Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

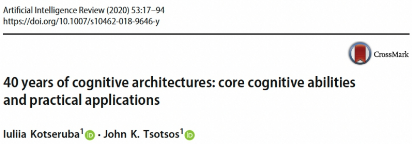
Institute for AI and Autonomous Systems

Cognitive Architectures

A General Overview

Dr. Alex Mitrevski
Master of Autonomous Systems

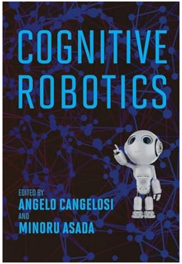
Structure


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

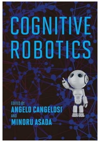
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



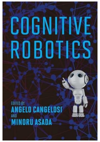
- Preliminaries
- Cognition paradigms
- Elements of cognitive architectures
- Cognitive architecture types and robotics applications of cognitive architectures
- Elements of cognitive architectures in our own codebase

What is a Cognitive Architecture?



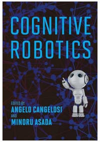
- As any other software architecture, a cognitive architecture defines **components, their representations, the connections between them**, as well as **the communication mechanisms** that define a system

What is a Cognitive Architecture?



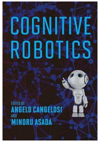
- As any other software architecture, a cognitive architecture defines **components, their representations, the connections between them**, as well as **the communication mechanisms** that define a system
- A cognitive architecture specifies how knowledge is represented, used for decision making, and learned from experience

What is a Cognitive Architecture?



- As any other software architecture, a cognitive architecture defines **components, their representations, the connections between them**, as well as **the communication mechanisms** that define a system
- A cognitive architecture specifies how knowledge is represented, used for decision making, and learned from experience
- The difference with an "ordinary" architecture is that the components of a cognitive architecture enable one or more of the essential cognitive process (as discussed last time)

What is a Cognitive Architecture?



- As any other software architecture, a cognitive architecture defines **components, their representations, the connections between them**, as well as **the communication mechanisms** that define a system

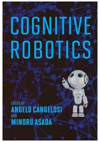
- A cognitive architecture specifies how knowledge is represented, used for decision making, and learned from experience

- The difference with an “ordinary” architecture is that the components of a cognitive architecture enable one or more of the essential cognitive process (as discussed last time)

”A cognitive architecture is a software framework that integrates all the elements required for a system to exhibit the attributes considered to be characteristic of a cognitive agent.” (Cangelosi and Asada ed. 2022, p. 193)

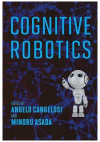
Cognition Paradigms

Cognition Paradigms



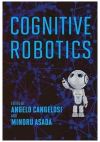
- A cognition paradigm is a set of principles that define how cognition is understood and represented computationally

Cognition Paradigms



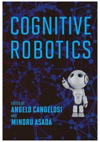
- A cognition paradigm is a set of principles that define how cognition is understood and represented computationally
- Two main paradigms:

Cognition Paradigms



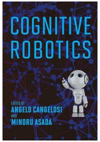
- A cognition paradigm is a set of principles that define how cognition is understood and represented computationally
- Two main paradigms:
  - **Cognitivist**: Cognition is a symbolic manipulation process

Cognition Paradigms



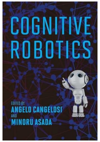
- A cognition paradigm is a set of principles that define how cognition is understood and represented computationally
- Two main paradigms:
  - **Cognitivist**: Cognition is a symbolic manipulation process
  - **Emergent**: Cognition is established by physical interaction with the world and acquiring knowledge through the interaction

Cognition Paradigms



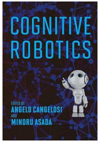
- A cognition paradigm is a set of principles that define how cognition is understood and represented computationally
- Two main paradigms:
  - **Cognitivist**: Cognition is a symbolic manipulation process
  - **Emergent**: Cognition is established by physical interaction with the world and acquiring knowledge through the interaction
- A **hybrid approach** attempts to combine the cognitivist and emergent paradigms into a common framework

Cognitivist Paradigm



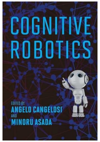
- Views cognition as a **process of manipulating symbolic representations of knowledge**

Cognitivist Paradigm



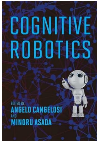
- Views cognition as a **process of manipulating symbolic representations of knowledge**
- In this view, a cognitive agent performs two main operations on data:

Cognitivist Paradigm



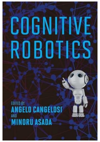
- Views cognition as a **process of manipulating symbolic representations of knowledge**
- In this view, a cognitive agent performs two main operations on data:
  - **converting sensory data to a symbolic representation**

Cognitivist Paradigm



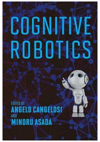
- Views cognition as a **process of manipulating symbolic representations of knowledge**
- In this view, a cognitive agent performs two main operations on data:
  - **converting sensory data to a symbolic representation** and
  - **performing reasoning** on this representation

Cognitivist Paradigm



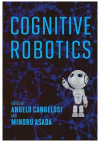
- Views cognition as a **process of manipulating symbolic representations of knowledge**
- In this view, a cognitive agent performs two main operations on data:
  - **converting sensory data to a symbolic representation** and
  - **performing reasoning** on this representation
- According to this paradigm, **cognitive processing happens independent of a concrete embodiment** (computational functionalism)

Emergent Paradigm



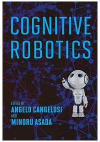
- According to this view, **cognition is tightly coupled with acting in the world** — knowledge about the world is acquired by interacting with it

Emergent Paradigm



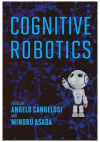
- According to this view, **cognition is tightly coupled with acting in the world** — knowledge about the world is acquired by interacting with it
- Equivalent to the embodied cognition that we talked about in the last lecture

Emergent Paradigm



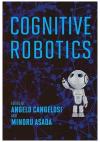
- According to this view, **cognition is tightly coupled with acting in the world** — knowledge about the world is acquired by interacting with it
- Equivalent to the embodied cognition that we talked about in the last lecture
- Since acting in the world is essential for making sense of the world, a cognitive robot based on the emergent paradigm needs to be able to **anticipate actions and their effects**

