Intelligence Without Representation

Written by Rodney A. Brooks
Presented by Arun Kirubarajan

Intelligence Without Representation

- Published in September 1987 by Rodney Brooks
- Tries to answer the question of **how to design artificial intelligence systems**
- Tackles the necessity of **whether we need symbolic representations**
- Shortly before the *“AI Winter”* at the end of the 1980s
    - Reduced funding/interest in AI research
    - Increased pessimism towards existing artificial intelligence ideas

Rodney Brooks

- MIT Roboticist / Entrepreneur
- Previous CTO of *iRobot*
- Popularized the **actionist approach** to artificial intelligence

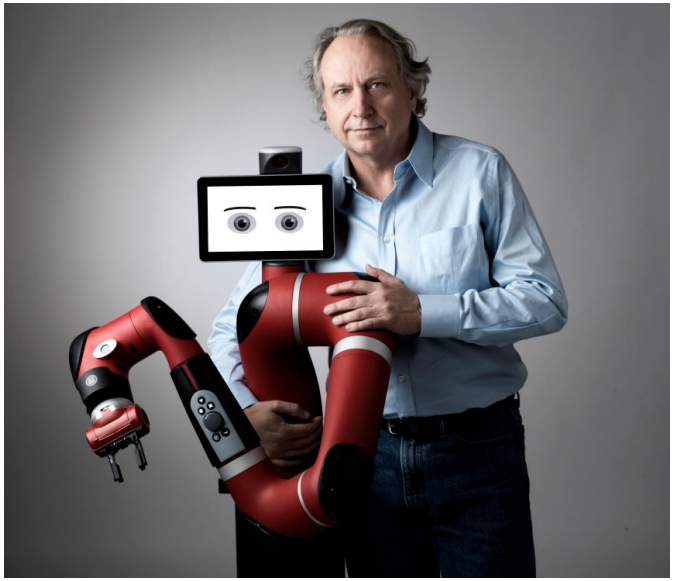
“Over time there's been a realization that vision, sound-processing, and early language are maybe the keys to how our brain is organized.”

Rodney Brooks' View of Intelligence

- Considers human beings (and many animals) to be intelligent
- Necessary components of intelligence:
  - problem solving behavior
  - language
  - expert knowledge and application
  - **reasoning**
- Components are simple once the *essence of being and reacting* are available
  - What is the evidence of this?

Inspiration from Evolution (according to Brooks)

- Intelligent creatures were once non-intelligent and their intelligence evolved
- Evolution did not occur out of a centralized knowledge base (representation)
    - Occurred by interacting with the world
    - Sensing and executing actions



Are these robots *intelligent*?

Rodney Brooks does! In this presentation, we will explore **why**.

View of Artificial Intelligence in the *1980s*

- Isolated and disjoint sub-problems
  - e.g. natural language understanding, plan verification, computer vision
- Some researchers are still optimistic about the possibility of a unified view of artificial intelligence
- Researchers still divided over symbolic representations for artificial intelligence

- Brooks believes that intelligence is too large and misunderstood to be broken down into appropriate sub-problems
- Overall, division of labour is good for practical applications, but doesn't move towards general intelligence

“Amongst the dreamers still in the field of AI (those not dreaming about dollars, that is), there is a feeling that one day all these pieces will all fall into place and we will see "truly" intelligent systems emerge.”

- **Rodney Brooks**

Abstract

“When intelligence is approached in an incremental manner, with strict reliance on interfacing to the real world through perception and action, reliance on representation disappears.”

The authors believe that intelligence shouldn’t be decomposed into independent information processing modules that interface with each other through some representation. Instead, they believe that it should be decomposed into parallel activity producers, **which interface directly to the real world**.

In parallel, actuators propel the agent while still sensing for collisions.

Upshot

“Based on these principles we have built a very successful series of mobile robots which operate without supervision as Creatures in standard office environments.”

The authors value world application. Their interpretation of intelligence makes it easier to develop fruitful intelligent agents that act directly, without being concerned with adequate central representation.

Representation

What is it?

My working definition:

How an agent internally stores some knowledge or concept about the world or environment.

Symbolic Representation

GOFAI

Representation that involves the use and manipulation of symbols to represent entities in the real world. Usually involves logical operations and constraints of represented entities.

Dominant paradigm in the AI community from the post-War era until the late 1980s.

An example of symbolic representation

CAN (SIT-ON PERSON CHAIR)), (CAN(STAND-ON PERSON CHAIR))

Discussion

- What is this a representation of?
  - Is this an adequate form of representation?
- What are the benefits to this type of representation?
  - Would you trust a system that relies on representations such as this?

