Portions of this slideshow: Philippe Pasquier

Image credit: Kıvanç Tatar

Carlos Castellanos | School of Interactive Games & Media | Rochester Institute of Technology

Current Topics in Interactive Development - IGME 480

RIT

Outline of Today's Lecture

- Autonomous Agents

Autonomous Agents & Multi-Agent Systems

What is an Agent?

- An **entity** which is placed in an **environment** and senses different **parameters** that are used to make a decision based on the goal of the entity. The entity performs the necessary **action** on the environment based on this decision. (Dorri et al., 2018)
- A computer system capable of autonomous action in some environment in order to meet some goal(s)

Autonomous Agents & Multi-Agent Systems

What are Multi-Agent Systems?

- Multiple agents that interact with their environment (including one another) are known as Multi-Agent Systems (MAS)
- MAS is a kind of Distributed Artificial Intelligence (DAI)
  - distributed problem solving and task allocation
  - solve complex problems
  - efficient
- MAS is (still) one of the dominant paradigms of AI

Autonomous Agents & Multi-Agent Systems

Agent Formalization

- What is an agent (formalized)?
  - At the extreme, an agent can be formalized as a function linking percepts to actions:
    - F_t: P* -> A

Autonomous Agents & Multi-Agent Systems

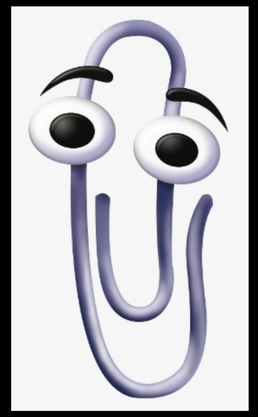
Agent Features

- Situatedness: perceiving the environment via sensors and being able to affect the environment via effectors
  - Agents can have the ability to obtain complete, accurate and up-to-date information on the environment (accessible) or incomplete or noisy information of the environment (inaccessible)
  - Environments can be deterministic (no uncertainty about the state that will result from performing an action) or non-deterministic (outcome of an action is not entirely predictable)
- Autonomy: capability of action without intervention or control over internal state
- Flexibility:
  - Responsiveness: ability to respond in a timely fashion to changes in the environment
  - Pro-activeness: do not simply act in response to their environment, they are able to exhibit goal-directed behavior by *taking the initiative*. (Wooldridge & Jennings, 1995)
- Sociability: ability to interact and share information with other agents (and possibly humans), often via some kind of communication language or protocol and usually for mutual benefit

Autonomous Agents & Multi-Agent Systems

Agents can be:

- software agents: situated in a digital information environment (disembodied)
- virtual agents: represented via graphical icons, animated cartoons or avatars
  - conversational agents: use written or natural language
- physically embodied agents: robots or some other physical body



Autonomous Agents & Multi-Agent Systems

Bots

- Twitter bots as art

Autonomous Agents & Multi-Agent Systems (Pt. 1)

Agent Architectures

- Three basic kinds:
  - Cognitive: maintain internal symbolic representations (of environment, other agents, etc); often involve reasoning and planning
  - Reactive: no internal representations; focus on behavioral rules or modules; typically have some internal state (but not cognitive); environmental data is generated from sensors; effectors are used to take action
  - Hybrid: mixing cognitive deliberativeness with reactiveness

Autonomous Agents & Multi-Agent Systems

Why agents?

- Agents can be thought of as having intentions, desires and beliefs. Take these statements about a human activity:
  - “Carlos took his umbrella because he *believed* it was going to rain.”
  - “InHwa *wanted* a glass of water because she was thirsty.”

“To ascribe beliefs, free will, intentions, consciousness, abilities or wants to a machine is legitimate when such an ascription expresses the same information about the machine that it expresses about a person. It is useful when the ascription helps us understand the structure of the machine, its past or future behavior, or how to repair or improve it.”

John McCarthy - McCarthy, J. (1979). “Ascribing mental qualities to machines.” In: *Philosophical perspectives in artificial intelligence*, ed. M. Ringle. Atlantic Highlands, N.J.: Humanities Press

“It is perfectly coherent to treat a light switch as a (very cooperative) agent with the capability of transmitting current at will, who invariably transmits current when it believes that we want it transmitted and not otherwise; flicking the switch is simply our way of communicating our desires’. (Shoham, 1990, p6) .”

