Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

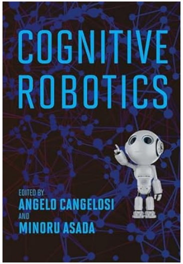
Institute for AI and
Autonomous Systems

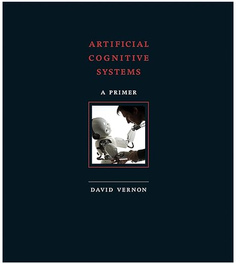
Cognitive Robotics

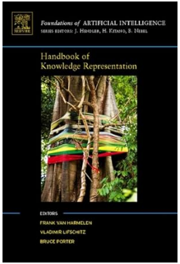
Introduction

Dr. Alex Mitrevski
Master of Autonomous Systems

Structure



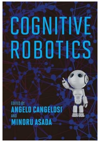




- Cognitive robotics in general terms
- Definitions
- A brief look at the human brain
- Cognitive agent architecture example: Soar
- Cognitive robot learning

Cognitive Robotics in General Terms

Cognitive Robotics



- Different perspectives can be found in the literature
  - Some researchers focus more on modelling various cognitive functionalities, others on concrete bio-inspired aspects, and yet others on the integration of cognitive functionalities into complete systems
- Some important underlying principles are that:
  - **robots are typically considered to be embodied physical agents**
  - the overall objective is to develop robots that have **perceptual, sensorimotor, and cognitive capabilities that are similar to, or inspired from, humans and other animals**

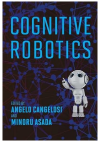
"Cognitive robotics is the field that combines insights and methods from AI, as well as cognitive and biological sciences, to robotics." (Cangelosi and Asada 2022, p. 4)

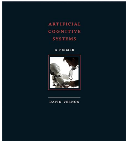
Cognitive Science

- A general discipline that is concerned with **understanding the working of the brain and its use for intelligent behaviour**
- The underlying hypothesis is that **thinking involves creating structures that are then used in computational processes**
- **Many different approaches to represent knowledge and define "thinking" procedures** (such as logic, analogical reasoning, connectionist approaches, Bayesian reasoning, deep learning)

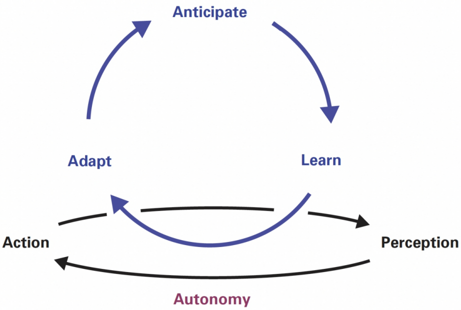
"Cognitive science is the interdisciplinary study of mind and intelligence." (Stanford Encyclopedia of Philosophy, rev. 31.01.2023)

Attributes of Cognition

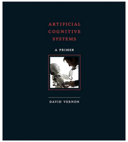




- **Autonomy**: Acting without external assistance
- **Perception**: Interpreting the environment from sensory observations
- **Action**: Affecting the world by taking into account prior knowledge (memories) and current observations of the environment
- **Anticipation**: Predicting the effects of actions
- **Learning**: Using information about the observed effects of actions for updating the knowledge about actions
- **Adaptation**: Modifying the behaviour as a result of updated knowledge



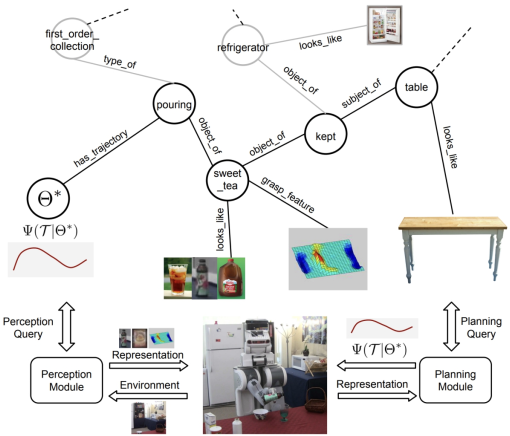
Capabilities of a Cognitive Robot



- Creating meaningful **representations of the world (world models)**
- **Remembering** relevant aspects about the world
- **Reasoning** about the world, based on **all the available information** and **taking uncertainty into account**
- Understanding or forming **goals** and identifying ways of accomplishing them
- **Predicting the effects of actions** (of the robot itself and of other agents)
- **Anticipating** the intentions and actions of other agents
- **Learning from experience** (with respect to which actions to take)
- **Recognising unexpected events** and **reacting** to them accordingly (including taking corrective actions)
- **Adapting** to changes

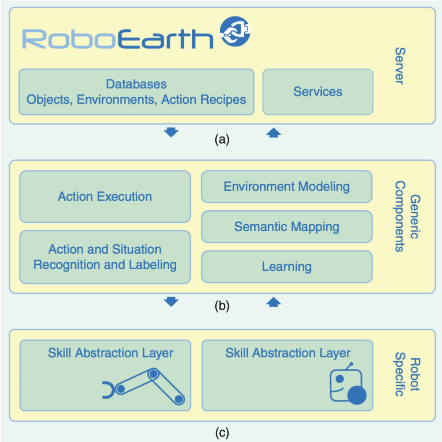
Examples of Cognitive Robot Systems

RoboBrain



A. Saxena et al., "RoboBrain: Large-Scale Knowledge Engine for Robots," *CoRR*, vol. abs/1412.0691, 2014. Available: https://arxiv.org/abs/1412.0691

RoboEarth



M. Waibel et al., "RoboEarth," in *IEEE Robotics & Automation Magazine*, vol. 18, no. 2, pp. 69–82, June 2011. Available: https://doi.org/10.1109/MRA.2011.941632

Why is Cognitive Robotics Important?

