

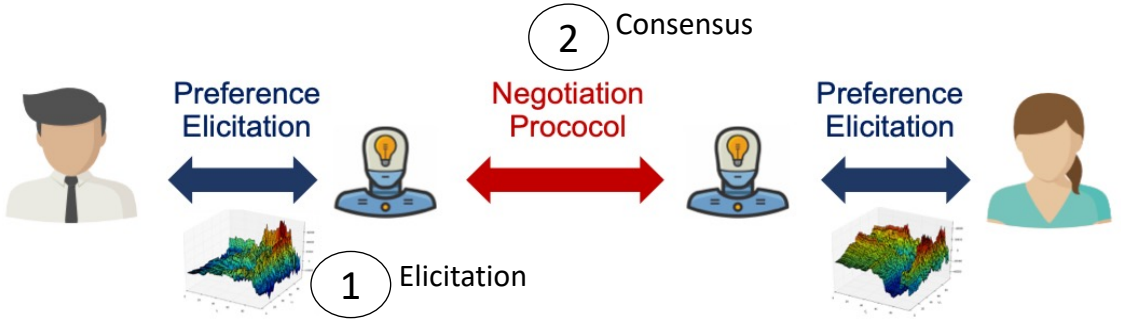
14 February 2023

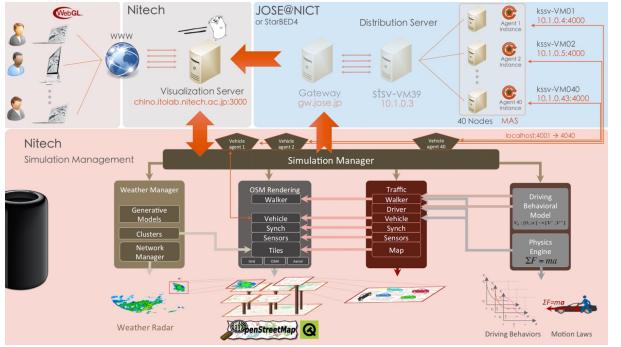
Background

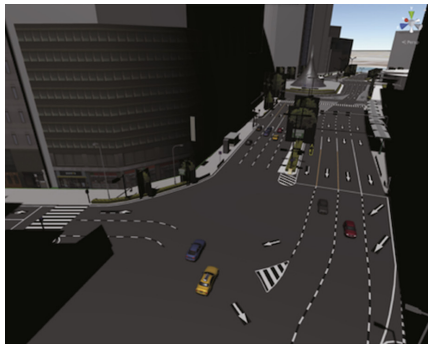
Feb 2023 –	**Associate Professor**, Kyoto University Department of Social Informatics
2020 – 2023	**Assistant Professor**, Kyoto University Department of Social Informatics
2019 – 2020	**Assistant Professor**, Nagoya Institute of Technology Department of Computer Science and Engineering
2017 – 2019	**Research Fellow**, Monash University School of Psychological Sciences / Philosophy Department
2015 – 2017	**Postdoctoral Researcher**, Nagoya Institute of Technology Department of Computer Science and Engineering
March 2015	**Doctor of Engineering**, Nagoya Institute of Technology
March 2012	**Master of Engineering**, Nagoya Institute of Technology

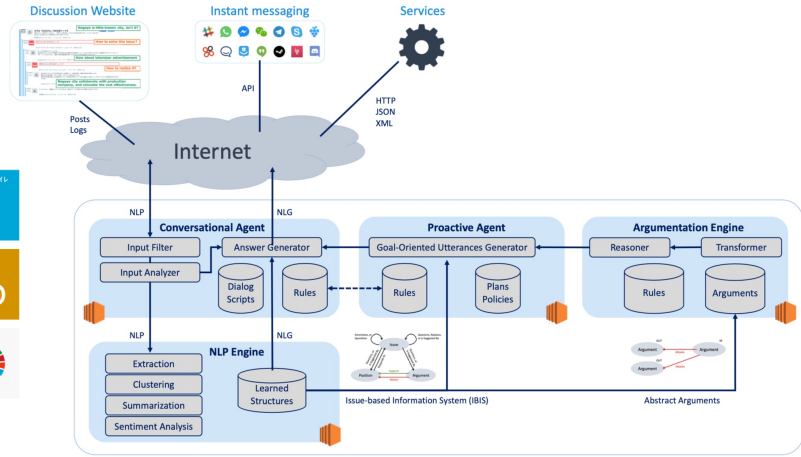
Research Interests

- **AI for Decision-making**
  - Automated Negotiation
- **Multiagent Social Simulation**
  - Mobility, Driving, Traffic, etc.
- **Conversational AI for Social Good**
  - Online Deliberation, Polarization, etc.









The Enactive Pathway to AI

Rafik Hadfi
Department of Social Informatics, Kyoto University
rafik.hadfi@i.kyoto-u.ac.jp

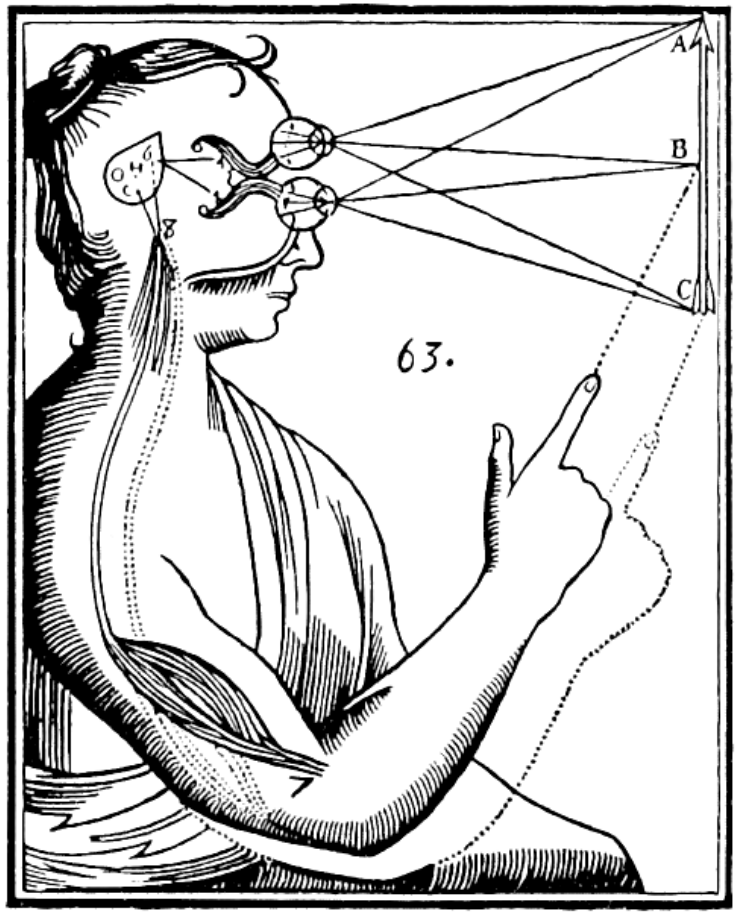
14 February 2023

Summary

- Mind-body Dualism
- The Embodied Alternative
- Enactive Cognition
- Enactive AI

Mind-body dualism

- Philosophy, neuroscience, physics … and AI?
- The Cartesian philosophy of mind
- The nature of the mind is different from that of the body, and so, it is possible for one to exist without the other