Yoav Shoham - Shoham, Y. (1990). Agent-oriented programming. Technical Report STAN–CS–1335–90, Computer Science Department, Stanford University, Stanford, CA 9430

Autonomous Agents & Multi-Agent Systems

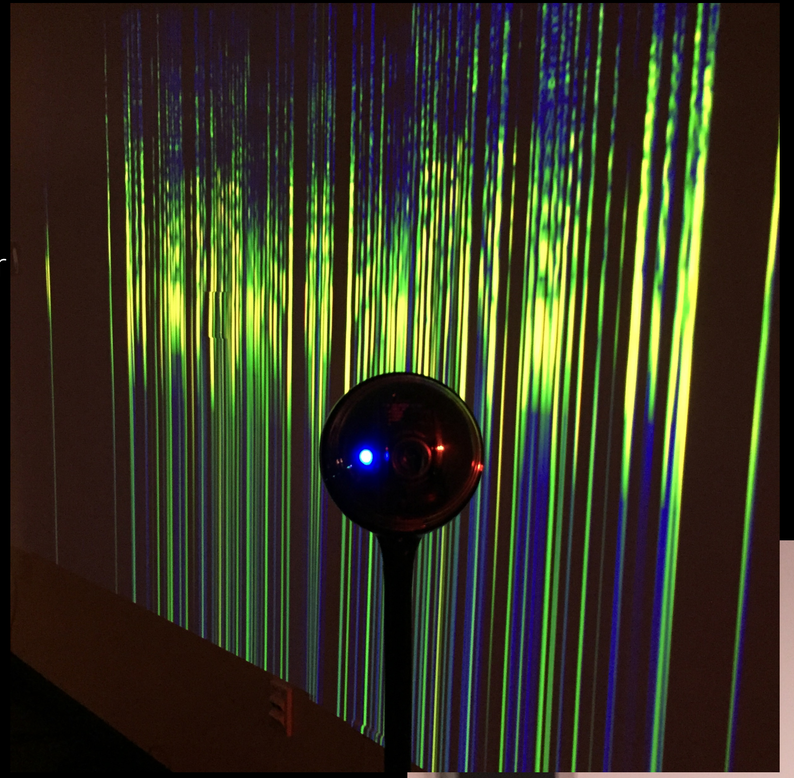
The intentional stance

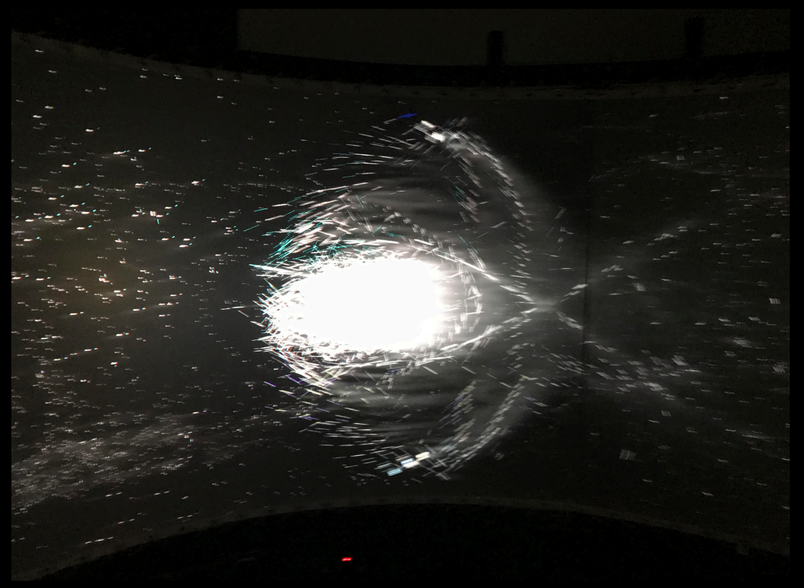
- The intentional stance is an abstraction tool
  - Cognitive architectures (and MAS more broadly) use it extensively as a way of relating to formal models
  - Like classes, objects and other abstractions, the intentional stance helps explain, understand and *program* computational systems of increasing complexity

CHATTERBOX

2020, Ronald Boersen, Aaron Liu-Rosenbaum, Kivanç Tatar, Philippe Pasquier

- Agents cluster a segmented corpus of vocal sounds in the latent space of a SOM
- trade “gibberish”
- simple rules (inspired by subsumption architecture) govern interaction between agents