Pros and Cons of Symbolic AI

Advantages:

- Consistency in execution
- Explicit representations
- Logic can be easily understood
    - Creates explainable systems

Disadvantages:

- Learning process (rules and knowledge) has to be hand coded which is a time-consuming problem
- Difficult to embed common-sense into knowledge bases
    - Harder to deploy into the real world

What's an example of **Non-Symbolic AI**?
What are the pros and cons?

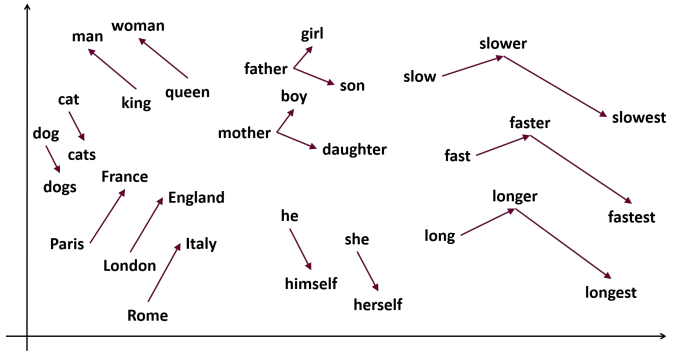
An example of non-symbolic representation

[0.76702922 -0.95673743 0.49207258 0. 0.08947326 -0.24245257]

Discussion

- What is this a representation of?
  - Is this an adequate form of representation?
- What are the benefits to this type of representation?
  - Would you trust a system that relies on representations such as this?

Currently, language representation is a **big** area!



So why has representation been such a big deal in artificial intelligence research lately?

According to Brooks, it's because representation learning is a good *"interface between isolated modules* and conference papers"*.

*this isn't true anymore...

So what's the *right* way to approach intelligence?

Rodney’s path for achieving true intelligence: building autonomous ***Creatures***.

Definition of ***Creatures***

What's the most obvious example of a *Creature?*

What would the analogous concept be for natural language systems?

1) A Creature must cope appropriately and in a timely fashion with changes in its dynamic environment.

2) A Creature should be robust with respect to its environment; minor changes in the properties of the world should not lead to total collapse of the Creature's behavior;

3) A Creature should be able to maintain multiple goals and, depending on the circumstances it finds itself in, change which particular goals it is actively pursuing;

4) A Creature should do something in the world; it should have some purpose in being.

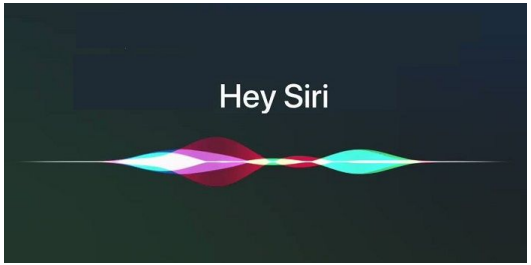
Chatbots (NLU) as a *Creature*

A Creature must cope appropriately and in a timely fashion with changes in its dynamic environment.

- Ability to respond to a wide variety of inputs

A Creature should be robust with respect to its environment.

- *Consistency in responses, irrespective of inputs*



A Creature should be able to maintain multiple goals and, depending on the circumstances it finds itself in, change which particular goals it is actively pursuing.

- Ability to be task-driven or open chit-chat

A Creature should do something in the world; it should have some purpose in being.

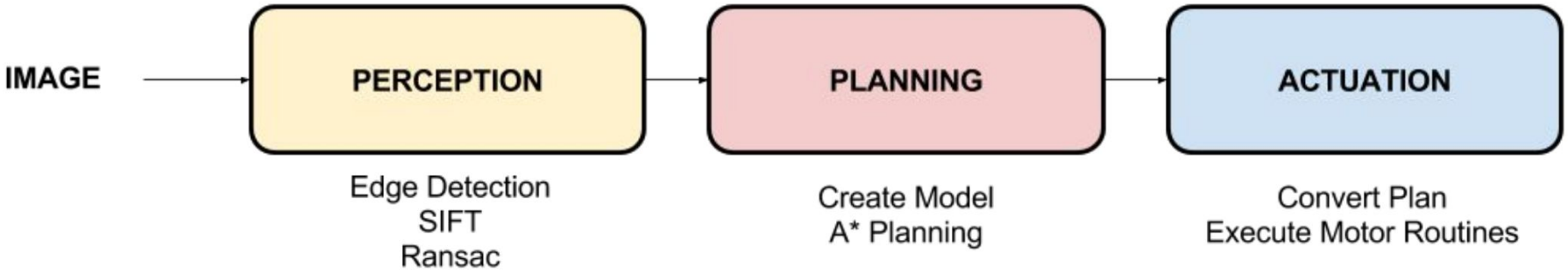
- Should be coherent/interesting text!

