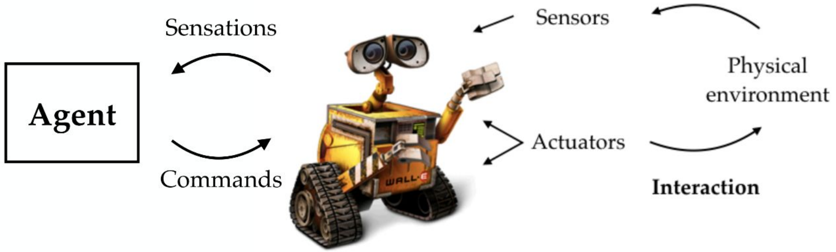
Developmental Robotics
How and why make robots like animals?

Valentin Marcel & Matej Hoffmann
Department of Cybernetics, Faculty of Electrical Engineering
Czech Technical University in Prague
valentin.marcel@fel.cvut.cz

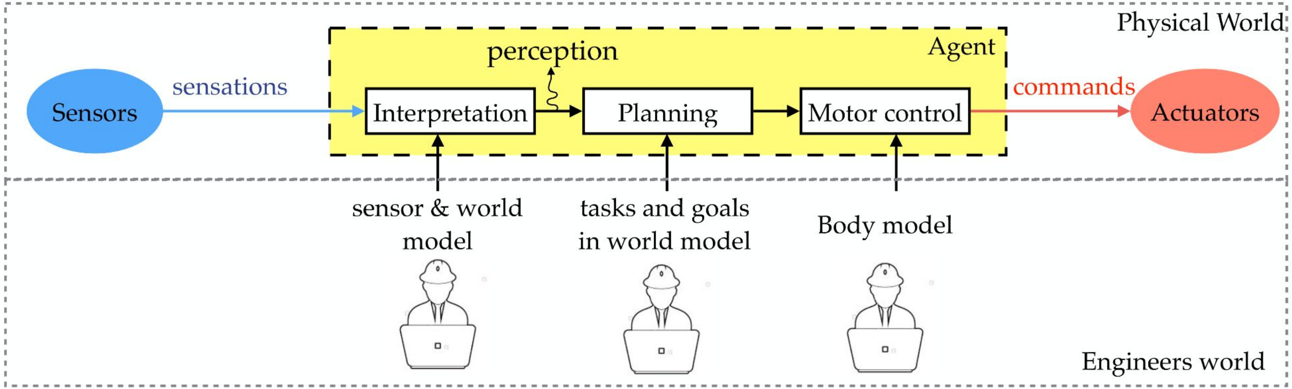
Outline

- Robotics Classical view
- What is developmental robotics
- A Body-Brain co-design, inspiration from life
- Learning-Adaptation: method and fashion
- Research in the lab

Classical view of a robotic agent



SENSE - PLAN - ACT



- Limitations

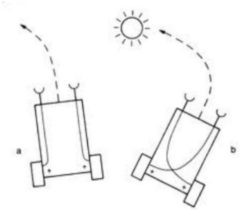
Old times

Stanford Cart, 1975

https://youtu.be/dcS6OI5xXqY

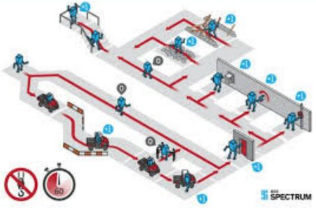
Required Knowledge

The more complex the more knowledge required



Breintenberg's vehicles

Robot floor cleaner



DARPA Challenge

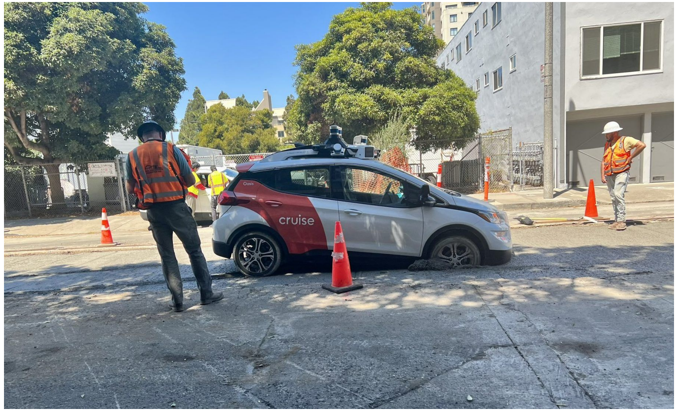
Kitty Challenge

Environment unpredictability and tasks complexity

Required Knowledge

- Knowledge about the world
- Knowledge about sensors
- Knowledge about the tasks
- Knowledge about the body