Emergent Paradigm



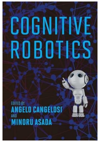
- According to this view, **cognition is tightly coupled with acting in the world** — knowledge about the world is acquired by interacting with it
- Equivalent to the embodied cognition that we talked about in the last lecture
- Since acting in the world is essential for making sense of the world, a cognitive robot based on the emergent paradigm needs to be able to **anticipate actions and their effects**
- Can be implemented through **connectionist systems** (neural networks) or via **dynamical systems theory**

Hybrid Paradigm



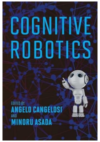
- In the hybrid view, **the merits of both the cognitivist and the emergent paradigms are recognised and exploited**

Hybrid Paradigm



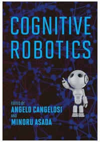
- In the hybrid view, **the merits of both the cognitivist and the emergent paradigms are recognised and exploited**
- In other words, both symbolic and subsymbolic knowledge is combined within a system:

Hybrid Paradigm



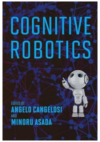
- In the hybrid view, **the merits of both the cognitivist and the emergent paradigms are recognised and exploited**
- In other words, both symbolic and subsymbolic knowledge is combined within a system:
  - Symbolic knowledge is used for **representing the world** and planning **actions for achieving goals** in it

Hybrid Paradigm



- In the hybrid view, **the merits of both the cognitivist and the emergent paradigms are recognised and exploited**
- In other words, both symbolic and subsymbolic knowledge is combined within a system:
  - Symbolic knowledge is used for **representing the world** and planning **actions for achieving goals** in it
  - Subsymbolic models are used for **interpreting perceptual data** and for **interacting with the world**

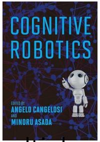
Hybrid Paradigm



- In the hybrid view, **the merits of both the cognitivist and the emergent paradigms are recognised and exploited**
- In other words, both symbolic and subsymbolic knowledge is combined within a system:
  - Symbolic knowledge is used for **representing the world** and planning **actions for achieving goals** in it
  - Subsymbolic models are used for **interpreting perceptual data** and for **interacting with the world**
- A hybrid system also requires an embodiment to interact with the world

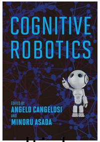
Elements of Cognitive Architectures

Desiderata for Cognitive Architectures



- According to Sun (the creator of the CLARION architecture), cognitive architectures have multiple desired characteristics:

Desiderata for Cognitive Architectures

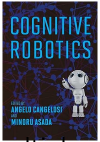


- According to Sun (the creator of the CLARION architecture), cognitive architectures have multiple desired characteristics:

Cognitive realism

An architecture should integrate multidisciplinary ideas about human cognition from a variety of fields

Desiderata for Cognitive Architectures



- According to Sun (the creator of the CLARION architecture), cognitive architectures have multiple desired characteristics:

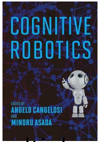
Cognitive realism

An architecture should integrate multidisciplinary ideas about human cognition from a variety of fields

Bioevolutionary realism

An architecture modelling human intelligence should be able to represent animal intelligence in general (due to evolution)

Desiderata for Cognitive Architectures



- According to Sun (the creator of the CLARION architecture), cognitive architectures have multiple desired characteristics:

Cognitive realism

An architecture should integrate multidisciplinary ideas about human cognition from a variety of fields

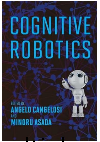
Bioevolutionary realism

An architecture modelling human intelligence should be able to represent animal intelligence in general (due to evolution)

Ecological realism

An architecture should allow a system to exist and act in its "natural" environment

Desiderata for Cognitive Architectures



- According to Sun (the creator of the CLARION architecture), cognitive architectures have multiple desired characteristics:

Cognitive realism

An architecture should integrate multidisciplinary ideas about human cognition from a variety of fields

Bioevolutionary realism

An architecture modelling human intelligence should be able to represent animal intelligence in general (due to evolution)

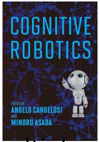
Ecological realism

An architecture should allow a system to exist and act in its "natural" environment

Eclecticism of methodologies

New models should be based on or improve on old models of cognition

Desiderata for Cognitive Architectures



- According to Sun (the creator of the CLARION architecture), cognitive architectures have multiple desired characteristics:

Cognitive realism

An architecture should integrate multidisciplinary ideas about human cognition from a variety of fields

Bioevolutionary realism

An architecture modelling human intelligence should be able to represent animal intelligence in general (due to evolution)

Ecological realism

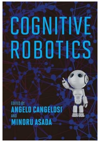
An architecture should allow a system to exist and act in its "natural" environment

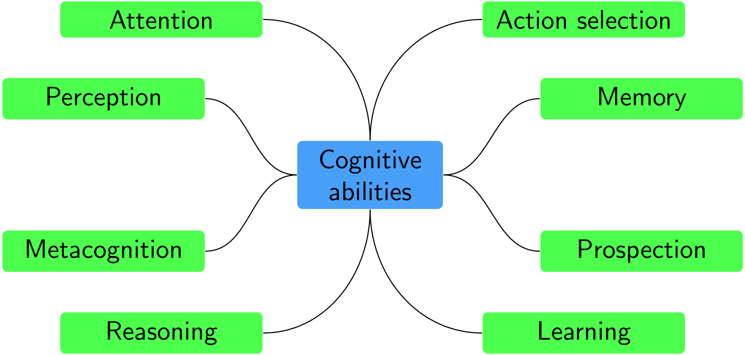
Eclecticism of methodologies

New models should be based on or improve on old models of cognition

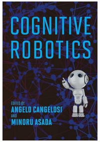
- Note: These desiderata are for "unified cognition theories", but cognitive robot architectures usually only focus on a few characteristics that are of relevance to a robot — they do not aim to be general theories of cognition

Element Overview



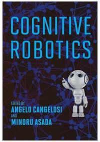


Perception



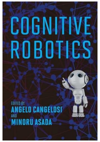
- Most traditional cognitive architectures focus on **visual input and range measurements**; some also consider audio input, while other sensory modalities are usually ignored
- In most architectures, **multiple modalities are treated separately instead of together**; improvements in multimodal neural networks may change this trend though
- Descriptions of cognitive architectures often ignore the **complexity of dealing with real-world noisy data**