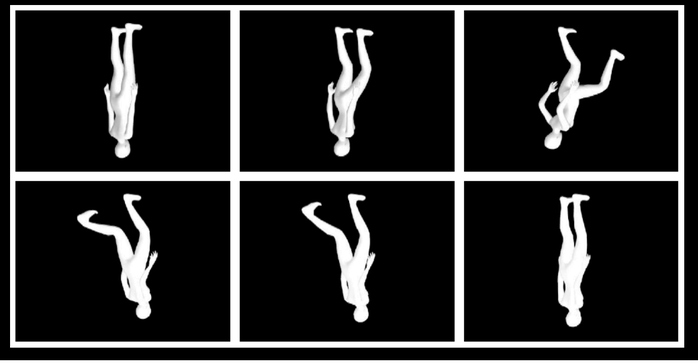


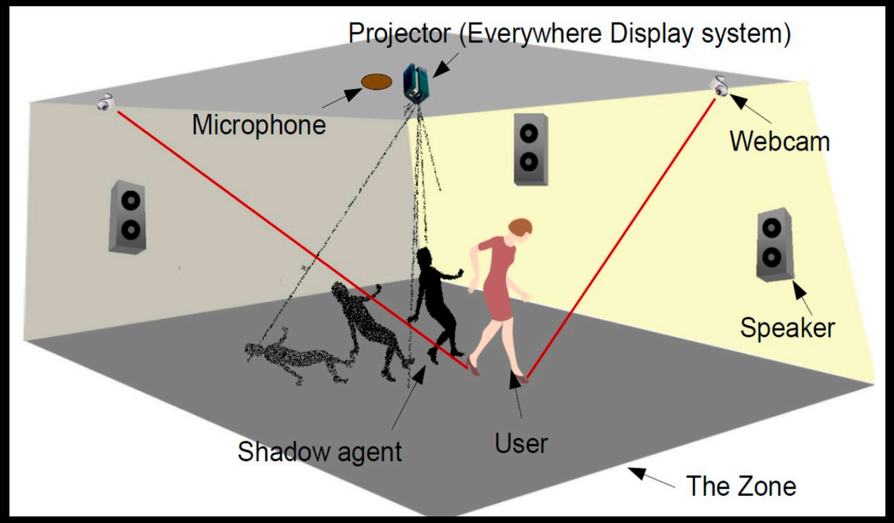
Zeta (2018), Kıvanç Tatar & Philippe Pasquier

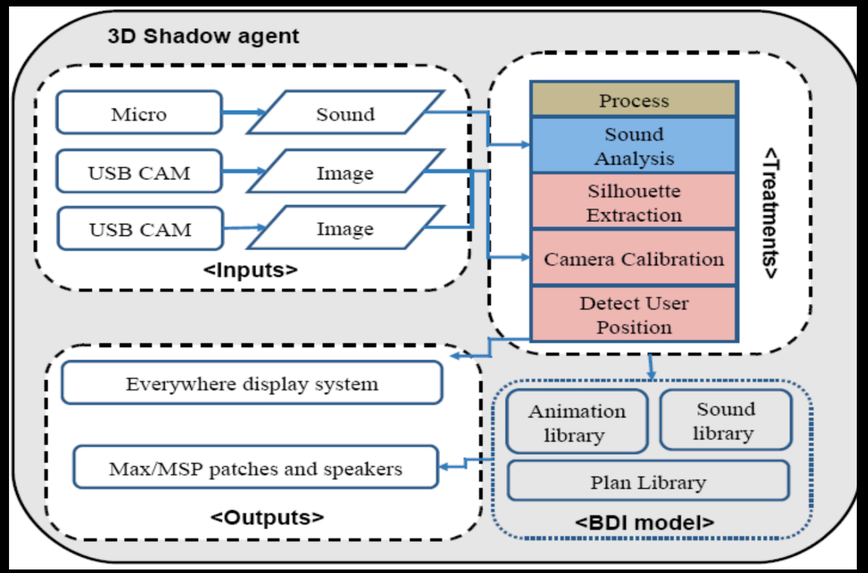
SHADOW AGENT

2008, Philippe Pasquier, Eunjung Han, Kirak Kim, Keechul Jung

- Shadow as medium
- shadow agent interacts with the human user by using pre-defined context-sensitive plans.
- BDI (Belief, Desires, Intentions) agent architecture