- Cognitive robotics is concerned with endowing agents with an ability to act autonomously, but also to collaborate with other agents
  - **Very relevant for human-centred environments** — a robot does not exist on its own island

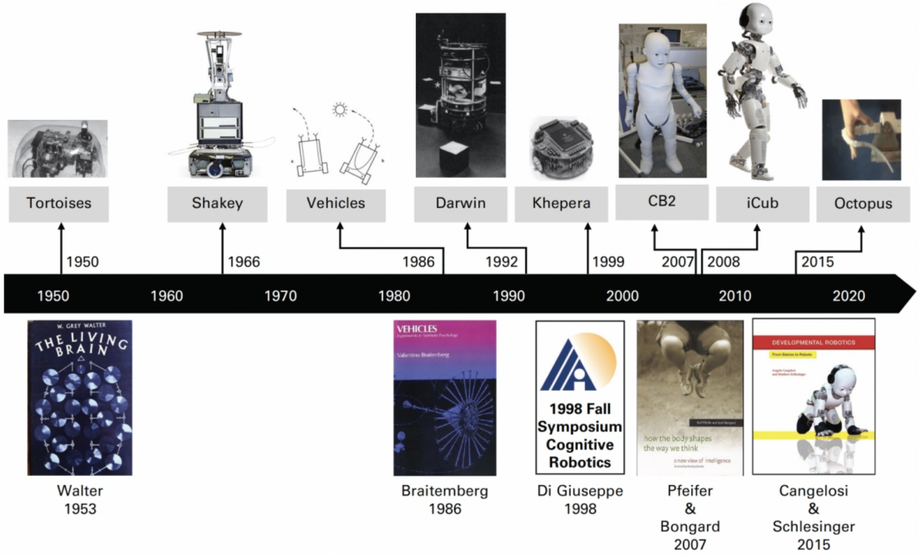
Why is Cognitive Robotics Important?

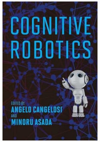
- Cognitive robotics is concerned with endowing agents with an ability to act autonomously, but also to collaborate with other agents
  - **Very relevant for human-centred environments** — a robot does not exist on its own island
- As a side goal, cognitive robotics usually pursues **more transparent robots** whose operation can be ascribed to well-understood principles
  - This is in contrast to systems that mimic intelligence using black-box systems, particularly end-to-end systems that are inherently non-transparent

Why is Cognitive Robotics Important?

- Cognitive robotics is concerned with endowing agents with an ability to act autonomously, but also to collaborate with other agents
  - **Very relevant for human-centred environments** — a robot does not exist on its own island
- As a side goal, cognitive robotics usually pursues **more transparent robots** whose operation can be ascribed to well-understood principles
  - This is in contrast to systems that mimic intelligence using black-box systems, particularly end-to-end systems that are inherently non-transparent
- In principle, **we can learn more about our own intelligence** by studying cognitive robotics
  - In principle, robots can be used to test hypotheses about human or animal intelligence
  - Even if an approach is not meant to be biologically accurate, we can use insights from our understanding of biological intelligence to guide the operation of a robotic system

Cognitive Robotics Timeline





Are ChatGPT-Based Robots Cognitive?

- Creating meaningful representations of the world
  - unknown - its internal structure isn't really prone to analysis and interpretation

Are ChatGPT-Based Robots Cognitive?

- Creating meaningful representations of the world
  - unknown - its internal structure isn't really prone to analysis and interpretation
- Remembering relevant aspects about the world
  - ✓ — large language models are trained on enormous amounts of data, so they — by all means — have a large memory

Are ChatGPT-Based Robots Cognitive?

- Creating meaningful representations of the world
  - unknown - its internal structure isn't really prone to analysis and interpretation
- Remembering relevant aspects about the world
  - ✓ — large language models are trained on enormous amounts of data, so they — by all means — have a large memory
- Reasoning about the world, based on
  - all the available information (✓) and
  - taking uncertainty into account (✗ - large language models have no notion of uncertainty about their output)
  - although it is questionable whether we can refer to the processing of GPT models as "reasoning"

Are ChatGPT-Based Robots Cognitive?

- Creating meaningful representations of the world
  - unknown - its internal structure isn't really prone to analysis and interpretation
- Remembering relevant aspects about the world
  - ✓ — large language models are trained on enormous amounts of data, so they — by all means — have a large memory
- Reasoning about the world, based on
  - all the available information (✓) and
  - taking uncertainty into account (✗ - large language models have no notion of uncertainty about their output)
  - although it is questionable whether we can refer to the processing of GPT models as "reasoning"
- Predicting the effects of actions and anticipating the intentions of other agents
  - ✗ — ChatGPT is a language model, so it can neither predict the effects of its actions nor does it know that other agents exist or might be acting as well

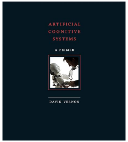
Are ChatGPT-Based Robots Cognitive?