Attention



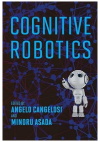
- A process based on which **“irrelevant” information is filtered out of sensory data** and only “relevant” data is left to be processed by the cognitive system

Attention



- A process based on which **"irrelevant" information is filtered out of sensory data** and only "relevant" data is left to be processed by the cognitive system
- Three general attention mechanisms, for instance guided by task relevance or motivation:

Attention



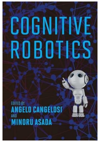
- A process based on which **“irrelevant” information is filtered out of sensory data** and only “relevant” data is left to be processed by the cognitive system

- Three general attention mechanisms, for instance guided by task relevance or motivation:

Selection

A process of selecting one measurement or entity from the complete set

Attention



- A process based on which **“irrelevant” information is filtered out of sensory data** and only “relevant” data is left to be processed by the cognitive system
- Three general attention mechanisms, for instance guided by task relevance or motivation:

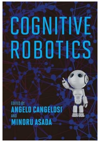
Selection

A process of selecting one measurement or entity from the complete set

Restriction

A process of selecting a subset of measurements or entities, e.g. by limiting the field of view

Attention



- A process based on which **“irrelevant” information is filtered out of sensory data** and only “relevant” data is left to be processed by the cognitive system
- Three general attention mechanisms, for instance guided by task relevance or motivation:

Selection

A process of selecting one measurement or entity from the complete set

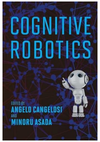
Restriction

A process of selecting a subset of measurements or entities, e.g. by limiting the field of view

Suppression

A process of removing measurements of entities, e.g. by inhibiting features or stimuli

Attention



- A process based on which **"irrelevant" information is filtered out of sensory data** and only "relevant" data is left to be processed by the cognitive system
- Three general attention mechanisms, for instance guided by task relevance or motivation:

Selection

A process of selecting one measurement or entity from the complete set

Restriction

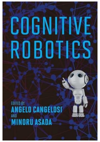
A process of selecting a subset of measurements or entities, e.g. by limiting the field of view

Suppression

A process of removing measurements of entities, e.g. by inhibiting features or stimuli

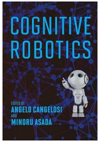
- Most cognitive architectures consider selection and restriction

Action Selection



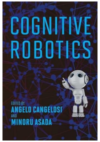
- A process of determining **which action to perform next to achieve some goal**

Action Selection



- A process of determining **which action to perform next to achieve some goal**
- Two general action selection strategies:

Action Selection

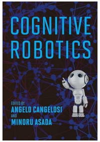


- A process of determining **which action to perform next to achieve some goal**
- Two general action selection strategies:

Planning

A sequence of actions is selected in advance (using traditional AI task planning)

Action Selection



- A process of determining **which action to perform next to achieve some goal**
- Two general action selection strategies:

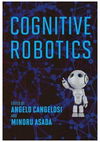
Planning

A sequence of actions is selected in advance (using traditional AI task planning)

Dynamic action selection

Actions are selected online as they are needed, for instance probabilistically, based on relevance, or using some heuristics

Action Selection



- A process of determining **which action to perform next to achieve some goal**
- Two general action selection strategies:

Planning

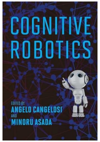
A sequence of actions is selected in advance (using traditional AI task planning)

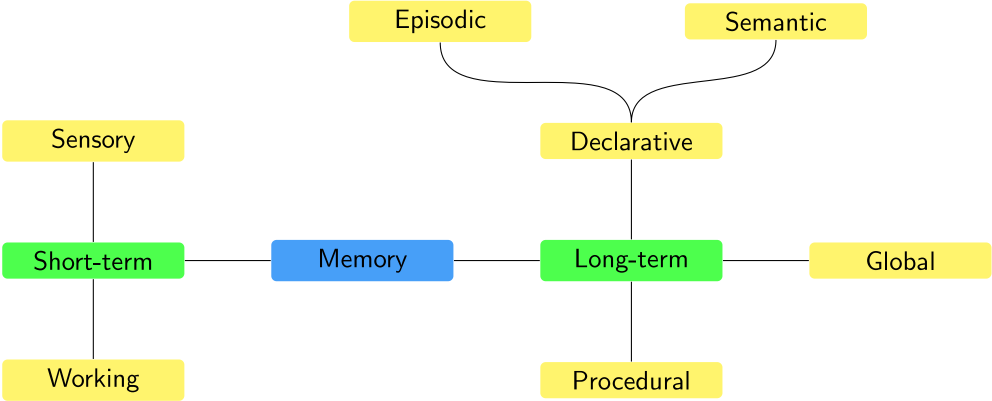
Dynamic action selection

Actions are selected online as they are needed, for instance probabilistically, based on relevance, or using some heuristics

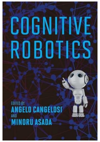
- Both approaches can also be combined: planning can be used for selecting a "global" sequence of actions, while dynamic selection can be performed for reactive adaptation

Memory


COGNITIVE
ROBOTICS
ANGELO CANGELOSI
MINORU ASADA

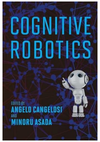

Episodic
Semantic
Sensory
Declarative
Short-term
Memory
Long-term
Global
Working
Procedural

Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

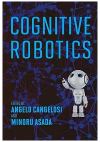
Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory

Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory | **Long-term memory**

Memory Types Overview

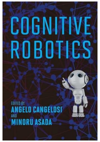
We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory | **Long-term memory**

Sensory

A short buffer of most recent sensor data

Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory | **Long-term memory**

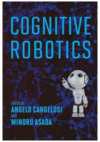
Sensory

A short buffer of most recent sensor data

Working

Stores information relevant for a robot's current task (e.g. recently seen objects on the surface the robot is currently working on)

Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory

Sensory

A short buffer of most recent sensor data

Working

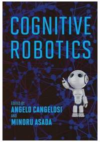
Stores information relevant for a robot's current task (e.g. recently seen objects on the surface the robot is currently working on)

Long-term memory

Declarative

- **Episodic**: Stores previous experiences (observed states and taken actions)
- **Semantic**: Contains general environment information (e.g. objects in the environment and their relations, or connections between areas that the robot can navigate to)

Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory

Sensory

A short buffer of most recent sensor data

Working

Stores information relevant for a robot's current task (e.g. recently seen objects on the surface the robot is currently working on)