How do we design such Creatures?

Brooks introduces the **Subsumption Architecture**. He expands upon this in his paper *Elephants Don't Play Chess* (1990).

Subsumption Architecture

- Decomposition of activities into task-achieving layers
- Each layer is a static topology of finite state machines (FSM)
  - Explicit purposes (unlike neural networks)
- Layers are connected through **suppression** and **inhibition** mechanisms
- Emphasis on distributive and parallel control
  - Integrates perception, control, and action similarly to animals.



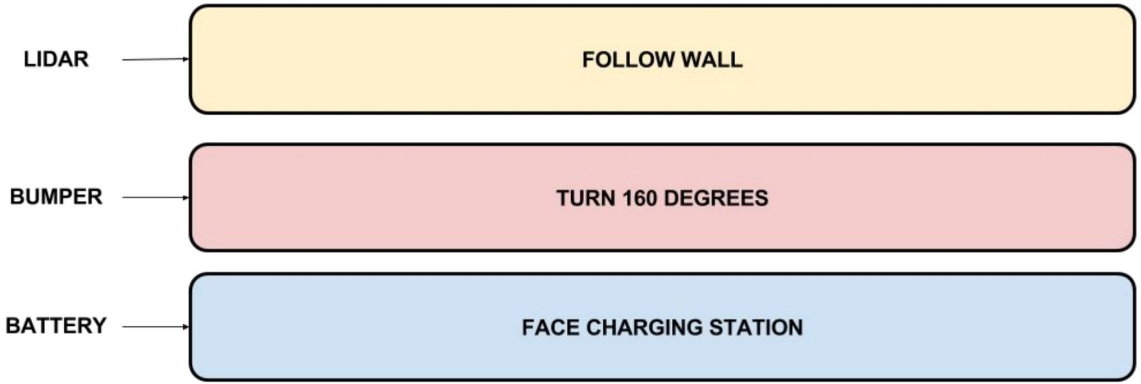


Image Credit: Alex Kuefler & Rachel Luo

Subsumption Architecture

Suppression

Messages arriving on the new layer suppress messages incoming from existing layers.

Inhibition

Messages on the new layer inhibit messages being emitted on existing layers.