- Creating meaningful representations of the world
  - unknown - its internal structure isn't really prone to analysis and interpretation
- Remembering relevant aspects about the world
  - ✓ — large language models are trained on enormous amounts of data, so they — by all means — have a large memory
- Reasoning about the world, based on
  - all the available information (✓) and
  - taking uncertainty into account (✗ - large language models have no notion of uncertainty about their output)
  - although it is questionable whether we can refer to the processing of GPT models as "reasoning"
- Predicting the effects of actions and anticipating the intentions of other agents
  - ✗ — ChatGPT is a language model, so it can neither predict the effects of its actions nor does it know that other agents exist or might be acting as well

In general, we cannot consider a ChatGPT-based robot to be cognitive directly, without significant additional machinery that would make it so

Definitions

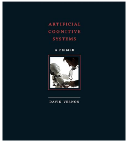
Cognition



- There is no single accepted definition of the term[1], but there are some common threads

[1] T. Bayne et al. "What is cognition?," Current Biology, vol. 29, no. 13, pp. R608-R615, 2019. Available: https://www.cell.com/current-biology/pdf/S0960-9822(19)30614-1.pdf

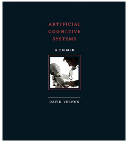
Cognition



- There is no single accepted definition of the term[1], but there are some common threads
- Cognition has to do with **representing and processing information** — obtained from perceptual observations and other learned experiences — **in order to infer suitable decisions about acting in the world**

[1] T. Bayne et al. "What is cognition?," Current Biology, vol. 29, no. 13, pp. R608-R615, 2019. Available: https://www.cell.com/current-biology/pdf/S0960-9822(19)30614-1.pdf

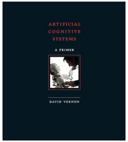
Cognition



- There is no single accepted definition of the term[1], but there are some common threads
- Cognition has to do with **representing and processing information** — obtained from perceptual observations and other learned experiences — **in order to infer suitable decisions about acting in the world**
- Cognition is thus not a purely reactive / associative learning process (of mapping sensory observations to actions), but also involves reasoning about the available information, generating predictions about the world, and subsequently acting to achieve some (potentially distant) goals

[1] T. Bayne et al. "What is cognition?," Current Biology, vol. 29, no. 13, pp. R608-R615, 2019. Available: https://www.cell.com/current-biology/pdf/S0960-9822(19)30614-1.pdf

Cognition

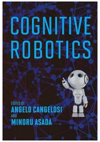


- There is no single accepted definition of the term[1], but there are some common threads
- Cognition has to do with **representing and processing information** — obtained from perceptual observations and other learned experiences — **in order to infer suitable decisions about acting in the world**
- Cognition is thus not a purely reactive / associative learning process (of mapping sensory observations to actions), but also involves reasoning about the available information, generating predictions about the world, and subsequently acting to achieve some (potentially distant) goals

"Cognition is the process by which an autonomous system perceives its environment, learns from experience, anticipates the outcome of events, acts to pursue goals, and adapts to changing circumstances" (Vernon 2014, p. 8)

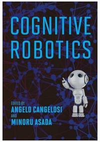
[1] T. Bayne et al. "What is cognition?," Current Biology, vol. 29, no. 13, pp. R608-R615, 2019. Available: https://www.cell.com/current-biology/pdf/S0960-9822(19)30614-1.pdf

Embodied Cognition



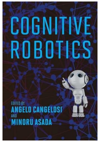
- Some authors consider the embodiment of a robot to be an essential element of its cognitive processes

Embodied Cognition



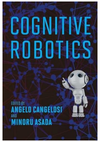
- Some authors consider the embodiment of a robot to be an essential element of its cognitive processes
- In this view, intelligent behaviour only makes sense in terms of the physical interaction of a robot with its environment

Embodied Cognition



- Some authors consider the embodiment of a robot to be an essential element of its cognitive processes
- In this view, intelligent behaviour only makes sense in terms of the physical interaction of a robot with its environment
- An embodied robot is involved in both online physical actions as well as offline reasoning

Embodied Cognition



- Some authors consider the embodiment of a robot to be an essential element of its cognitive processes
- In this view, intelligent behaviour only makes sense in terms of the physical interaction of a robot with its environment
- An embodied robot is involved in both online physical actions as well as offline reasoning

"Embodied cognition is the approach to studying natural intelligent systems that underscores the roles of sensorimotor knowledge and representation and the interaction between our own body and the environment in producing intelligent behavior." (Cangelosi and Asada 2022, p. 7)

Knowledge Representation

- Knowledge representation is concerned with how knowledge that an autonomous robot requires for its operation can be **encoded** so that a robot can use it during its operation

Knowledge Representation

- Knowledge representation is concerned with how knowledge that an autonomous robot requires for its operation can be **encoded** so that a robot can use it during its operation
- Traditionally, concerned with **encoding knowledge in logical languages** (e.g. first-order logic, ontologies, etc.)

Knowledge Representation

- Knowledge representation is concerned with how knowledge that an autonomous robot requires for its operation can be **encoded** so that a robot can use it during its operation
- Traditionally, concerned with **encoding knowledge in logical languages** (e.g. first-order logic, ontologies, etc.)
- Can also include aspects of how **prior knowledge** can be **embedded into statistical machine learning-based systems**

Knowledge Representation

- Knowledge representation is concerned with how knowledge that an autonomous robot requires for its operation can be **encoded** so that a robot can use it during its operation
- Traditionally, concerned with **encoding knowledge in logical languages** (e.g. first-order logic, ontologies, etc.)
- Can also include aspects of how **prior knowledge** can be **embedded into statistical machine learning-based systems**
- A cognitive robot should be able to use the available knowledge for making decisions and expand the knowledge to react to dynamic changes of the world

Knowledge Representation

- Knowledge representation is concerned with how knowledge that an autonomous robot requires for its operation can be **encoded** so that a robot can use it during its operation
- Traditionally, concerned with **encoding knowledge in logical languages** (e.g. first-order logic, ontologies, etc.)
- Can also include aspects of how **prior knowledge** can be **embedded into statistical machine learning-based systems**
- A cognitive robot should be able to use the available knowledge for making decisions and expand the knowledge to react to dynamic changes of the world

Knowledge representation is the encoding of aspects about the world in a formal (explicit or implicit) machine-readable form

Reasoning

- Given some knowledge, a robot should be able to use it so that it can **make appropriate decisions** on how to act; this is the task of reasoning

Reasoning

- Given some knowledge, a robot should be able to use it so that it can **make appropriate decisions** on how to act; this is the task of reasoning
- Reasoning makes it possible to **derive new pieces of information** given what is known about the world (in terms of explicitly encoded knowledge and in terms of observations)