Autonomous Agents & Multi-agent Systems

Cognitive Agents: The BDI Architecture

- Belief, Desires and Intentions (BDI) Architecture
  - Inspired by the philosophy of human practical reasoning
  - Considers mental attitudes: beliefs, desires and intentions
  - Beliefs represent the state of the environment & the know-how of the agent
  - Desires are motivational, not necessarily consistent or achievable & are usually hard-coded
  - Intentions represent problem-solving tasks that do not conflict are persistent and are believed to be possible to achieve

Autonomous Agents & Multi-agent Systems

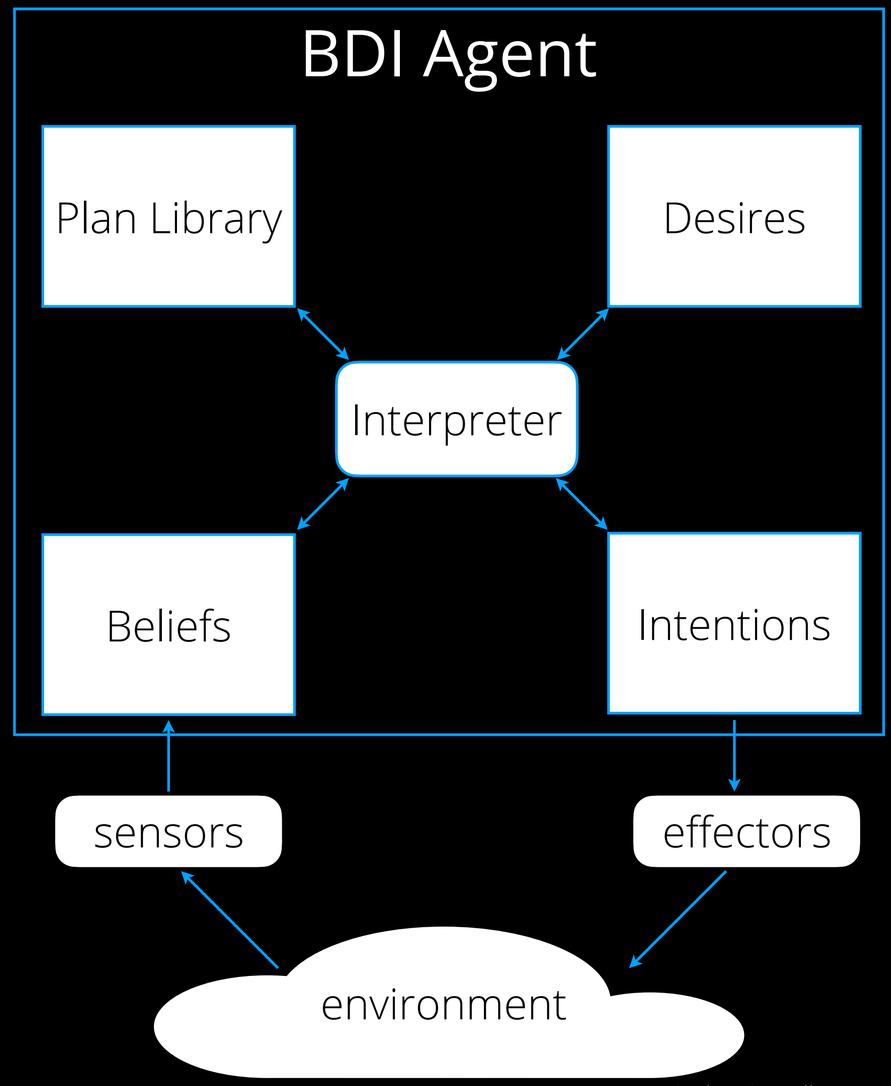
Cognitive Agents: The BDI Architecture

- BDI formalization
  - Means-end reasoning: agent generates a plan (a sequence of actions)
  - Plans can be deterministic or non-deterministic and are usually represented in a formal language
    - Deterministic actions are represented use a STRIPS style:
      - a name which may have arguments
      - a pre-condition list: facts which must be true for action to executed (usually expressed in terms of beliefs)
      - a delete list: facts that are no longer true after action is performed
      - an add list: facts made true buy the execution of an action