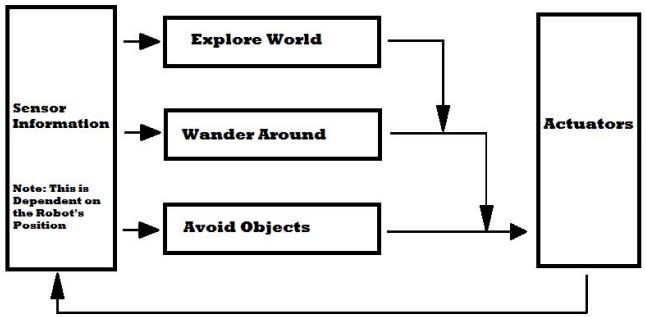


Image Credit: Wikipedia

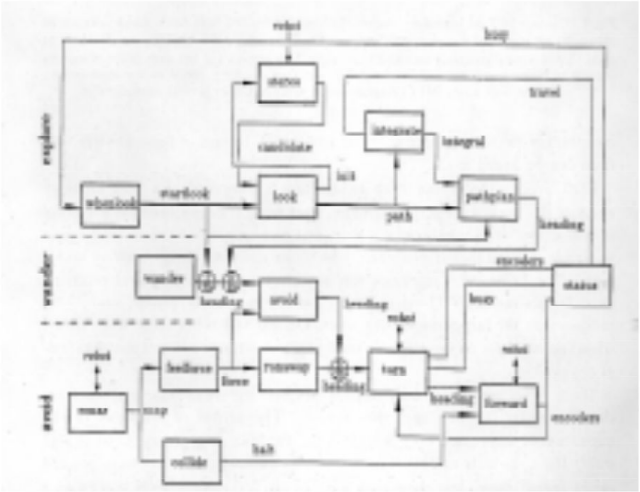
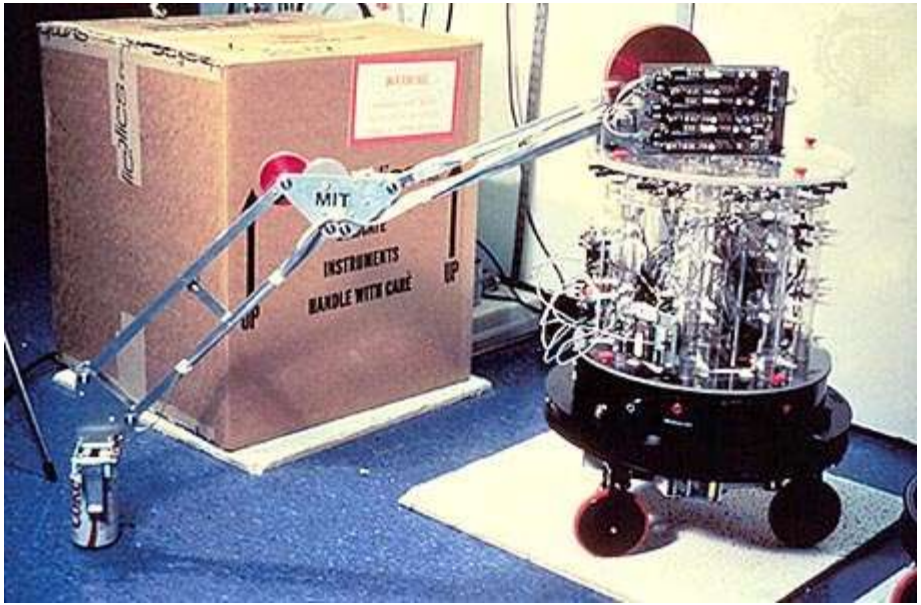


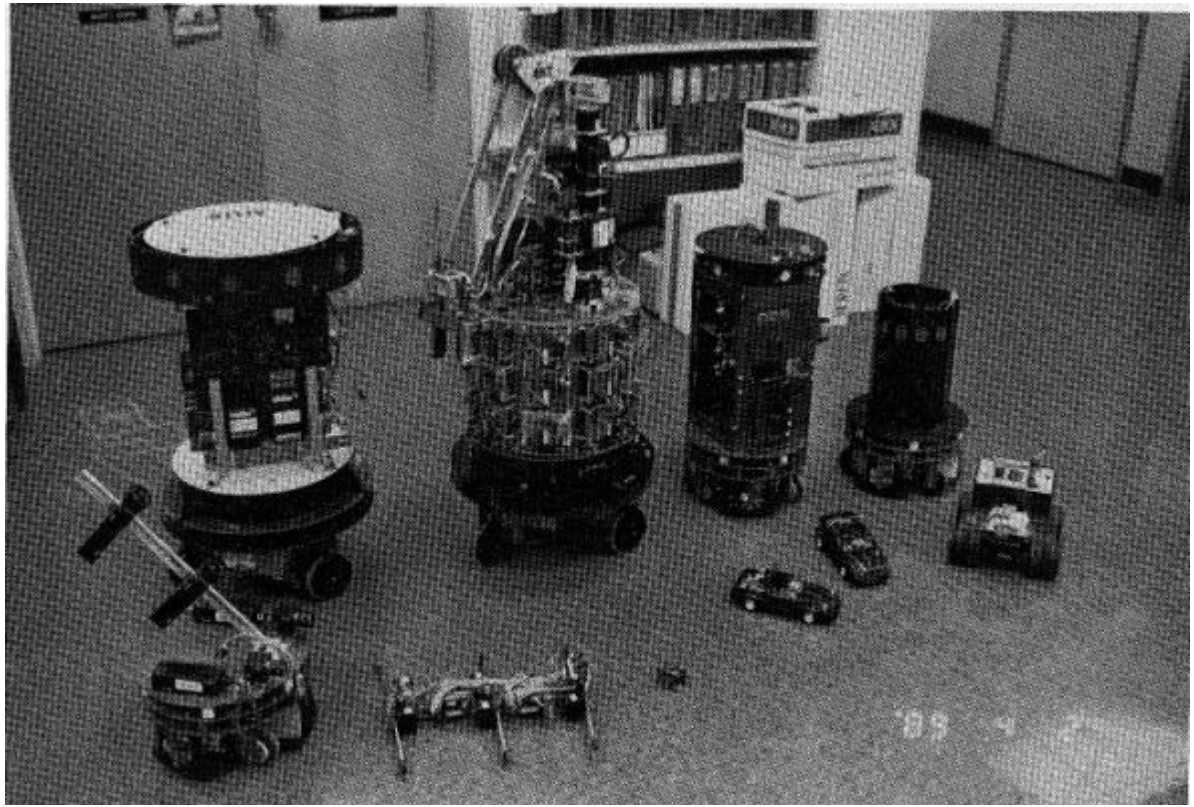
Fig. 2. We wire, finite state machines together into layers of control. Each layer is built on top of existing layers. Lower level layers never rely on the existence of higher level layers.