Reasoning

- Given some knowledge, a robot should be able to use it so that it can **make appropriate decisions** on how to act; this is the task of reasoning
- Reasoning makes it possible to **derive new pieces of information** given what is known about the world (in terms of explicitly encoded knowledge and in terms of observations)
- In principle, a **structured knowledge representation** is required to enable the development of computational reasoning procedures

Reasoning

- Given some knowledge, a robot should be able to use it so that it can **make appropriate decisions** on how to act; this is the task of reasoning
- Reasoning makes it possible to **derive new pieces of information** given what is known about the world (in terms of explicitly encoded knowledge and in terms of observations)
- In principle, a **structured knowledge representation** is required to enable the development of computational reasoning procedures
- Traditional logical reasoning and probabilistic reasoning are the two most prominent examples of reasoning for cognitive systems

Reasoning

- Given some knowledge, a robot should be able to use it so that it can **make appropriate decisions** on how to act; this is the task of reasoning
- Reasoning makes it possible to **derive new pieces of information** given what is known about the world (in terms of explicitly encoded knowledge and in terms of observations)
- In principle, a **structured knowledge representation** is required to enable the development of computational reasoning procedures
- Traditional logical reasoning and probabilistic reasoning are the two most prominent examples of reasoning for cognitive systems

Reasoning is a process of using all available information (e.g. scene information, prior knowledge about the world, prior experiences) to make decisions about the best course of action to take in a given situation

Anticipation

- A cognitive robot knows that (a) **its actions have an effect on the world** and (b) **other agents exist and act in the world** as well

Anticipation

- A cognitive robot knows that (a) **its actions have an effect on the world** and (b) **other agents exist and act in the world** as well
- Actions thus need to be performed so that they can take both of these types of information into account

Anticipation

- A cognitive robot knows that (a) **its actions have an effect on the world** and (b) **other agents exist and act in the world** as well
- Actions thus need to be performed so that they can take both of these types of information into account
- When the actions of other agents are predicted, we also talk about **intent recognition**

Anticipation

- A cognitive robot knows that (a) **its actions have an effect on the world** and (b) **other agents exist and act in the world** as well
- Actions thus need to be performed so that they can take both of these types of information into account
- When the actions of other agents are predicted, we also talk about **intent recognition**
- In principle, anticipation needs to go beyond classical checking of action effects - prediction is a temporal prediction problem that should be performed continuously

Anticipation

- A cognitive robot knows that (a) **its actions have an effect on the world** and (b) **other agents exist and act in the world** as well
- Actions thus need to be performed so that they can take both of these types of information into account
- When the actions of other agents are predicted, we also talk about **intent recognition**
- In principle, anticipation needs to go beyond classical checking of action effects - prediction is a temporal prediction problem that should be performed continuously

Anticipation is a process of predicting the effects of a robot's own actions and predicting how other agents will act in a given situation

Learning

- **The ability to learn is an essential property of cognitive robots**, distinguishing them from intelligent robots that simply rely on hand-coded / prior knowledge

Learning

- **The ability to learn is an essential property of cognitive robots**, distinguishing them from intelligent robots that simply rely on hand-coded / prior knowledge
- Learning can be performed using offline data, but **a cognitive robot never stops learning**, namely it continues learning throughout its operation

Learning

- **The ability to learn is an essential property of cognitive robots**, distinguishing them from intelligent robots that simply rely on hand-coded / prior knowledge
- Learning can be performed using offline data, but **a cognitive robot never stops learning**, namely it continues learning throughout its operation
- A cognitive robot can **learn based on its own observations**, but also **using feedback** that is received by other agents

Learning

- **The ability to learn is an essential property of cognitive robots**, distinguishing them from intelligent robots that simply rely on hand-coded / prior knowledge
- Learning can be performed using offline data, but **a cognitive robot never stops learning**, namely it continues learning throughout its operation
- A cognitive robot can **learn based on its own observations**, but also **using feedback** that is received by other agents
- In a sense, **learning should mimic the operation of the hippocampus** in the brain, where information is continually stored and updated

Learning

- **The ability to learn is an essential property of cognitive robots**, distinguishing them from intelligent robots that simply rely on hand-coded / prior knowledge
- Learning can be performed using offline data, but **a cognitive robot never stops learning**, namely it continues learning throughout its operation
- A cognitive robot can **learn based on its own observations**, but also **using feedback** that is received by other agents
- In a sense, **learning should mimic the operation of the hippocampus** in the brain, where information is continually stored and updated

Learning is a process of updating a robot's model of (various aspects of) the world with new information based on observations and feedback

Adaptivity

- A cognitive robot needs to learn and update its models of the world, but it also needs to be able to use that acquired knowledge to **update its behaviour accordingly**

E. von Glasersfeld, "Learning and adaptation in the theory of constructivism," *Communication and Cognition*, vol. 26, no. 3/4, pp. 393–402, 1993. Available: https://dbis-digivis.uibk.ac.at/mediawiki/images/3/3a/Learning_and_Adaptation_in_the_Theory_of_Constructivism.pdf

Adaptivity

- A cognitive robot needs to learn and update its models of the world, but it also needs to be able to use that acquired knowledge to **update its behaviour accordingly**
- Ideally, a robot should **adapt to both changing physical characteristics of the world** and **changing human behaviour and preferences**

E. von Glasersfeld, "Learning and adaptation in the theory of constructivism," *Communication and Cognition*, vol. 26, no. 3/4, pp. 393–402, 1993. Available: https://dbis-digivis.uibk.ac.at/mediawiki/images/3/3a/Learning_and_Adaptation_in_the_Theory_of_Constructivism.pdf