Descartes' mind-body dualism

Mind-body dualism in perception

The tree, as perceived by an external observer

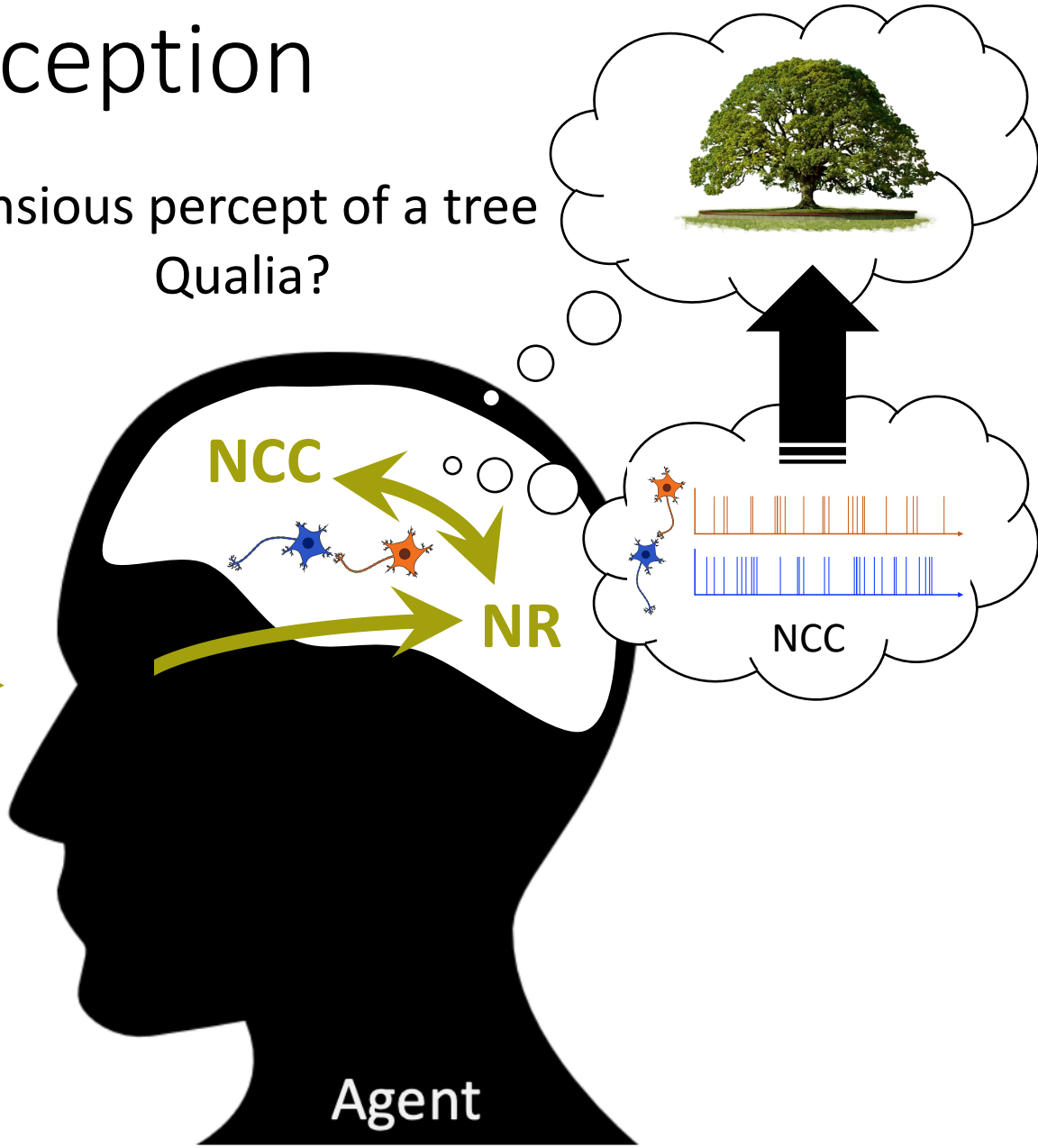
Light rays

Mind-body dualism in perception

The tree, as perceived by an external observer

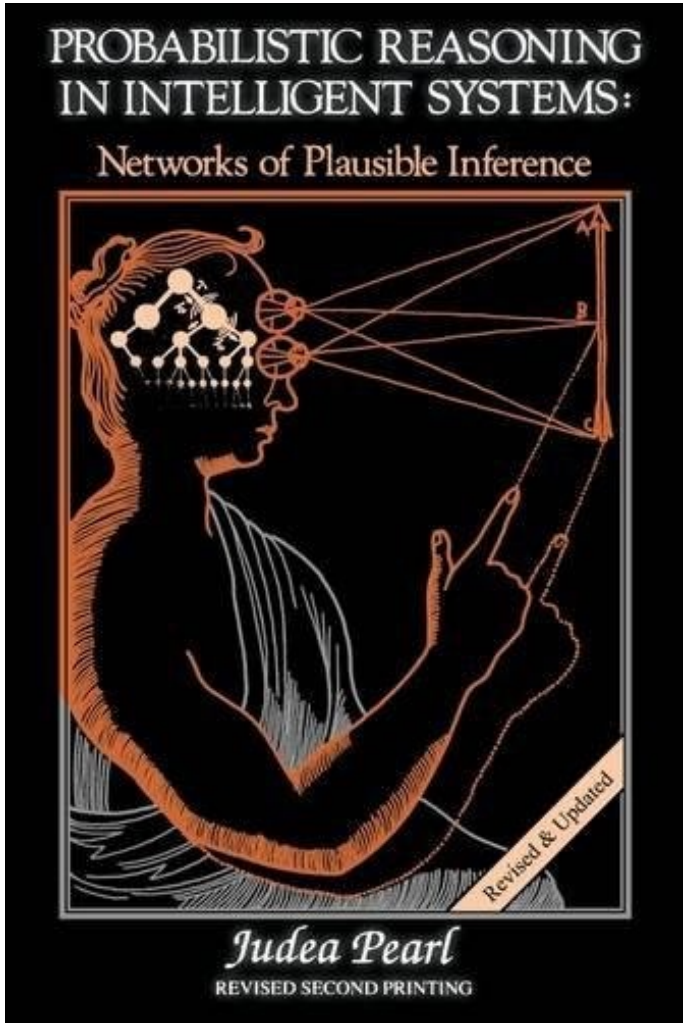
Light rays

Consious percept of a tree
Qualia?