Autonomous Agents & Multi-agent Systems

Cognitive Agents: The BDI Architecture

- BDI Agent
  - Interpreter selects a desire to pursue as an intention (based on current beliefs)
  - Interpreter selects or generates a plan to be executed in order to achieve that intention (means-end reasoning)



Autonomous Agents & Multi-agent Systems

Cognitive Agents: The BDI Architecture

- BDI Interpreter - basic algorithm
  - (B, D, I): Initialize-state();
  - While true
    - Update(B, D, I); // according to internal and external perceptions
    - Options: option-generator(B, D, I);
    - Selected-options: deliberate(B, D, I);
    - Update-intentions(Selected-options, I);
    - Plan: Planning(I, B);
    - Execute(Plan);
    - Get-new-perceptions();
  - end While

Autonomous Agents & Multi-agent Systems

Cognitive Agents: The BDI Architecture

- Final thoughts on BDI
  - The BDI model for (rational) cognitive agents is a paradigmatic approach and one of the most advanced ones: it gives you a flavor of the agent world!
  - Not much exploited in computational creativity and the arts

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

- There are many unsolved (some would say unsolvable) problems associated with symbolic AI (based on knowledge representation)
- These problems have led some researchers to question the viability of the whole paradigm, and to the development of reactive architectures
- Robiticist/AI researcher Rodney Brooks put forward three theses:
  - Intelligent behavior can be generated *without explicit representations* of the kind that symbolic AI proposes
  - Intelligent behavior can be generated *without explicit abstract reasoning* of the kind that symbolic AI proposes
  - Intelligence is an *emergent property* of certain complex systems
- “Elephants Don’t Play Chess"

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

- Brooks identified three key ideas that have informed his research:
    1. Situatedness and embodiment: ‘Real’ intelligence is situated in the world, not in disembodied systems such as theorem provers or expert systems.
    2. Intelligence is ‘in the eye of the beholder’; it is not an innate, isolated property.
    3. Intelligence and emergence: ‘Intelligent’ behavior arises as a result of an agent’s interaction with its environment.

Autonomous Agents & Multi-agent Systems

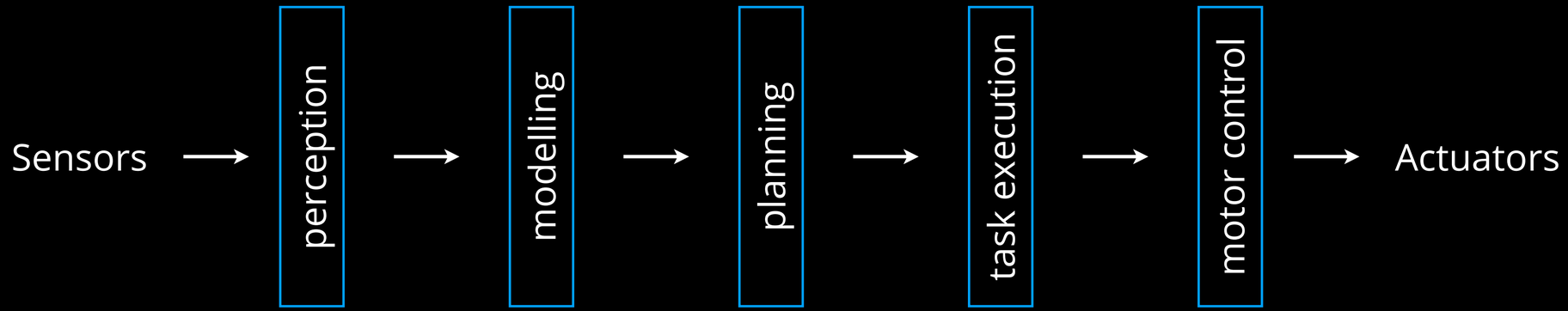
Reactive Agents: The Subsumption Architecture

- Subsumption Architecture
  - A hierarchy or layered set of task-accomplishing behaviors
  - Each behavior is a rather simple rule-like structure (e.g. avoid object)
  - Each behavior ‘competes’ with others to exercise control over the agent
  - Lower layers represent more primitive kinds of behavior (thought to be critical, such as avoiding obstacles), and have precedence over layers further up the hierarchy
  - The resulting systems are, in terms of the amount of computation they do, extremely simple

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

- A traditional decomposition of a mobile robot control system into functional modules:
  - Note the similarity to the BDI architecture



from Brooks (1986)