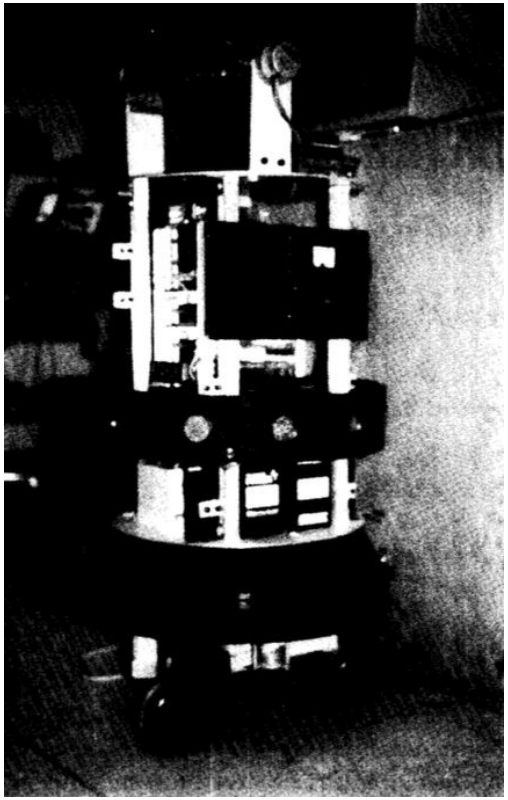
See: **end-to-end principle** in network design



Existed at MIT!

What kind of representation did it use?





What Subsumption is **NOT** (according to Brooks)

1. Connectionism
   a. No distributed representations
2. Neural Networks
   a. No biological inspirations
3. Production Rules
   a. No explicit if/else statements
4. “German Philosophy”
   a. Similar to Martin Heidegger, but intended to be a real/deployed system

Unexpected Conclusion and Radical Hypothesis

Unexpected Conclusion:

“When we examine very simple level intelligence we find that explicit representations and models of the world simply get in the way. It turns out to be better to use the world as its own model.”

Radical Hypothesis:

“Representation is the wrong unit of abstraction in building the bulkiest parts of intelligent systems.”

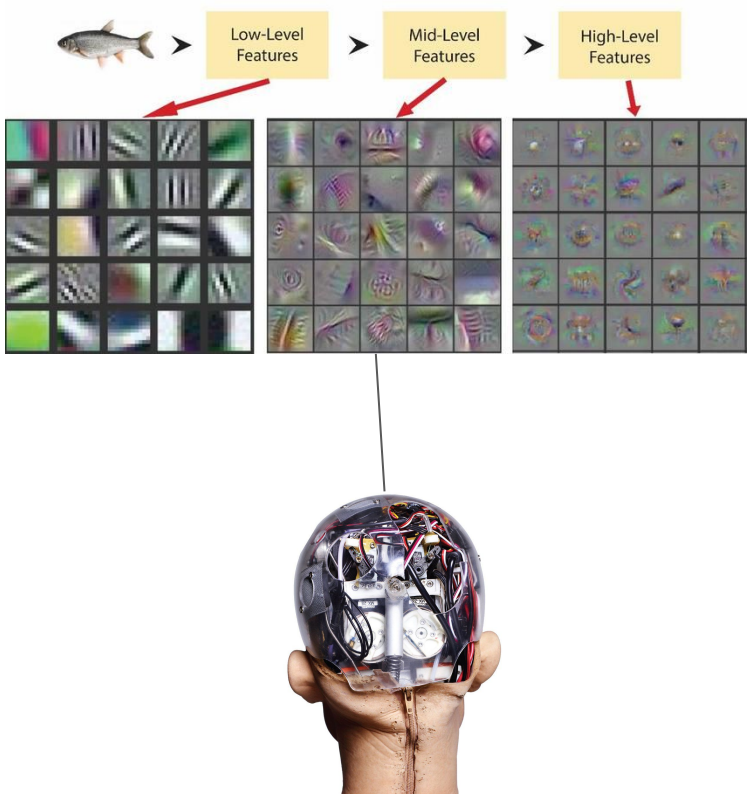
My Personal Take

Representations are still important, especially for NLU systems.

Representation is harder to ignore in NLU tasks since representation of characters/words has to be considered for all systems.

One of Brooks' criticisms is that symbolic representations are universal and static, but recent developments in contextualized embeddings show that representations can by dynamic.

Model interpretability is more important than ever, now that more people have access to "AI" systems.



Thank you!
Any questions?