Mind-body dualism in perception & action

- The Bayesian brain
- Mental states are treated as **hypotheses** updated based on **evidence** from the physical world
- Active Inference
  - The Free-energy Principle



Mind-body dualism in perception & action

Perception

Bayesian Inference about the world

External Hidden states about the world $\boldsymbol{p(y)}$

Free energy $\boldsymbol{F(y,\zeta) \geq surprise}$

$\boldsymbol{argmin_{\zeta}\, F(y,\zeta)}$

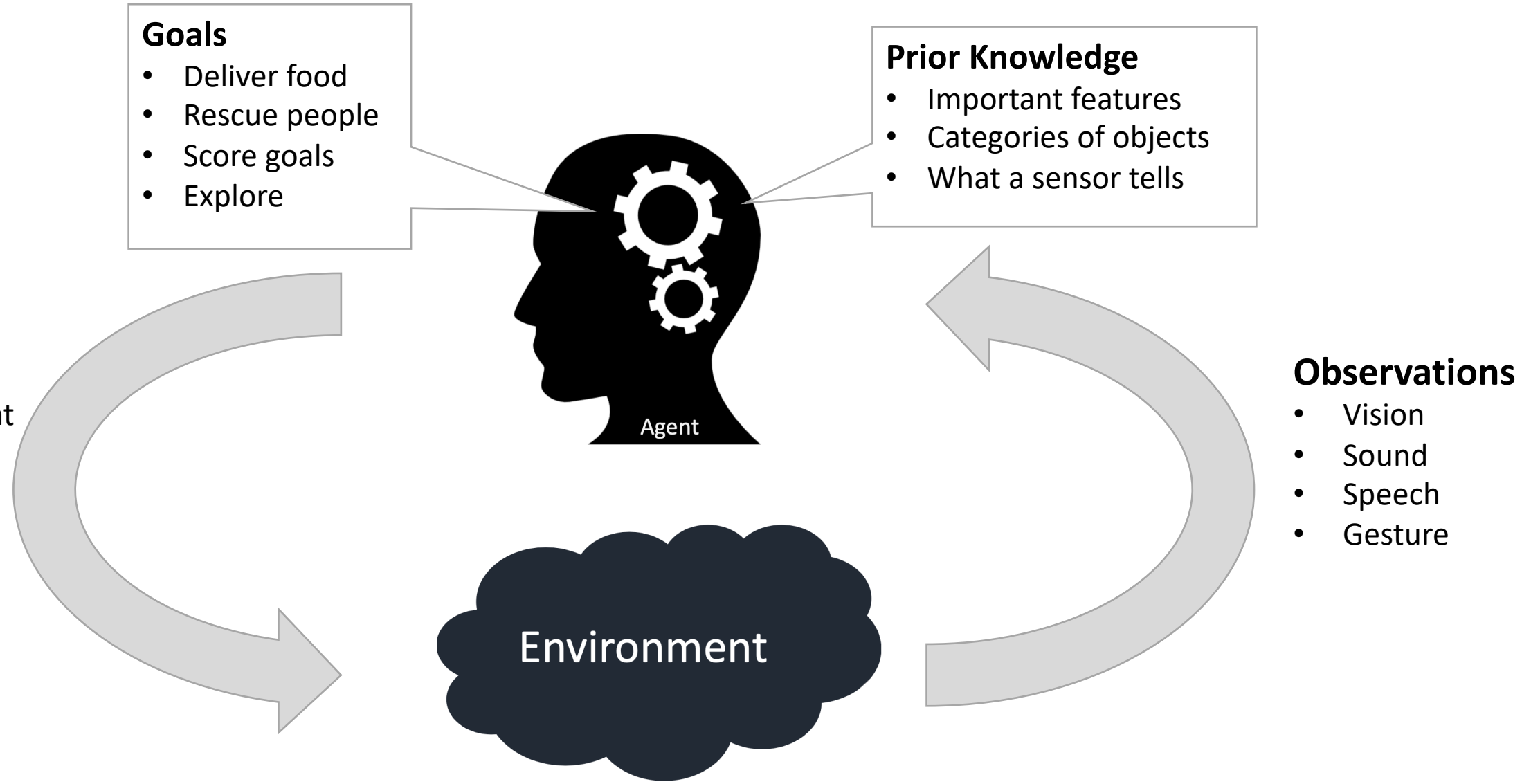
How the brain models the environmental states by aligning **internal** to **external** states through free energy minimization

Action

Acting on the world to avoid surprises

Friston, et al. "Active inference: a process theory." Neural computation 29.1 (2017): 1-49.

Mind-body dualism in AI?



Mind-body dualism rejection

- Cartesian dualism is rooted in a misunderstanding of the activities of the mind

- Perceiving the mind as an embedded non-bodily entity in the body is admitting that there are two separate substances, one is abstract, and the other one is concrete (Gilbert Ryle, 1949)

- Ryle's **Ghost in the Machine** dogma

The Embodied Alternative

- Cognition is not something that happens internally but also involves a continuous **interaction** between **mind**, **body** and **environment**

The Embodied Alternative

- Cognition is not something that happens internally but also involves a continuous **interaction** between **mind**, **body** and **environment**

- All the concepts that humans have are based on bodily experiences

The Embodied Alternative

- Cognition is not something that happens internally but also involves a continuous **interaction** between **mind**, **body** and **environment**

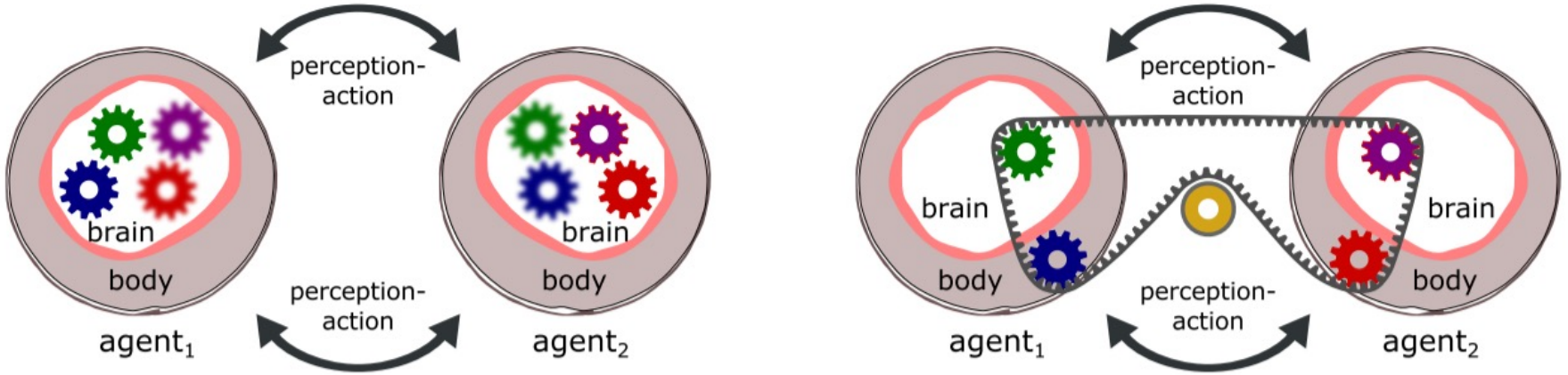
- All the concepts that humans have are based on bodily experiences

- Concepts used to think and express thoughts are shaped by the gestural features of the body
  - When we refer to comprehending something in all its aspects, we use the verb **grasp**, which also means holding it tightly

The Embodied Alternative

- Cognition is not something that happens internally but also involves a continuous **interaction** between **mind**, **body** and **environment**

- And by extension, social cognition too



Abramova et al. "Enactive mechanistic explanation of social cognition." (2017)