Adaptivity

- A cognitive robot needs to learn and update its models of the world, but it also needs to be able to use that acquired knowledge to **update its behaviour accordingly**
- Ideally, a robot should **adapt to both changing physical characteristics of the world** and **changing human behaviour and preferences**
- **Adaptation needs to be built into the decision-making processes of a robot**, namely adaptivity can be observed as an innate characteristic that enables behaviour updates to deal with a dynamic world

E. von Glasersfeld, "Learning and adaptation in the theory of constructivism," *Communication and Cognition*, vol. 26, no. 3/4, pp. 393–402, 1993. Available: https://dbis-digivis.uibk.ac.at/mediawiki/images/3/3a/Learning_and_Adaptation_in_the_Theory_of_Constructivism.pdf

Adaptivity

- A cognitive robot needs to learn and update its models of the world, but it also needs to be able to use that acquired knowledge to **update its behaviour accordingly**
- Ideally, a robot should **adapt to both changing physical characteristics of the world** and **changing human behaviour and preferences**
- **Adaptation needs to be built into the decision-making processes of a robot**, namely adaptivity can be observed as an innate characteristic that enables behaviour updates to deal with a dynamic world

> Adaptivity is a property that enables a robot to modify its behaviour according to the needs of the environment in which it is active

---

E. von Glasersfeld, "Learning and adaptation in the theory of constructivism," *Communication and Cognition*, vol. 26, no. 3/4, pp. 393–402, 1993. Available: https://dbis-digivis.uibk.ac.at/mediawiki/images/3/3a/Learning_and_Adaptation_in_the_Theory_of_Constructivism.pdf

Theory of Mind

- Particularly in human-robot interaction (but also in robot-robot interaction), it is important for a robot to be able to **understand that a human (or a different robot) may have different beliefs about the world than oneself** (based on the perceptual stance or based on any other factors that may affect the belief)

C. Frith and U. Frith, "Theory of mind: Quick guide," *Current Biology*, vol. 15, no. 7, pp. R644–R645, Sept. 2005. Available: https://www.cell.com/current-biology/pdf/S0960-9822(05)00960-7.pdf

Theory of Mind

- Particularly in human-robot interaction (but also in robot-robot interaction), it is important for a robot to be able to **understand that a human (or a different robot) may have different beliefs about the world than oneself** (based on the perceptual stance or based on any other factors that may affect the belief)
- A theory of mind is not a model of the real state of the world, but **a model of another agent's belief of the world**

C. Frith and U. Frith, "Theory of mind: Quick guide," *Current Biology*, vol. 15, no. 7, pp. R644–R645, Sept. 2005. Available: https://www.cell.com/current-biology/pdf/S0960-9822(05)00960-7.pdf

Theory of Mind

- Particularly in human-robot interaction (but also in robot-robot interaction), it is important for a robot to be able to **understand that a human (or a different robot) may have different beliefs about the world than oneself** (based on the perceptual stance or based on any other factors that may affect the belief)

- A theory of mind is not a model of the real state of the world, but **a model of another agent's belief of the world**

- A cognitive robot needs to have a theory of mind in order to co-exist with other agents successfully; otherwise, interactions will often result in misunderstandings

C. Frith and U. Frith, "Theory of mind: Quick guide," *Current Biology*, vol. 15, no. 7, pp. R644–R645, Sept. 2005. Available: https://www.cell.com/current-biology/pdf/S0960-9822(05)00960-7.pdf

Theory of Mind

- Particularly in human-robot interaction (but also in robot-robot interaction), it is important for a robot to be able to **understand that a human (or a different robot) may have different beliefs about the world than oneself** (based on the perceptual stance or based on any other factors that may affect the belief)
- A theory of mind is not a model of the real state of the world, but **a model of another agent's belief of the world**
- A cognitive robot needs to have a theory of mind in order to co-exist with other agents successfully; otherwise, interactions will often result in misunderstandings

A theory of mind is a robot's ability to understand that other agents may have a different set of beliefs about the world than the robot itself

C. Frith and U. Frith, "Theory of mind: Quick guide," *Current Biology*, vol. 15, no. 7, pp. R644–R645, Sept. 2005. Available: https://www.cell.com/current-biology/pdf/S0960-9822(05)00960-7.pdf

A Brief Look at the Human Brain

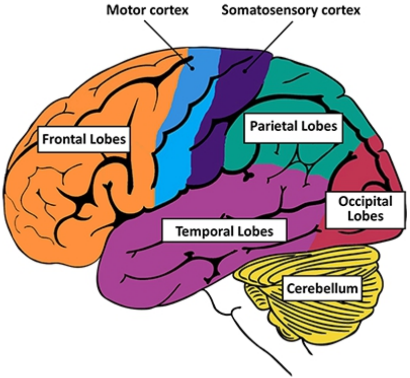
Cognitive Robotics and the Human Brain

- **A significant body of work in cognitive robotics is inspired by or modelled on humans (and other living animals)**; it can thus be useful to look at the structure of the human brain that serves as an inspiration
- Inspiration about the working of the brain is particularly used when developing **cognitive robot architectures** (more on that a bit later and in the next lecture)
- The parallel to the operation of the human brain can be seen in the:
  - components that are included in a cognitive architecture
  - mechanisms by which certain operations are established, or
  - manner in which components communicate with each other

The Human Brain

Our brain has multiple regions that are in charge of different activities:

- **Frontal lobes**: Store working memory and are active during problem solving
- **Motor cortex**: Takes care of planning and executing motions
- **Somatosensory cortex**: Processes signals about touch, temperature, and body movement
- **Parietal lobes**: Multifunctional parts that are active during reading and arithmetic, but also processes signals while eating
- **Occipital lobes**: Process visual input
- **Temporal lobes**: Process auditory signals, but also contain various other sensory memories
- **Cerebellum**: Controls essential body functions (e.g. breathing)