Long-term memory

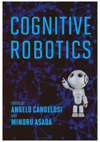
Declarative

- **Episodic**: Stores previous experiences (observed states and taken actions)
- **Semantic**: Contains general environment information (e.g. objects in the environment and their relations, or connections between areas that the robot can navigate to)

Procedural

Enables a robot to select appropriate actions in given states for performing its tasks

Memory Types Overview



We can distinguish between **two general types of memory** in cognitive architectures:

Short-term memory

Sensory

A short buffer of most recent sensor data

Working

Stores information relevant for a robot's current task (e.g. recently seen objects on the surface the robot is currently working on)

Long-term memory

Declarative

- **Episodic**: Stores previous experiences (observed states and taken actions)
- **Semantic**: Contains general environment information (e.g. objects in the environment and their relations, or connections between areas that the robot can navigate to)

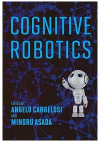
Procedural

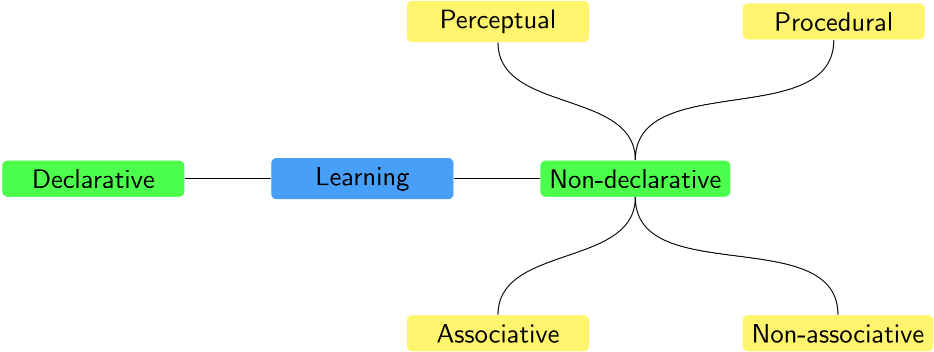
Enables a robot to select appropriate actions in given states for performing its tasks

Global

A global memory structure for all knowledge (not used very commonly)

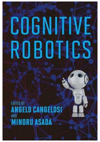
Learning


COGNITIVE ROBOTICS
EDITED BY
ANGELO CANGELOSI
AND
MINORU ASADA


Perceptual
Procedural
Declarative
Learning
Non-declarative
Associative
Non-associative

Learning Types Overview

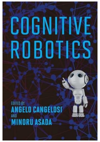
There are **two general learning types** in cognitive architectures:



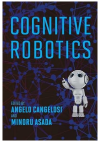
Learning Types Overview

There are **two general learning types** in cognitive architectures:

Declarative



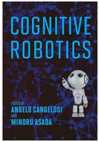
Learning Types Overview



There are **two general learning types** in cognitive architectures:

Declarative | **Non-declarative**

Learning Types Overview



There are **two general learning types** in cognitive architectures:

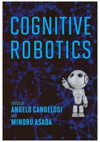
Declarative | **Non-declarative**

Fact learning

Adding facts or rules to declarative memory (**chunks** are added to the memory)

Learning Types Overview

There are **two general learning types** in cognitive architectures:



Declarative	**Non-declarative**

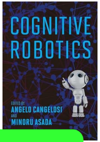
Fact learning

Adding facts or rules to declarative memory (**chunks** are added to the memory)

Percept-to-concept association

Sensory data is mapped to explicit concepts (e.g. objects) - used in emergent and hybrid architectures

Learning Types Overview



There are **two general learning types** in cognitive architectures:

Declarative

Fact learning

Adding facts or rules to declarative memory (**chunks** are added to the memory)

Percept-to-concept association

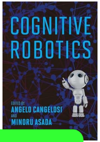
Sensory data is mapped to explicit concepts (e.g. objects) - used in emergent and hybrid architectures

Non-declarative

Perceptual

Learning information about the environment from sensory data (e.g. learning navigation maps)

Learning Types Overview



There are **two general learning types** in cognitive architectures:

Declarative

Fact learning

Adding facts or rules to declarative memory (**chunks** are added to the memory)

Percept-to-concept association

Sensory data is mapped to explicit concepts (e.g. objects) - used in emergent and hybrid architectures

Non-declarative

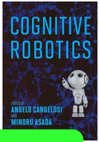
Perceptual

Learning information about the environment from sensory data (e.g. learning navigation maps)

Procedural

Skill learning from experience

Learning Types Overview



There are **two general learning types** in cognitive architectures:

Declarative

Fact learning

Adding facts or rules to declarative memory (**chunks** are added to the memory)

Percept-to-concept association

Sensory data is mapped to explicit concepts (e.g. objects) - used in emergent and hybrid architectures

Non-declarative

Perceptual

Learning information about the environment from sensory data (e.g. learning navigation maps)

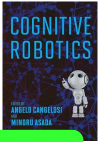
Procedural

Skill learning from experience

Associative

Learning decision-making rules based on rewards (e.g. reinforcement learning) —— performed at a “higher level” than procedural learning

Learning Types Overview



There are **two general learning types** in cognitive architectures:

Declarative

Fact learning

Adding facts or rules to declarative memory (**chunks** are added to the memory)

Percept-to-concept association

Sensory data is mapped to explicit concepts (e.g. objects) - used in emergent and hybrid architectures

Non-declarative

Perceptual

Learning information about the environment from sensory data (e.g. learning navigation maps)

Procedural

Skill learning from experience

Associative

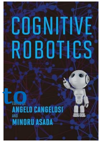
Learning decision-making rules based on rewards (e.g. reinforcement learning) —— performed at a "higher level" than procedural learning

Non-associative

Learning how to react to specific stimuli:

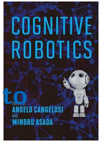
- **Habituation**: With a reduced strength
- **Sensitisation**: With an increased strength

Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

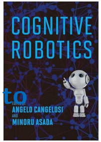
Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

- Three types of reasoning:

Reasoning



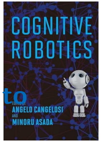
- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

- Three types of reasoning:

Deductive

Drawing conclusions (about individual instances) based on given premises

Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

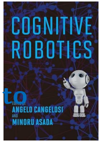
- Three types of reasoning:

Deductive

Drawing conclusions (about individual instances) based on given premises

The correctness of a deductive argument is **independent of the validity of the premises** (i.e. the reasoning process itself can be correct even with incorrect premises); however, a **sound argument** is one in which the premises are actually **valid**

Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

- Three types of reasoning:

Deductive

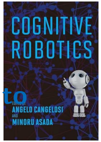
Drawing conclusions (about individual instances) based on given premises

The correctness of a deductive argument is **independent of the validity of the premises** (i.e. the reasoning process itself can be correct even with incorrect premises); however, a **sound argument** is one in which the premises are actually **valid**

Inductive

Drawing generalised conclusions based on individual observations

Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

- Three types of reasoning:

Deductive

Drawing conclusions (about individual instances) based on given premises

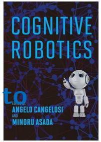
The correctness of a deductive argument is **independent of the validity of the premises** (i.e. the reasoning process itself can be correct even with incorrect premises); however, a **sound argument** is one in which the premises are actually **valid**

Inductive

Drawing generalised conclusions based on individual observations

The conclusion of an inductive argument cannot be (logically) guaranteed, but the amount of evidence determines the strength of the argument

Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

- Three types of reasoning:

Deductive

Drawing conclusions (about individual instances) based on given premises

The correctness of a deductive argument is **independent of the validity of the premises** (i.e. the reasoning process itself can be correct even with incorrect premises); however, a **sound argument** is one in which the premises are actually **valid**

Inductive

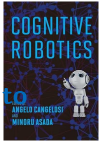
Drawing generalised conclusions based on individual observations

The conclusion of an inductive argument cannot be (logically) guaranteed, but the amount of evidence determines the strength of the argument

Abductive

Reasoning based on incomplete information to obtain the likeliest conclusion about a causal relation

Reasoning



- A process of **knowledge manipulation** using logical rules in order to **infer conclusions or to choose actions to perform**

- Three types of reasoning:

Deductive

Drawing conclusions (about individual instances) based on given premises

The correctness of a deductive argument is **independent of the validity of the premises** (i.e. the reasoning process itself can be correct even with incorrect premises); however, a **sound argument** is one in which the premises are actually **valid**

Inductive

Drawing generalised conclusions based on individual observations

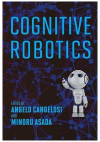
The conclusion of an inductive argument cannot be (logically) guaranteed, but the amount of evidence determines the strength of the argument

Abductive

Reasoning based on incomplete information to obtain the likeliest conclusion about a causal relation

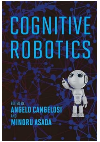
- Mostly present in cognitivist and hybrid systems; emergent architectures typically do not / cannot facilitate reasoning

Metacognition



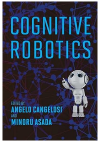
- Defined as “knowledge and beliefs about one’s own cognitive processes” (Oxford Reference)

Metacognition



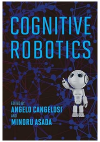
- Defined as “knowledge and beliefs about one’s own cognitive processes” (Oxford Reference)
- Consists of **monitoring the operation of a cognitive system and its cognitive processes**

Metacognition



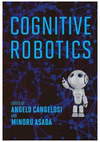
- Defined as “knowledge and beliefs about one’s own cognitive processes” (Oxford Reference)
- Consists of **monitoring the operation of a cognitive system and its cognitive processes**
- Needed if a cognitive robot is supposed to **improve its cognitive processes over time** (e.g. become better at planning or prediction)

Metacognition



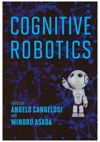
- Defined as "knowledge and beliefs about one's own cognitive processes" (Oxford Reference)
- Consists of **monitoring the operation of a cognitive system and its cognitive processes**
- Needed if a cognitive robot is supposed to **improve its cognitive processes over time** (e.g. become better at planning or prediction)
- Not present very commonly in cognitive architectures — presumably because of the difficulty of embedding it and performing it meaningfully

Prospection



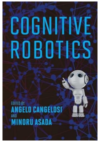
- Enables a cognitive robot to **predict the effects of actions** as well as **anticipate what other agents might do**

Prospection



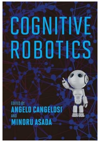
- Enables a cognitive robot to **predict the effects of actions** as well as **anticipate what other agents might do**
- **Important for all cognitive processes**, but particularly for (active) perception, action selection, attention, and anticipation

Prospection



- Enables a cognitive robot to **predict the effects of actions** as well as **anticipate what other agents might do**
- **Important for all cognitive processes**, but particularly for (active) perception, action selection, attention, and anticipation
- Metacognition needs prospection as well — cognitive processes can be improved meaningfully only if the effects of the current cognitive processes can be predicted

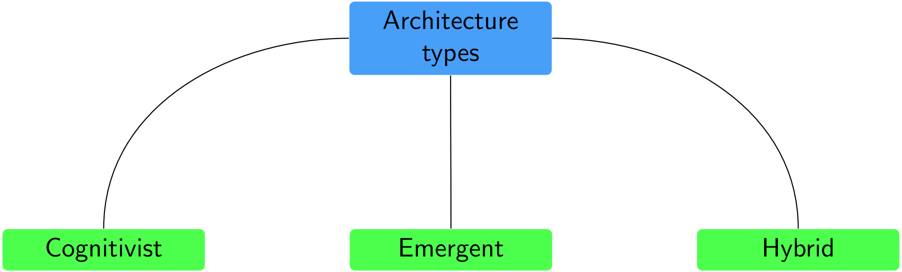
Prospection



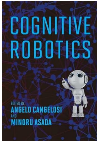
- Enables a cognitive robot to **predict the effects of actions** as well as **anticipate what other agents might do**
- **Important for all cognitive processes**, but particularly for (active) perception, action selection, attention, and anticipation
- Metacognition needs prospection as well — cognitive processes can be improved meaningfully only if the effects of the current cognitive processes can be predicted
- **Achieved through an internal (mental) simulation**
  - Note that this does not necessarily have to be done in a full-fledged simulator — simulation models can also be performed by computational models in an architecture

Cognitive Architecture Types and Robotics Applications of Cognitive Architectures

Classification of Cognitive Architectures

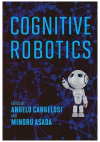


Cognitivist (Symbolic)