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

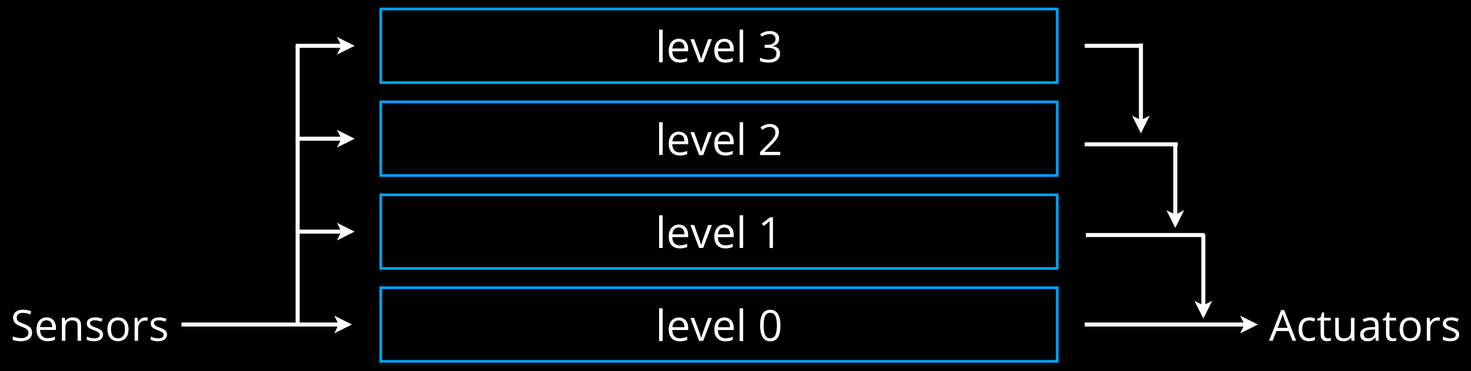
- A decomposition of a mobile robot control system based upon task-achieving behaviors:

from Brooks (1986)

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

- Layered control in the subsumption architecture:
  - The system can be partitioned at any level, the layers below will form an operational system
  - no central controller to coordinate all possible functions in the robot