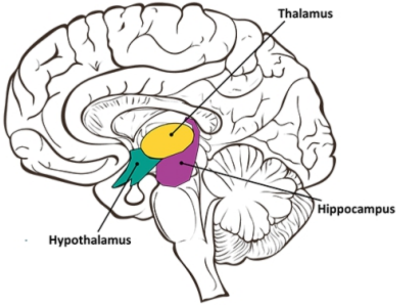


https://www.ninds.nih.gov/health-information/public-education/brain-basics/brain-basics-know-your-brain

The Inner Brain

Many important functions of the body are regulated by the inner brain, which consists of three main elements:

- **Hypothalamus**: Regulates the operation of the body and produces hormones that control emotions
- **Thalamus**: Controls signals to and from the spinal cord, and to and from the cerebrum[2]
- **Hippocampus**: Takes care of consolidating experiences as memories (so they can be preserved and retrieved as long-term memories)

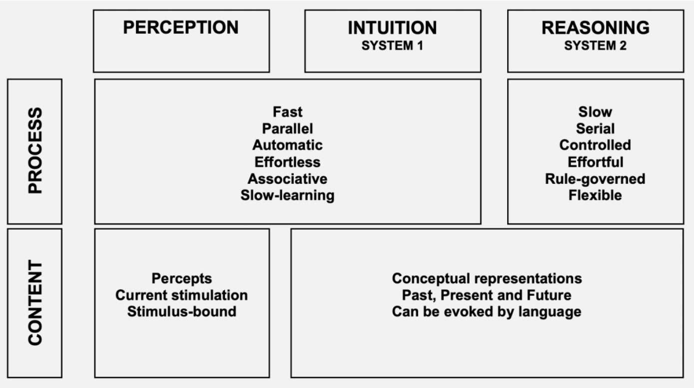


https://www.ninds.nih.gov/health-information/public-education/brain-basics/brain-basics-know-your-brain

[2] The umbrella term about the part of the brain that encompasses the regions discussed on the previous slide

System I and System II Thinking

To describe thinking processes in the brain, a distinction between system I (intuitive) thinking and system II (deliberative) thinking has been proposed

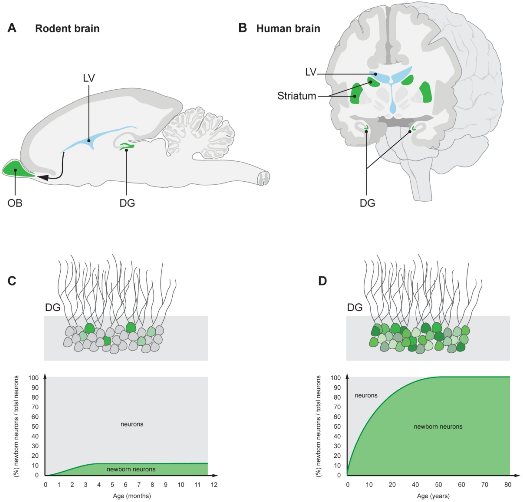


D. Kahneman, "Maps of Bounder Rationality: A Perspective on Intuitive Judgment and Choice," *The American Economic Review*, vol. 93, no. 5, pp. 1449–1475, Dec. 2003. Available: https://scholar.princeton.edu/sites/default/files/kahneman/files/maps_bounded_rationality_dk_2003.pdf

Some cognitive systems model this type of thinking (implicitly or explicitly)

Adult Neurogenesis[3]

- Certain parts of the brain, such as the hippocampus, have the **ability to form new neurons throughout one's life**
- This is a principle called **adult neurogenesis**
- Neurogenesis enables **structural plasticity** and is likely one reason why we can learn and adapt effectively throughout our lives, without losing too much of what we have learned before



[3] A. Ernst and J. Frisén, "Adult neurogenesis in humans-common and unique traits in mammals," *PLoS biology*, vol. 13, no. 1, p. e1002045, 2015. Available: https://doi.org/10.1371/journal.pbio.1002045

Cognitive Agent Architecture Example: Soar

Cognitive Architectures

- A cognitive robot (or, more general, system) architecture **encodes cognitive principles and computational cognitive operations**
- Usually, such **architectures differ in the concrete aspects that they represent**; some focus on the **representation of memories and experience learning**, others encode **principles of affective reasoning**, while yet others represent **motivation as a driver of actions**
- We will look into cognitive architectures in more detail in the next lecture
- At this point, it is useful to look at one example of a cognitive architecture so that the differences with traditional "agent architectures" become apparent

Soar Architecture

- To illustrate cognitive architectures, we will take a brief look at the **Soar architecture**[4]
- Soar is a **generic agent architecture**, not specifically a robot architecture
- This architecture focuses on two main aspects:
  - The representation and use of **memories**
  - **Continual learning** of controllers for execution
- Soar includes a **combination of relational and continuous elements**
- The architecture does not aim to be a faithful representation of an animal brain

[4] J. E. Laird, K. R. Kinkade, S. Mohan, and J. Z. Xu, "Cognitive Robotics Using the Soar Cognitive Architecture," in *Cognitive Robotics Workshop at the 26th AAAI Conf. Artificial Intelligence*, 2012.