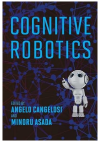
- Focus on **modelling task-independent aspects of cognition**

Cognitivist (Symbolic)



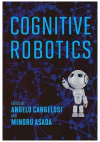
- Focus on **modelling task-independent aspects of cognition**
- **Task-specific knowledge needs to be explicitly provided** to an agent that is based on a cognitivist architecture
  - But learning can be used to improve on the provided knowledge

Cognitivist (Symbolic)



- Focus on **modelling task-independent aspects of cognition**
- **Task-specific knowledge needs to be explicitly provided** to an agent that is based on a cognitivist architecture
  - But learning can be used to improve on the provided knowledge
- A **cognitive model** is obtained when a cognitive architecture is combined with domain- or task-specific knowledge

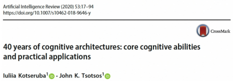
Cognitivist (Symbolic)



- Focus on **modelling task-independent aspects of cognition**
- **Task-specific knowledge needs to be explicitly provided** to an agent that is based on a cognitivist architecture
  - But learning can be used to improve on the provided knowledge
- A **cognitive model** is obtained when a cognitive architecture is combined with domain- or task-specific knowledge

> "A **cognitivist cognitive** architecture is a generic computational model that is neither domain-specific nor task-specific, and it needs to be provided with knowledge to perform any given task." (Cangelosi and Asada ed. 2022, p. 194)

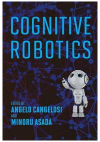
Emergent (Subsymbolic)


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

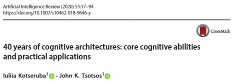
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



- In an emergent cognitive architecture, **an agent starts in a basic cognitive state that is developed throughout the agent's operation**

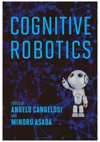
Emergent (Subsymbolic)


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

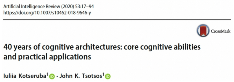
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



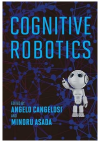
- In an emergent cognitive architecture, **an agent starts in a basic cognitive state that is developed throughout the agent's operation**
- The cognitive model of an emergent agent is developed by **exposing the agent to its operational environment** so that it can acquire relevant experiences (but without being overwhelmed by them)
  - Thus, unlike in cognitivist architectures, the knowledge is acquired by the agent itself

Emergent (Subsymbolic)


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

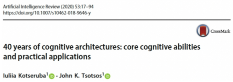
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



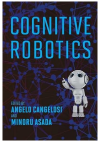
- In an emergent cognitive architecture, **an agent starts in a basic cognitive state that is developed throughout the agent's operation**
- The cognitive model of an emergent agent is developed by **exposing the agent to its operational environment** so that it can acquire relevant experiences (but without being overwhelmed by them)
  - Thus, unlike in cognitivist architectures, the knowledge is acquired by the agent itself
- The **embodiment** is an essential part of emergent agents since the agent needs to be able to explore its environment in order to acquire learning experiences

Emergent (Subsymbolic)


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

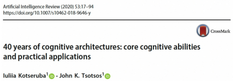
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



- In an emergent cognitive architecture, **an agent starts in a basic cognitive state that is developed throughout the agent's operation**
- The cognitive model of an emergent agent is developed by **exposing the agent to its operational environment** so that it can acquire relevant experiences (but without being overwhelmed by them)
  - Thus, unlike in cognitivist architectures, the knowledge is acquired by the agent itself
- The **embodiment** is an essential part of emergent agents since the agent needs to be able to explore its environment in order to acquire learning experiences
- Emergent architectures differ in the biological plausibility of the underlying models (some are based on artificial neural networks, others on more biologically accurate representations)

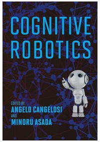
Emergent (Subsymbolic)


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

40 years of cognitive architectures: core cognitive abilities and practical applications

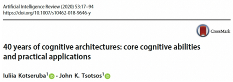
Iuliia Kotseruba[1] · John K. Tsotsos[1]



- In an emergent cognitive architecture, **an agent starts in a basic cognitive state that is developed throughout the agent's operation**
- The cognitive model of an emergent agent is developed by **exposing the agent to its operational environment** so that it can acquire relevant experiences (but without being overwhelmed by them)
  - Thus, unlike in cognitivist architectures, the knowledge is acquired by the agent itself
- The **embodiment** is an essential part of emergent agents since the agent needs to be able to explore its environment in order to acquire learning experiences
- Emergent architectures differ in the biological plausibility of the underlying models (some are based on artificial neural networks, others on more biologically accurate representations)

"An emergent cognitive architecture is the initial state from which an agent subsequently develops."
(Cangelosi and Asada ed. 2022, p. 194)

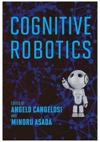
Hybrid


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

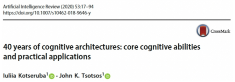
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



- Hybrid architectures are **based on the hybrid view of cognition** and include elements of both cognitivist and emergent architectures

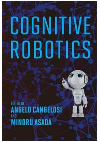
Hybrid


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

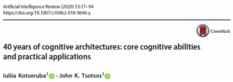
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



- Hybrid architectures are **based on the hybrid view of cognition** and include elements of both cognitivist and emergent architectures
- The difference between different hybrid architectures lies in the **elements that are included and in how they are combined**

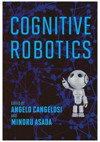
Hybrid


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

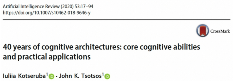
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



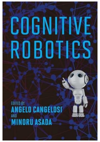
- Hybrid architectures are **based on the hybrid view of cognition** and include elements of both cognitivist and emergent architectures
- The difference between different hybrid architectures lies in the **elements that are included and in how they are combined**
  - Hybrid architectures are thus typically more **domain-specific**, as they focus on encoding aspects that are relevant for the domain of interest

Hybrid


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

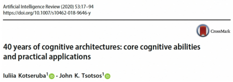
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



- Hybrid architectures are **based on the hybrid view of cognition** and include elements of both cognitivist and emergent architectures
- The difference between different hybrid architectures lies in the **elements that are included and in how they are combined**
  - Hybrid architectures are thus typically more **domain-specific**, as they focus on encoding aspects that are relevant for the domain of interest
- In some hybrid architectures, subsymbolic processing is only performed for extracting symbols from sensory measurements, but a more common case is to incorporate multiple symbolic and subsymbolic elements and processes