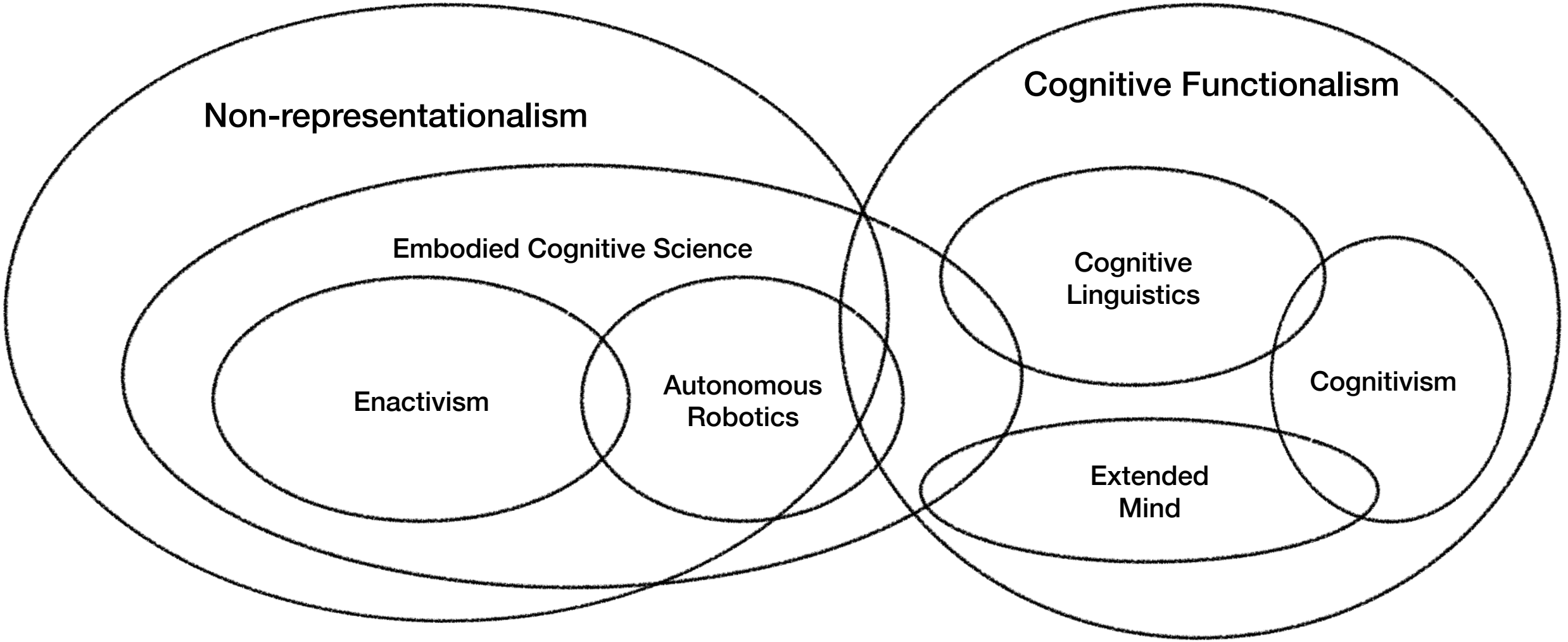
Enactive (4E) Cognition

- Cognition is
  - **Embodied:** involves bodily structures and processes, too
  - **Embedded:** functioning only in a related external environment
  - **Enacted:** involving not only neural processes, but (en)actions
  - **Extended:** into the organism's environment

Enactive Cognition

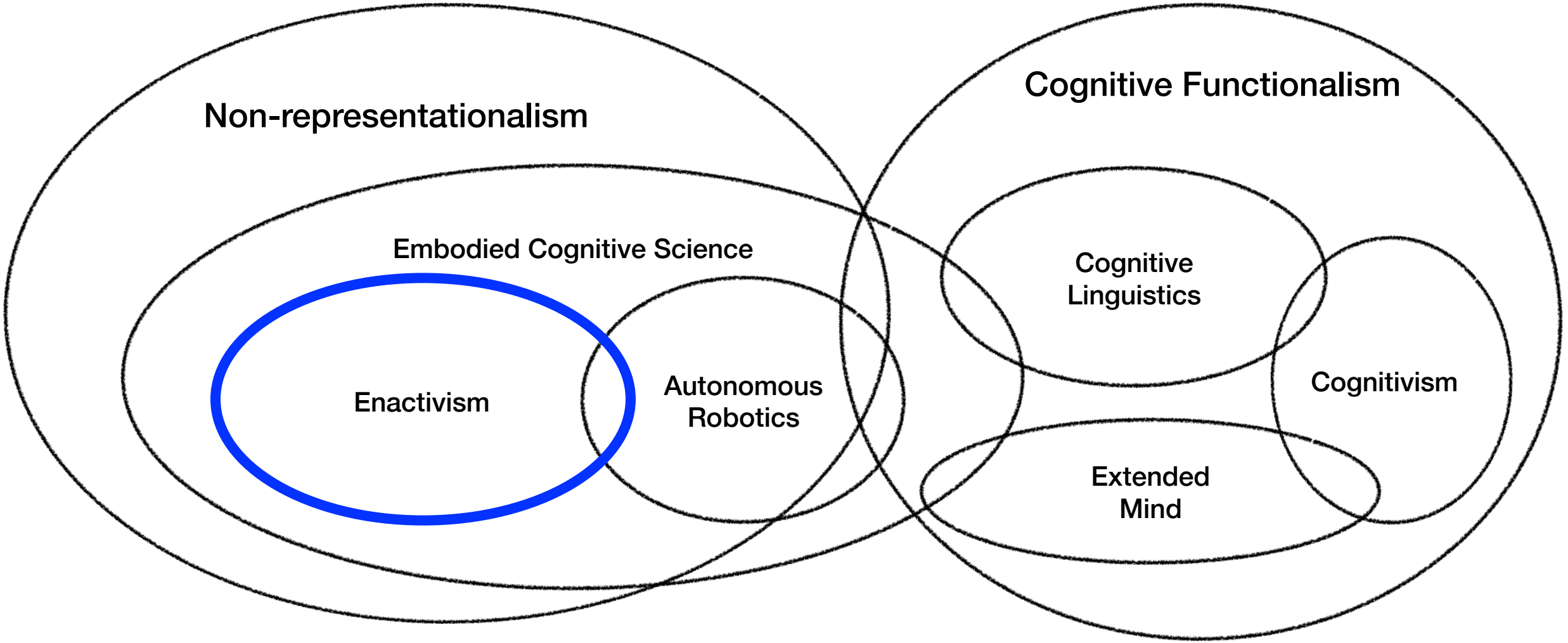
- Cognition emerges from sensorimotor activity
  - **Autopoietic:** conceives of cognition in terms of the **biodynamics** of living systems (Varela & Maturana 1980)
  - **Sensorimotor:** focuses on explaining the **intentionality** and **phenomenology** of perceptual experiences (O'Regan 2001)
  - **Radical:** replaces all **representational** explanations of cognition with embodied, **interactive** explanations (Hutto 2013)