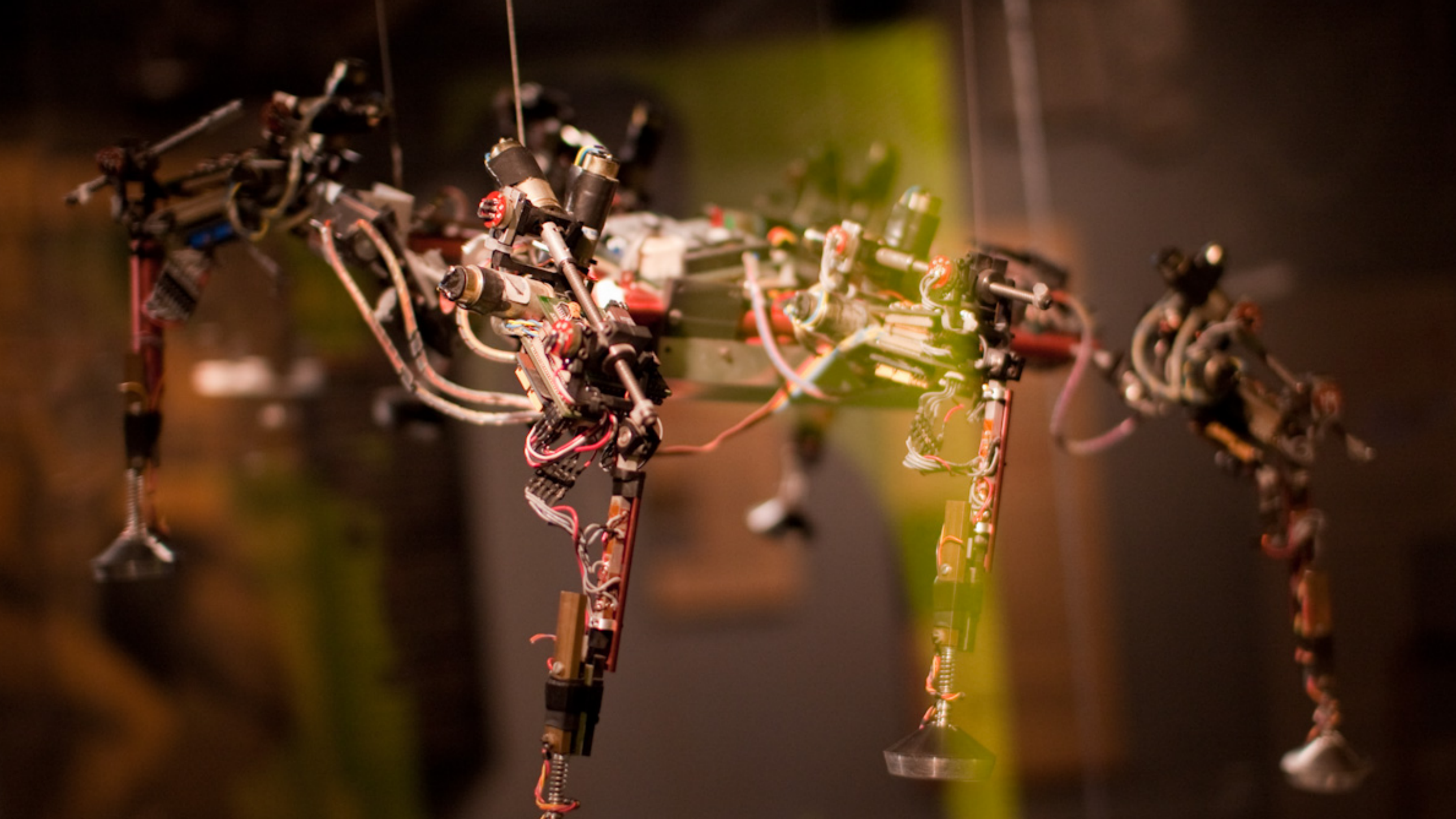


from Brooks (1986)

THE SUBSUMPTION ARCHITECTURE

Example: Genghis

- 30cm long, 6-legged insect-like robot, designed by Rodney Brooks




baxter

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

- Advantages of Reactive Architectures:
  - Simplicity
  - Economy
  - Computational tractability
  - Robustness against failure
  - Potential for complex, emergent group behavior
- Can achieve tasks that would be considered very impressive using symbolic AI methods (why not use them for creative tasks?)

Autonomous Agents & Multi-agent Systems

Reactive Agents: The Subsumption Architecture

- Limitations of Reactive Architectures:
  - Agents must be able to map local knowledge to appropriate action
  - Impossible to take non-local (or long-term) information into account
  - If it works, how do we know why it works? The departure from “knowledge level” implies a loss of transparency and readability
  - What if it doesn’t work? Purely reactive systems are typically hard to debug
  - Engineering emergence: lack of clear design methodology
  - Design becomes difficult with more than a few rules (dynamic interactions between rules become quickly too complex)
  - How about communication with humans?

Autonomous Agents & Multi-agent Systems

Robotic Art

- *CYSP 1* (1956), Nicolas Schöffer
  - first cybernetic sculpture
  - cybernetic concept of feedback played an important role for in the sculpture's behavior.
  - Sensors picked up signals for color, light intensity, and sound intensity from the environment. The signals then controlled the sculpture's behavior.
  - Programmed with simple reactive behaviors

Autonomous Agents & Multi-agent Systems

Robotic Art

- *Ménage* (1974), Norman White
  - Multi-robot piece with 5 autonomous robots (4 on ceiling, one on floor)
  - Each robot had a rotating antenna-like scanner to detect light sources (and move towards them)
  - Simplicity behaviors led to complex group behavior

Autonomous Agents & Multi-agent Systems

Robotic Art

- *Petit Mal* (1993-95), Simon Penny
  - a robot that has “seizures”
  - design inspired by subsumption architecture
  - reactive/embodied approach to human-computer interaction
  - “un-optimized” and “under-engineered”
  - Penny stresses the culturally situated nature of intelligent agents, which he feels is often missed by mainstream AI researchers because of their emphasis on the abstract over the concrete.
  - “it is a fallacy to assume that the characteristics of an agent are in the code and are limited to what is explicitly described in the code. In fact, the opposite is much closer to the truth” (Penny 1997, pp. 105).

Deep Swamp (2018), Tega Brain

“When many work together for a goal, great things may be accomplished. It is said a lion cub was killed by a single colony of ants.”

Saskya Pandita (1182-1251)