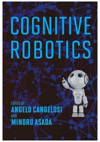
Hybrid


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



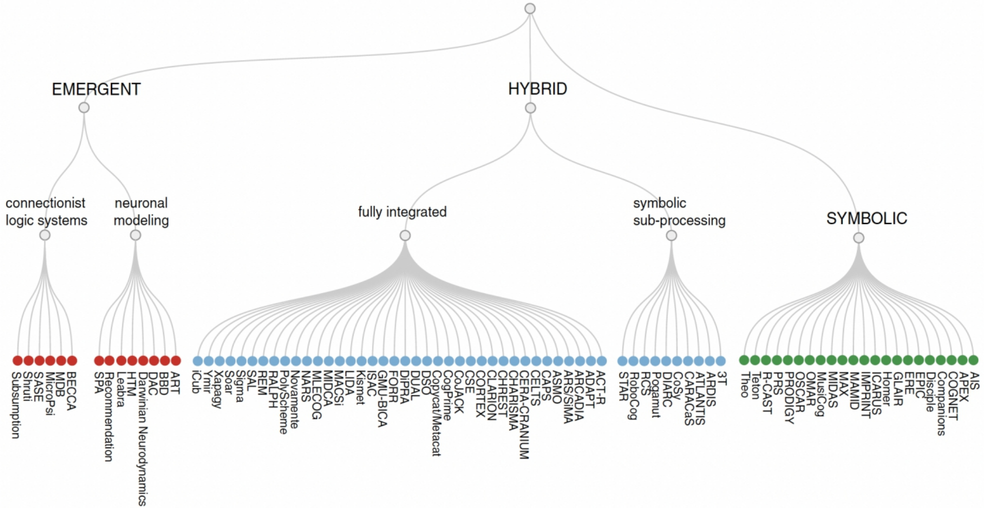
- Hybrid architectures are **based on the hybrid view of cognition** and include elements of both cognitivist and emergent architectures
- The difference between different hybrid architectures lies in the **elements that are included and in how they are combined**
  - Hybrid architectures are thus typically more **domain-specific**, as they focus on encoding aspects that are relevant for the domain of interest
- In some hybrid architectures, subsymbolic processing is only performed for extracting symbols from sensory measurements, but a more common case is to incorporate multiple symbolic and subsymbolic elements and processes
- Hybridisation can also involve **combining multiple architectures**

Architecture Taxonomy

Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

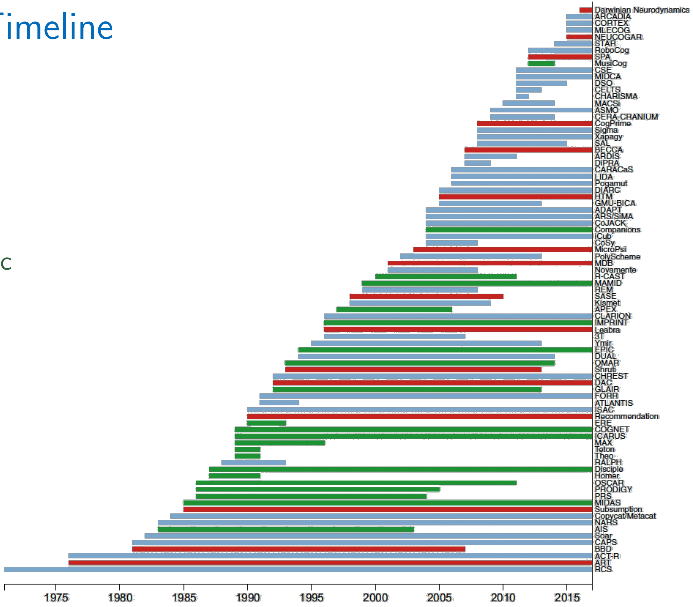
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



Architecture Timeline

- Green: symbolic
- Red: emergent
- Blue: hybrid

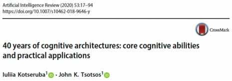


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

Architecture Timeline Comments


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

The previous graph reveals a few important aspects about cognitive architectures:

Architecture Timeline Comments

Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

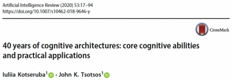
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

The previous graph reveals a few important aspects about cognitive architectures:

- **There is a large variety of cognitive architectures** — not all of these are tailored to or have been used in robotics

Architecture Timeline Comments


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

The previous graph reveals a few important aspects about cognitive architectures:

- **There is a large variety of cognitive architectures** — not all of these are tailored to or have been used in robotics
- Cognitivist architectures used to be more popular than emergent and hybrid architectures before the turn of the 21st century; since then, **hybrid architectures are most common**, particularly for architectures that are newly developed

Architecture Timeline Comments

Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

The previous graph reveals a few important aspects about cognitive architectures:

- **There is a large variety of cognitive architectures** — not all of these are tailored to or have been used in robotics
- Cognitivist architectures used to be more popular than emergent and hybrid architectures before the turn of the 21st century; since then, **hybrid architectures are most common**, particularly for architectures that are newly developed
- **The development of a cognitive architecture is a long process** (lasting several years), with architectures sometimes undergoing major changes and conceptual updates

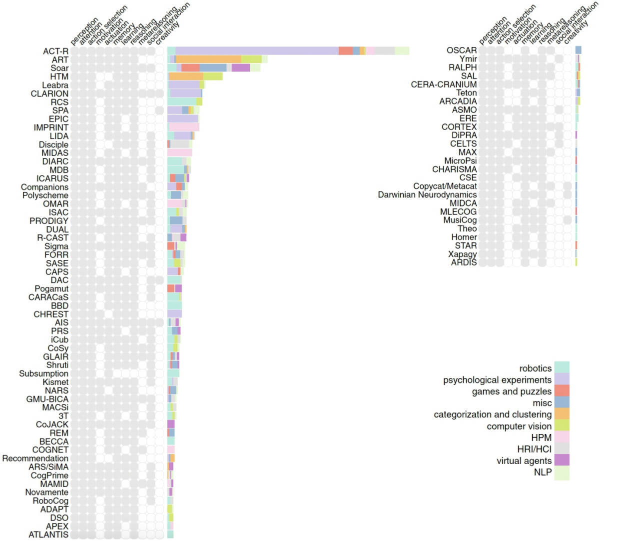
An Overview of Applications of Cognitive Architectures

Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

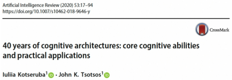
40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

- Architectures differ in the incorporated features and in the applications that they have been used for
- Robotics is not the number one use case for cognitive architectures — psychological experiments are
- Robotics is a close second application area though



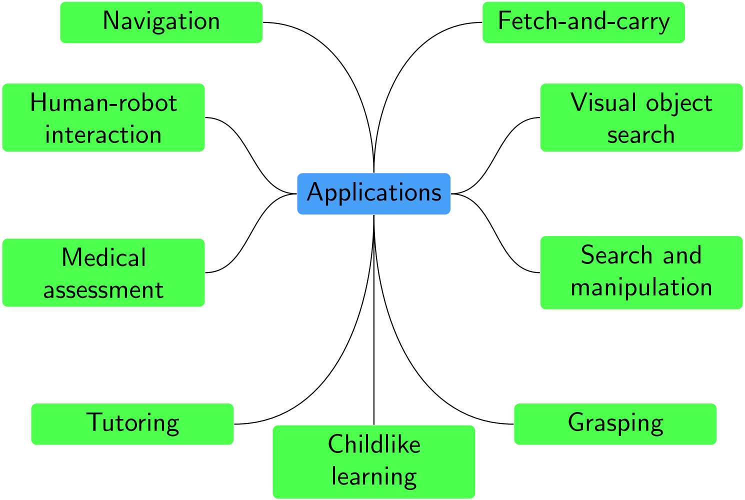
Overview of Demonstrated Robotics Applications



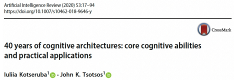
Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]



Comments on the Applications


Artificial Intelligence Review (2020) 53:17–94
https://doi.org/10.1007/s10462-018-9646-y

CrossMark

40 years of cognitive architectures: core cognitive abilities and practical applications

Iuliia Kotseruba[1] · John K. Tsotsos[1]

- Multiple architectures have been used to implement use cases such as navigation, fetch-and-carry, or visual search and manipulation
- More complex use cases are usually only demonstrated with a single architecture
  - Too time- and resource-intensive to reproduce large studies
  - Even if an architecture does, in theory, include all properties that would enable its use for a given task, the practical implementation may require significant effort
  - For concrete applications, architectures that are tailored to the use case are easier to develop and manage

Elements of Cognitive Architectures in our Codebase

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions
  - Memory:
    - Working memory in the form of a **fact database** that is updated by the robot while it is acting

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions
  - Memory:
    - Working memory in the form of a **fact database** that is updated by the robot while it is acting
    - Semantic memory in the form of an **environment ontology**

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions
  - Memory:
    - Working memory in the form of a **fact database** that is updated by the robot while it is acting
    - Semantic memory in the form of an **environment ontology**
    - Rudimentary procedural memory based on which **behaviours to be executed are specified through a declarative language** (in terms of state machines)

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions
  - Memory:
    - Working memory in the form of a **fact database** that is updated by the robot while it is acting
    - Semantic memory in the form of an **environment ontology**
    - Rudimentary procedural memory based on which **behaviours to be executed are specified through a declarative language** (in terms of state machines)
  - Learning:
    - **Percept-to-concept association** (using deep learning, concretely for object recognition)

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions
  - Memory:
    - Working memory in the form of a **fact database** that is updated by the robot while it is acting
    - Semantic memory in the form of an **environment ontology**
    - Rudimentary procedural memory based on which **behaviours to be executed are specified through a declarative language** (in terms of state machines)
  - Learning:
    - **Percept-to-concept association** (using deep learning, concretely for object recognition)
    - Procedural learning (**concretely for learning manipulation skills** — my own PhD work[1])

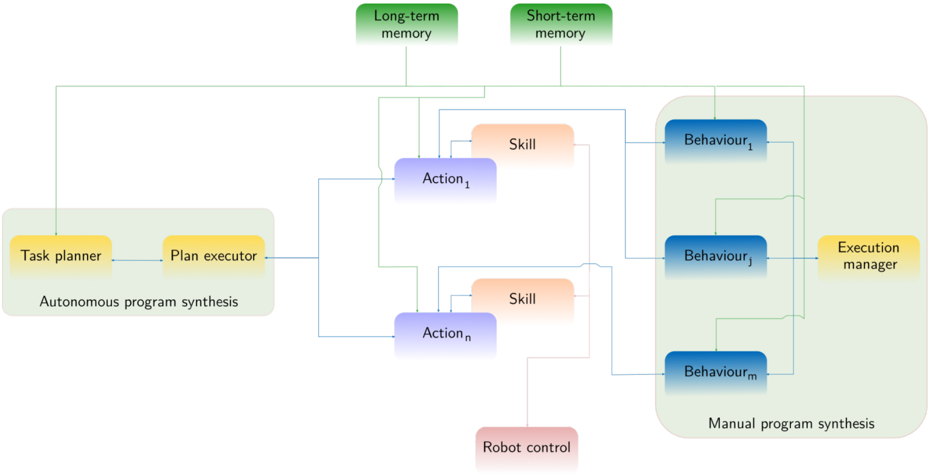
[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Are There Cognitive Elements in Our Codebase?

- **Yes!** Several choices in our domestic robotics architecture are influenced by various existing cognitive architectures
- Our architecture can be seen as a **hybrid architecture** with various elements discussed today:
  - Action selection: **Task planning** for selecting sequences of actions
  - Memory:
    - Working memory in the form of a **fact database** that is updated by the robot while it is acting
    - Semantic memory in the form of an **environment ontology**
    - Rudimentary procedural memory based on which **behaviours to be executed are specified through a declarative language** (in terms of state machines)
  - Learning:
    - **Percept-to-concept association** (using deep learning, concretely for object recognition)
    - Procedural learning (**concretely for learning manipulation skills** — my own PhD work[1])
  - Metacognition (**for the analysis of execution failures** — my own PhD work)

[1] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042

Action Selection and Procedural Memory[2]



[2] A. Mitrevski, "Skill generalisation and experience acquisition for predicting and avoiding execution failures," *Ph.D. dissertation*, Department of Computer Science, RWTH Aachen University, 2023. Available: https://publications.rwth-aachen.de/record/943042