In Contemporary Cognitive Sciences



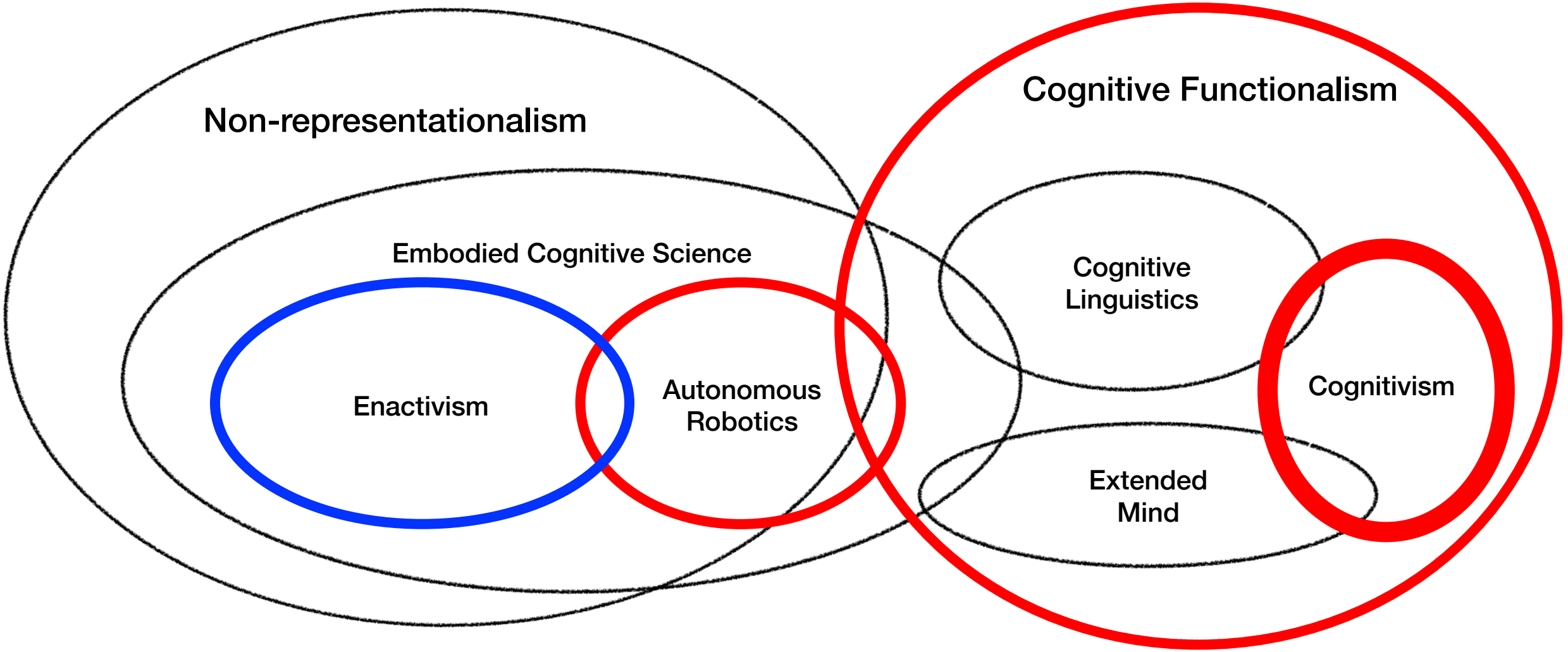
Cuffari, et al. "From participatory sense-making to language: there and back again." Phenomenology and the Cognitive Sciences 14 (2015)

In Contemporary Cognitive Sciences



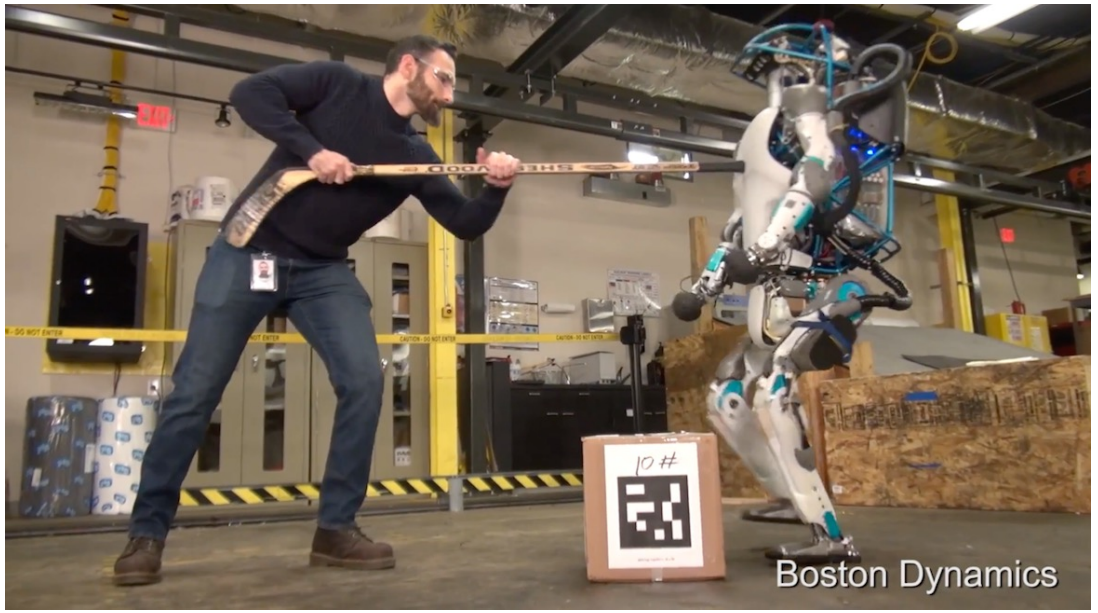
Cuffari, et al. "From participatory sense-making to language: there and back again." Phenomenology and the Cognitive Sciences 14 (2015)

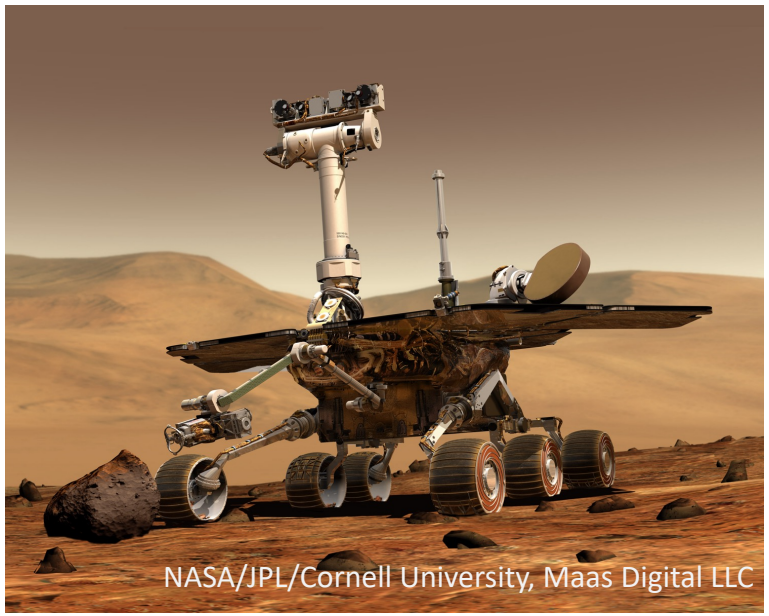
In Contemporary Cognitive Sciences, and AI?



Cuffari, et al. "From participatory sense-making to language: there and back again." Phenomenology and the Cognitive Sciences 14 (2015)

Enactive AI: the motivation?