Soar Architectural Diagram

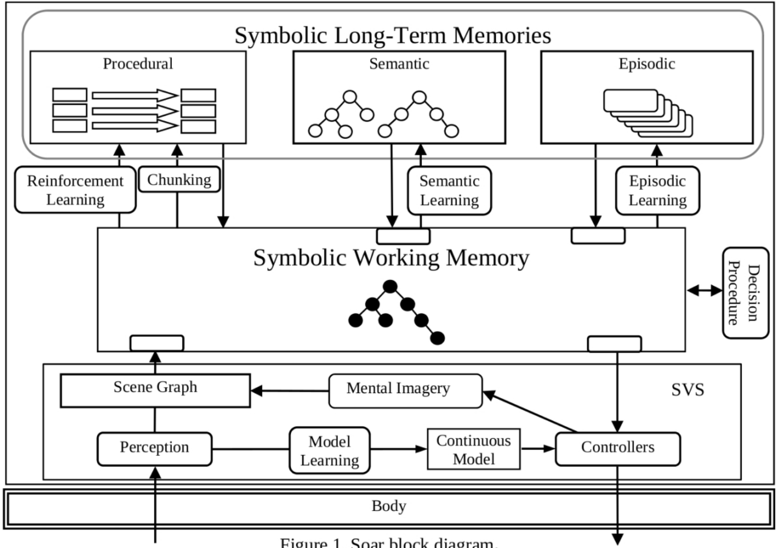


Figure 1. Soar block diagram.

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:
  - **Procedural memory**: Suggests actions based on if-then rules called **productions**

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:
  - **Procedural memory**: Suggests actions based on if-then rules called **productions**
  - **Semantic memory**: Includes ontologies and other knowledge graphs

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:
  - **Procedural memory**: Suggests actions based on if-then rules called **productions**
  - **Semantic memory**: Includes ontologies and other knowledge graphs
  - **Episodic memory**: Stores previous agent experiences, represented in a relational form (as states and actions that have been applied in those states)

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:
  - **Procedural memory**: Suggests actions based on if-then rules called **productions**
  - **Semantic memory**: Includes ontologies and other knowledge graphs
  - **Episodic memory**: Stores previous agent experiences, represented in a relational form (as states and actions that have been applied in those states)
- **SVS (Spatial Visual System)**: Includes a scene graph (detected objects and relations between them) and components for learning predictive models:

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:
  - **Procedural memory**: Suggests actions based on if-then rules called **productions**
  - **Semantic memory**: Includes ontologies and other knowledge graphs
  - **Episodic memory**: Stores previous agent experiences, represented in a relational form (as states and actions that have been applied in those states)
- **SVS (Spatial Visual System)**: Includes a scene graph (detected objects and relations between them) and components for learning predictive models:
  - **Continuous models**: Linear functions that predict the outcomes of continuous actions (multiple models are learned for different qualitative modes)

Soar Elements

- **Symbolic working memory**: Represents sensor measurements and agent goals
- **Symbolic long-term memories**: Long-term storage with three components:
  - **Procedural memory**: Suggests actions based on if-then rules called **productions**
  - **Semantic memory**: Includes ontologies and other knowledge graphs
  - **Episodic memory**: Stores previous agent experiences, represented in a relational form (as states and actions that have been applied in those states)
- **SVS (Spatial Visual System)**: Includes a scene graph (detected objects and relations between them) and components for learning predictive models:
  - **Continuous models**: Linear functions that predict the outcomes of continuous actions (multiple models are learned for different qualitative modes)
  - **Relational models**: Predict how a relational action changes a relational state (information is retrieved from the episodic memory)

Cognitive Robot Learning

Learning Paradigms in Cognitive Robotics

- A variety of learning paradigms can be — and have been — used on cognitive robots

Learning Paradigms in Cognitive Robotics

- A variety of learning paradigms can be — and have been — used on cognitive robots
- The appropriate approach for a given functionality depends on various factors, such as:

Learning Paradigms in Cognitive Robotics

- A variety of learning paradigms can be — and have been — used on cognitive robots
- The appropriate approach for a given functionality depends on various factors, such as:
  - The underlying **computational model**

Learning Paradigms in Cognitive Robotics

- A variety of learning paradigms can be — and have been — used on cognitive robots
- The appropriate approach for a given functionality depends on various factors, such as:
  - The underlying **computational model**
  - Whether learning is performed **independently** or **with / from other agents**

Learning Paradigms in Cognitive Robotics

- A variety of learning paradigms can be — and have been — used on cognitive robots
- The appropriate approach for a given functionality depends on various factors, such as:
  - The underlying **computational model**
  - Whether learning is performed **independently** or **with / from other agents**
  - The **type and amount of available data**

Learning Paradigms in Cognitive Robotics

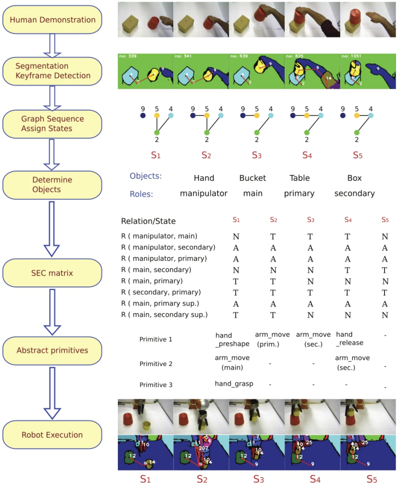
- A variety of learning paradigms can be — and have been — used on cognitive robots
- The appropriate approach for a given functionality depends on various factors, such as:
  - The underlying **computational model**
  - Whether learning is performed **independently** or **with / from other agents**
  - The **type and amount of available data**
  - The nature of the **desired output**

Learning Paradigms in Cognitive Robotics

- A variety of learning paradigms can be — and have been — used on cognitive robots
- The appropriate approach for a given functionality depends on various factors, such as:
  - The underlying **computational model**
  - Whether learning is performed **independently** or **with / from other agents**
  - The **type and amount of available data**
  - The nature of the **desired output**
- Let us briefly look at a few different learning paradigms and what they can be used for
  - More details about some of these techniques are discussed later in the course, or in my "Robot Learning" course