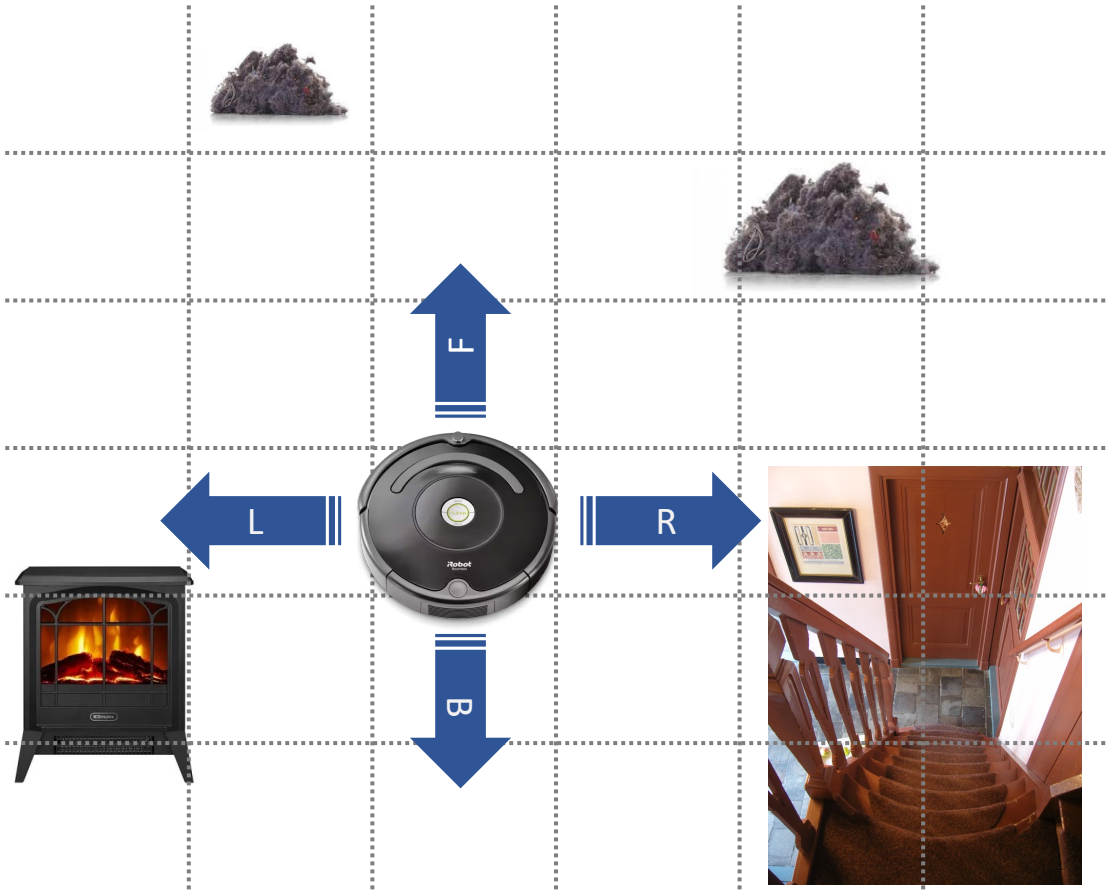




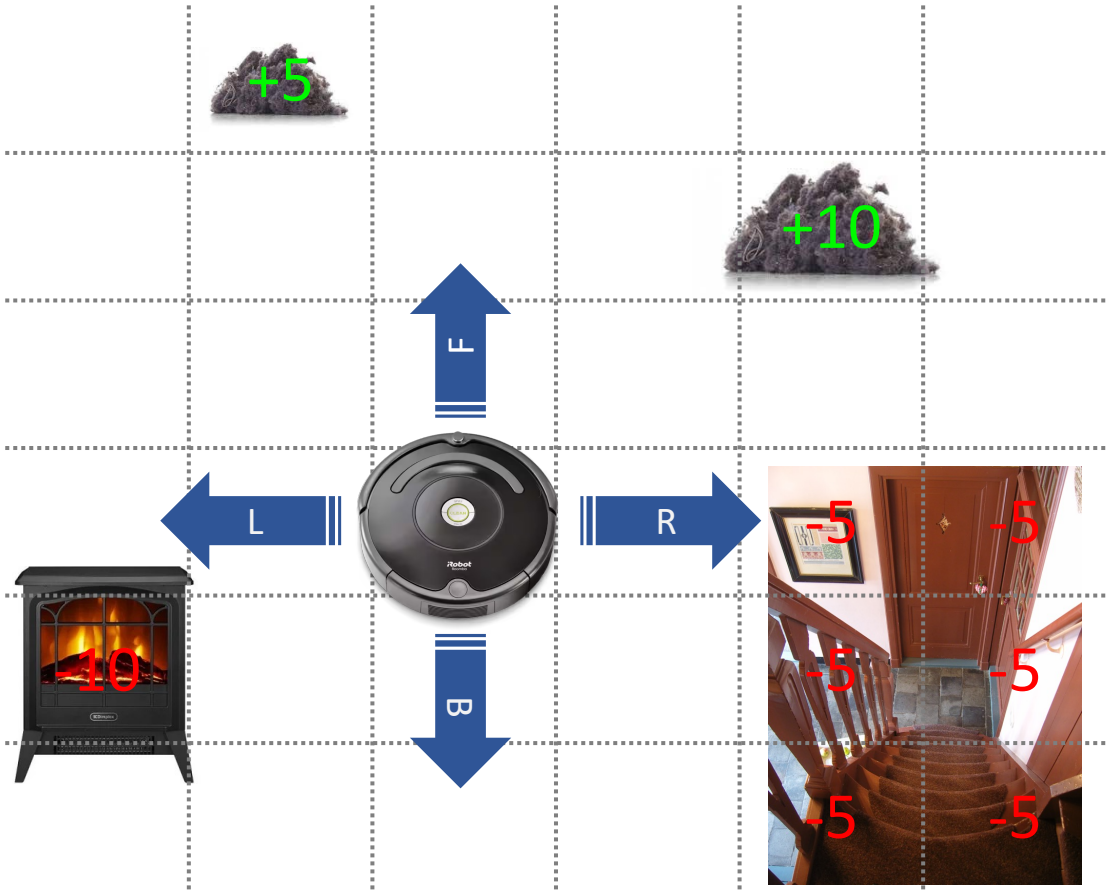
Enactive AI: the motivation?

Enactive AI: the motivation?

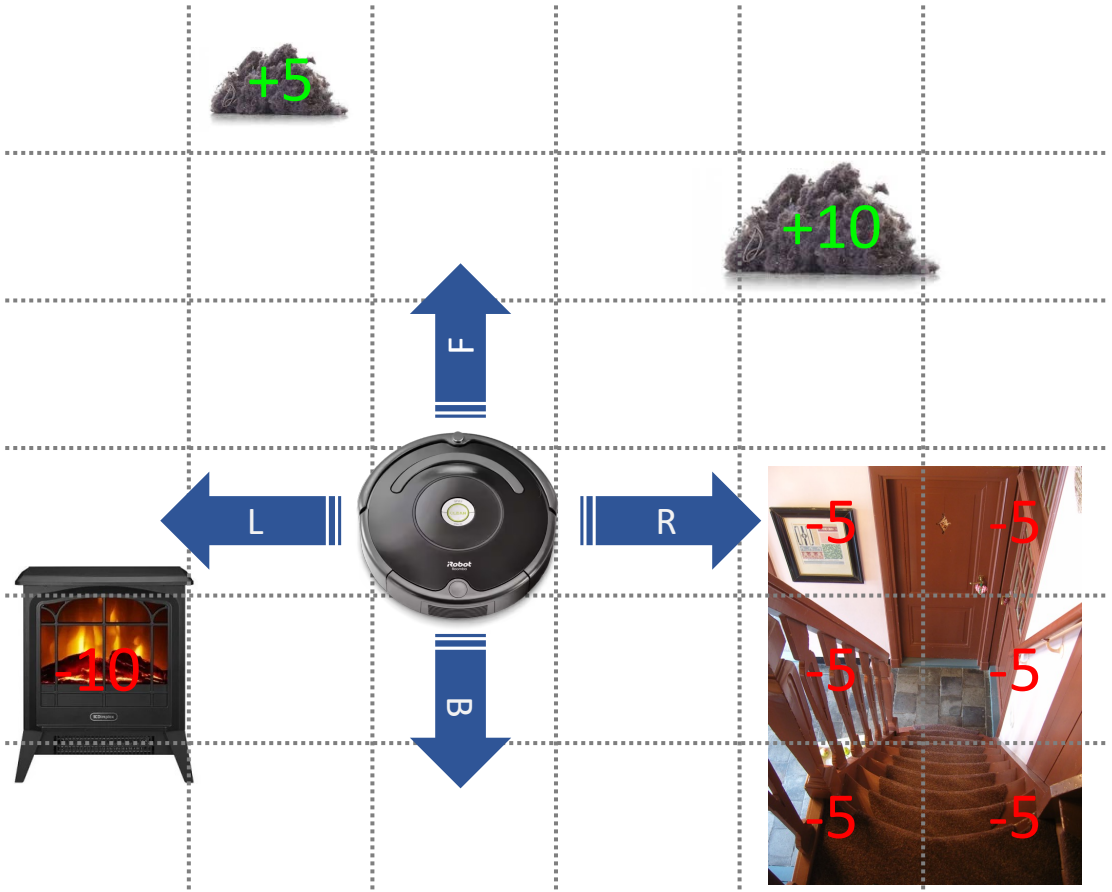
Enactive AI: the motivation?



Enactive AI: the motivation?



Enactive AI: the motivation?

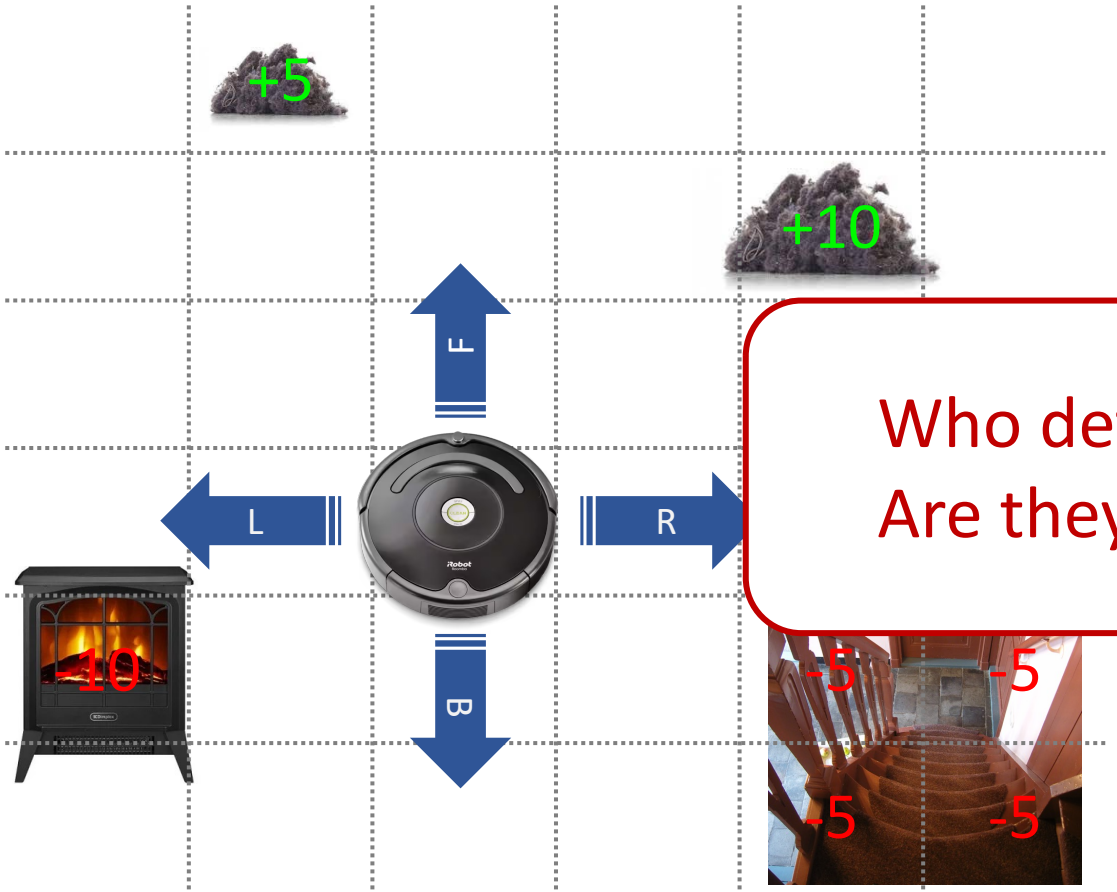


States $s \in [1, 6]^2$

Actions $a \in \{L,\ B,\ R,\ F\}$

Rewards $r \in [-10, 10]$

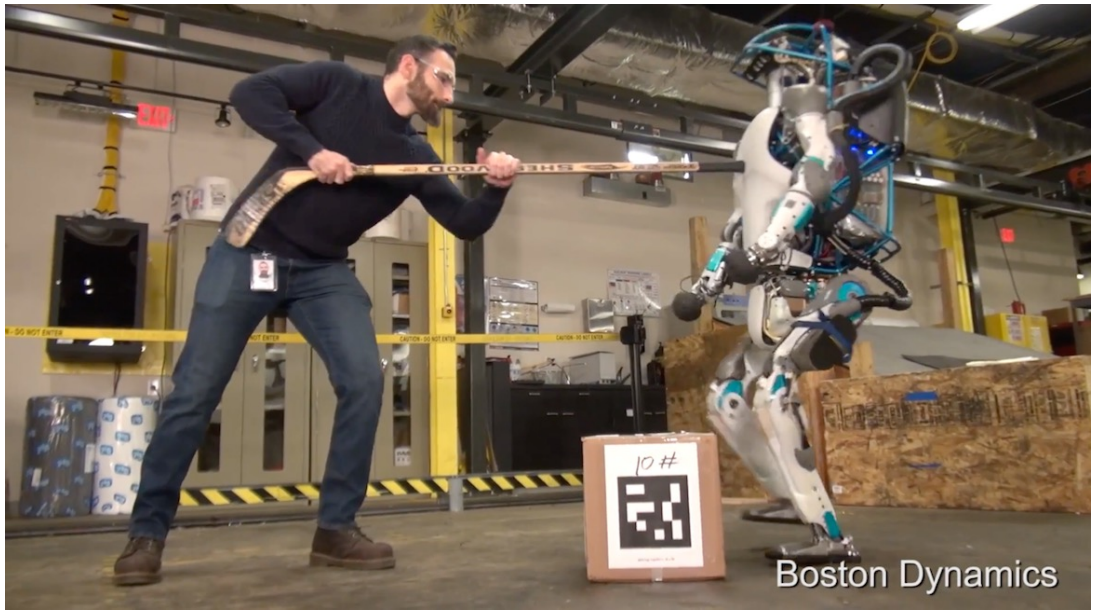
Enactive AI: the motivation?