Relational Learning

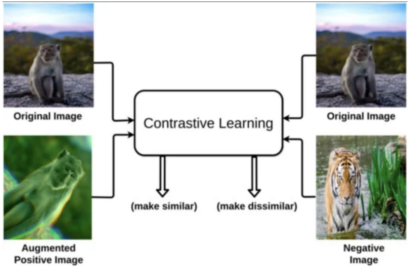
- Has the objective of learning certain **relations about the environment or the operation of a robot**
- In principle, relational learning results in **transparent / interpretable models**
- Usually performed as a **feature extraction procedure** - given known relations and available data, relations (or features) that explain the data are extracted
- Can also include **learning relations** — data are used both to learn relations and to extract those that explain the data



M. J. Aein et al., "Library of actions: Implementing a generic robot execution framework by using manipulation action semantics," *Int. Journal Robotics Research*, vol. 38, no. 8, pp. 910–934, 2019. Available: https://doi.org/10.1177/0278364919850295

(Self-)Supervised Learning

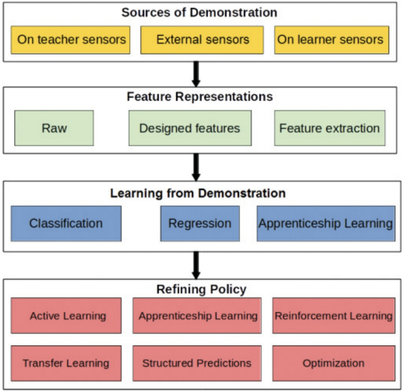
- A cognitive robot can benefit from models acquired using typical (self-)supervised learning in a variety of ways and for a variety of tasks
- **Useful if a large dataset to learn from is available**
- Most pretrained models for vision and natural language processing are based on this paradigm
- **Acquired models do not make a cognitive model on their own** (as discussed in the case of ChatGPT)
  - Only useful to a cognitive robot if they are embedded in a complete cognitive architecture



A. Jaiswal et al., "A survey on contrastive self-supervised learning," *Technologies*, vol. 9, no. 1, 2020. Available: https://doi.org/10.3390/technologies9010002

Imitation Learning

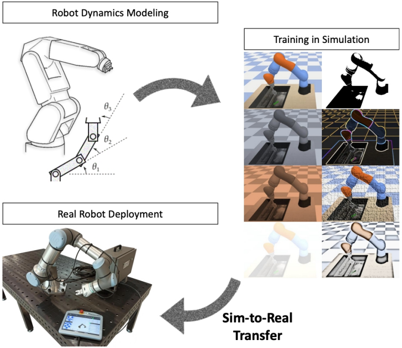
- Enables learning execution policies based on **observations of humans or other agents**
- Mimics the way in which people (and other animals) learn from each other by copying each other's actions
- Typically results in a **temporal predictive model of a robot's state** that allows generating motor actions
- Acquired policies often used to initialise a policy model that is then **refined in a subsequent learning process**



A. Hussein et al. "Imitation learning: A survey of learning methods," *ACM Computing Surveys*, vol. 50, no. 2, pp. 1–35, 2017. Available: https://doi.org/10.1145/3054912

Reinforcement Learning

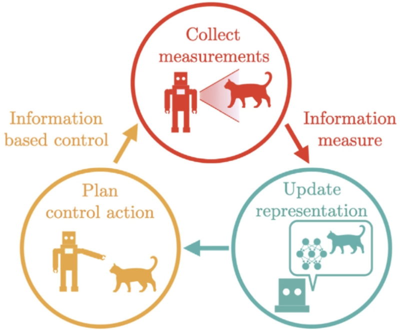
- Concerned with **learning a policy** (mapping from observations to actions) **based on rewards received by interacting with the environment**
- In principle, enables an exploration-guided autonomous learning process — provided that a sufficiently meaningful reward function can be written down
- Difficult to achieve directly on a physical robot platform due to **high data requirements**
  - Learning usually performed **in simulation**; fine-tuning is then done on the physical robot
  - It is possible to speed up learning by **pretraining a policy on offline data** or by **learning from supervisor guidance**



W. Zhao et al., "Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey," in *IEEE Symposium series on Computational Intelligence (SSCI)*, 2020, pp. 737–744. Available: https://doi.org/10.1109/SSCI47803.2020.9308468

Active Learning

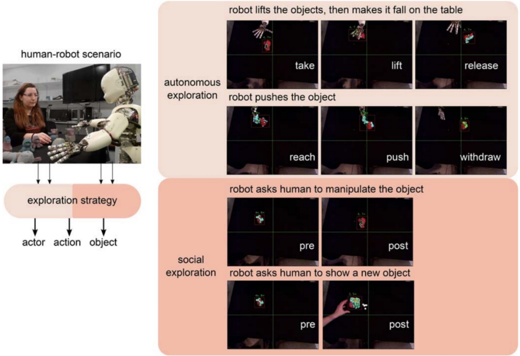
- Enables a learning process during which a robot can actively **focus on relevant learning experiences**
- Experiences are typically selected by **maximising an information measure**
- Aims to **reduce the data requirements of the learning procedure** since the robot does not explore randomly



A. T. Taylor et al., "Active learning in robotics: A review of control principles," *Mechatronics*, vol. 77, p. 102576, 2021. Available: https://doi.org/10.1016/j.mechatronics.2021.102576

Developmental Learning

- A learning paradigm that is **inspired by the way in which babies learn** — individually and in a social setting
- An important goal of developmental learning is to **enable continuous learning**
- **Visual-motor coordination** often a focus of study
- Usually includes a **motivation component that guides the learning process**
  - Note the similarities to active learning



S. Ivaldi et al., "Object Learning Through Active Exploration," *IEEE Trans. Autonomous Mental Development*, vol. 6, no. 1, pp. 56–72, 2014. Available: https://doi.org/10.1109/TAMD.2013.2280614