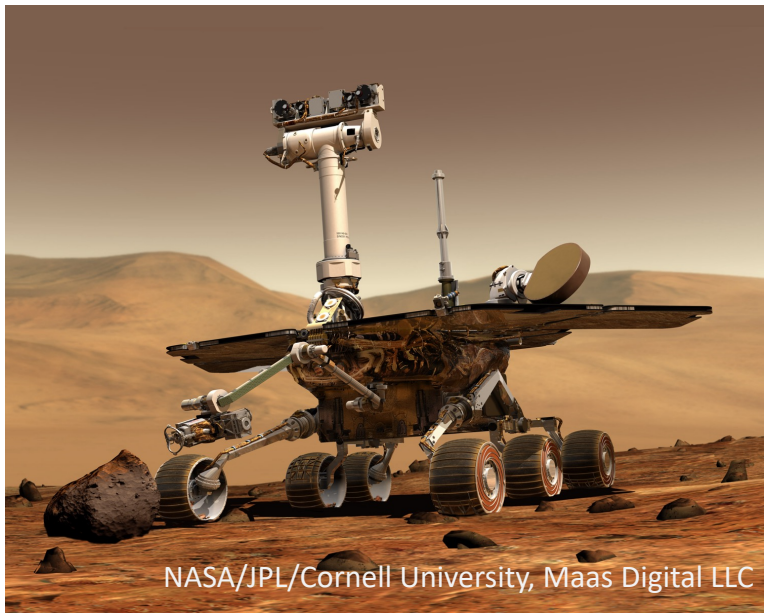


States $s \in [1,6]^2$
Actions $a \in \{L,\ B,\ R,\ F\}$
Rewards $r \in [-10, 10]$

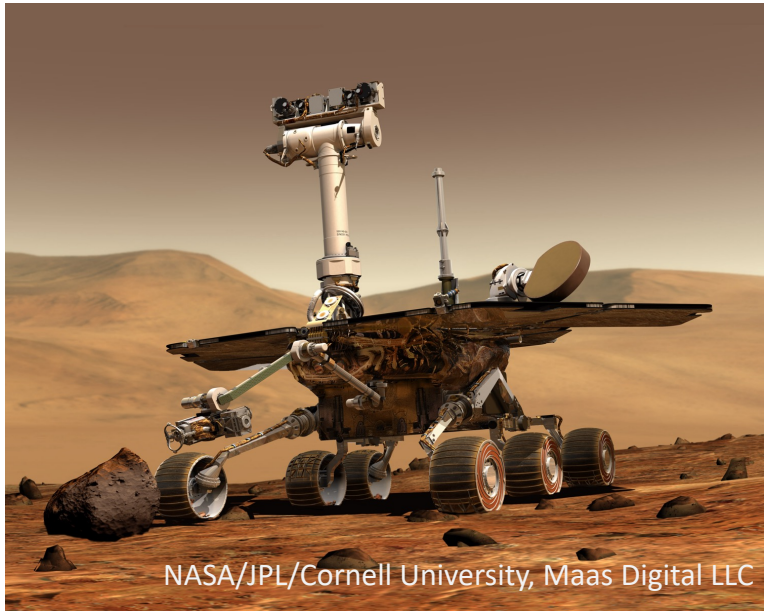
Who defines the states, actions, and rewards?
Are they always **known** and/or **knowable**?

Enactive AI: the motivation?





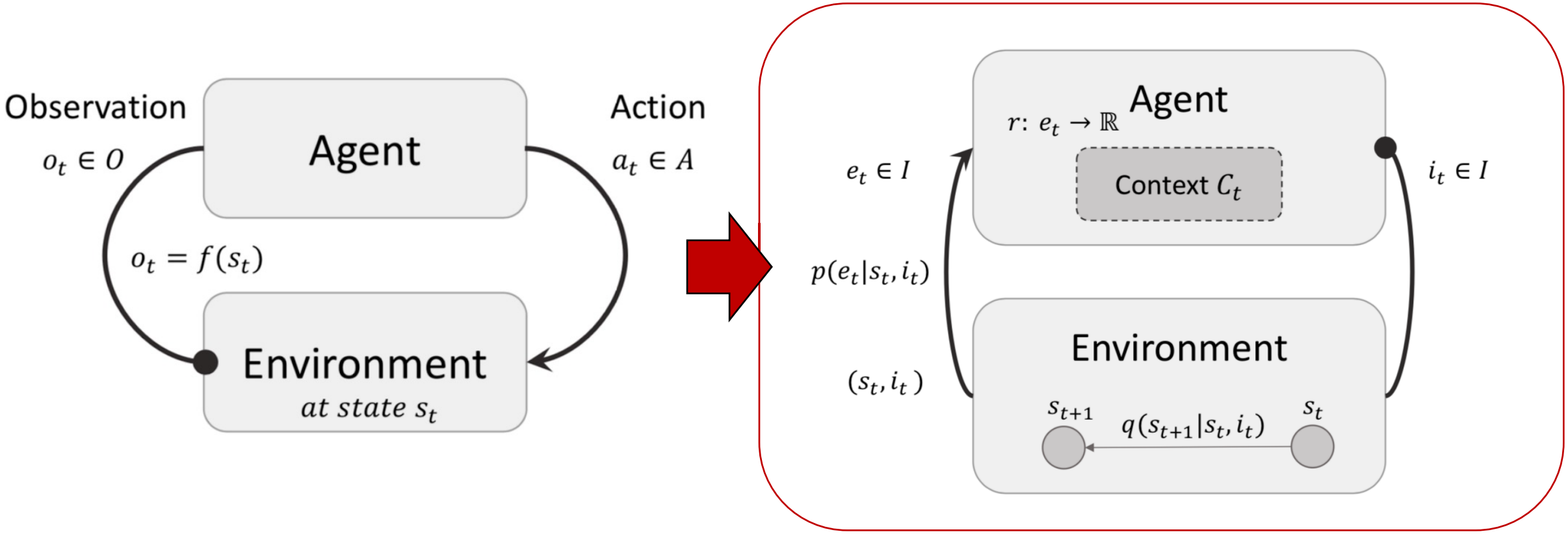
Enactive AI: t



An artistic rendition of the Kepler-35 planetary system, in which a Saturn-size planet orbits a pair of stars (*nasa.gov)*

Enactive AI: Reinforcement Learning

- Reinforcement Learning and its enactive formulation



Hadfi et al. "Investigating Enactive Learning for Autonomous Intelligent Agents." arXiv preprint arXiv:1810.04535 (2018).

Enactive AI: The enactive agent

- No prescribed model of the world
  - The agent **constructs** its own model from “nothing”
- No prescribed **re**-presentation of the world
  - The agent experiences the world as is, an impermanent act of discovery
- No prescribed problems
  - The agent is not designed to solve a specific problem
- No prescribed “utilitarian” goals
  - The agent establishes its own **intrinsic** goals
- No division between the agent and its environment
  - It is all about the **boundary**,… and agency, with the observer?

Enactive AI: what next?

- Vis-à-vis the current trends in ML, RL, and robotics
  - Requirements and practicality?
  - The untapped potential of robotics?

- Challenges
  - Autopoietic systems
  - Sentient AI

Thank you!

Rafik Hadfi
Department of Social Informatics, Kyoto University
rafik.hadfi@i.kyoto-u.ac.jp

14 February 2023