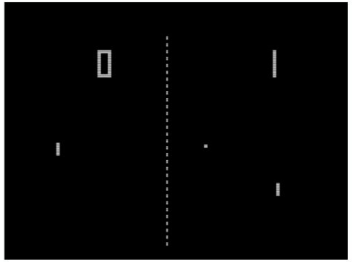
Problems



Cruise LLC self-driving cars in SF

DARPA Challenges

- Ad hoc models : specialized knowledge



Model Pong

Model Pac-Man

Model Space Invaders

⋮

+ **Experience**

Playing Atari with Deep Reinforcement Learning, Mnih et al., NIPS, 2013.

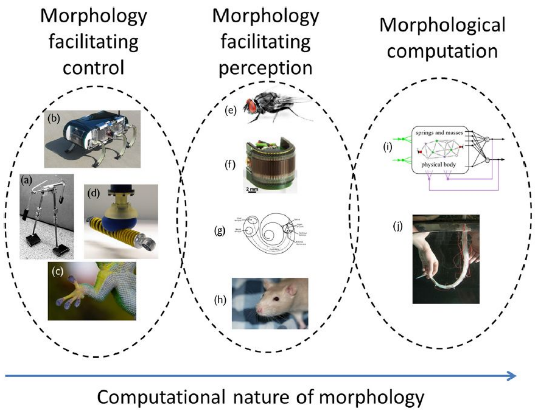
Simple observation:

Life solve these problems through:

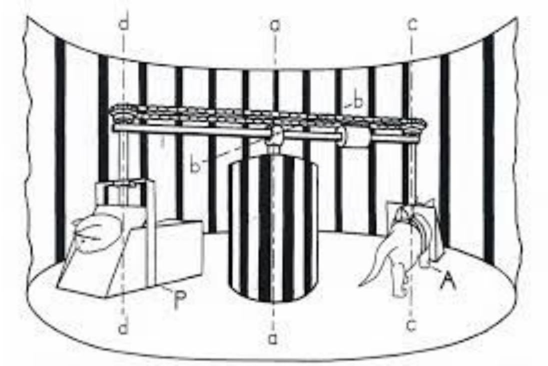
- Genetics and evolution + Environment Interaction

Genetics for body and brain architecture + Tuning mechanism + Adaptation/Social Abilities

Genetic evolution



Tuning through interaction



Held and Hein 1963

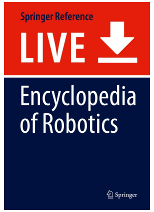
Learning by playing/teaching/imitation

What is Developmental Robotics

Developmental robotics:

Developmental Robotics is the interdisciplinary approach to the autonomous design of **behavioral** and **cognitive capabilities** in artificial agents and robots that takes direct inspiration from the developmental principles and mechanisms observed in natural cognitive systems such as children.

Angelo Cangelosi



Let’s design a fit robot

Main objective of a humanoid robot: a robot that performs human tasks

Behavior

Cognitive capabilities

Actuation

Sensing

Let’s design a fit robot

Main objective of a humanoid robot: a robot that performs human tasks

Behavior: interact with humans (speech, body language, facial expression, empathy, safe interactions…), navigate (climb stairs, enter car, move in house), manipulate (recognize object/tool, grasp and use, fine motor skills, picking up loads), service and assistance.

Cognitive capabilities: adapt to environment, understand the context, learn from experience, adapt behavior, self-maintenance, self-repair?, comply with social rules and regulation in public spaces, understands human needs (three laws of robotics by Asimov)

Actuation: Bipedal locomotion, articulated hands, facial movements. Soft but can be strong and precise. Efficient, robust, compliant, energy efficient.

Sensing: Vision, audio, tactile, proprioceptive, gravity, pleasure & pain?, and other: humidity, temperature, air quality etc. Quick, robust, energy and computation efficient!

How far are we?

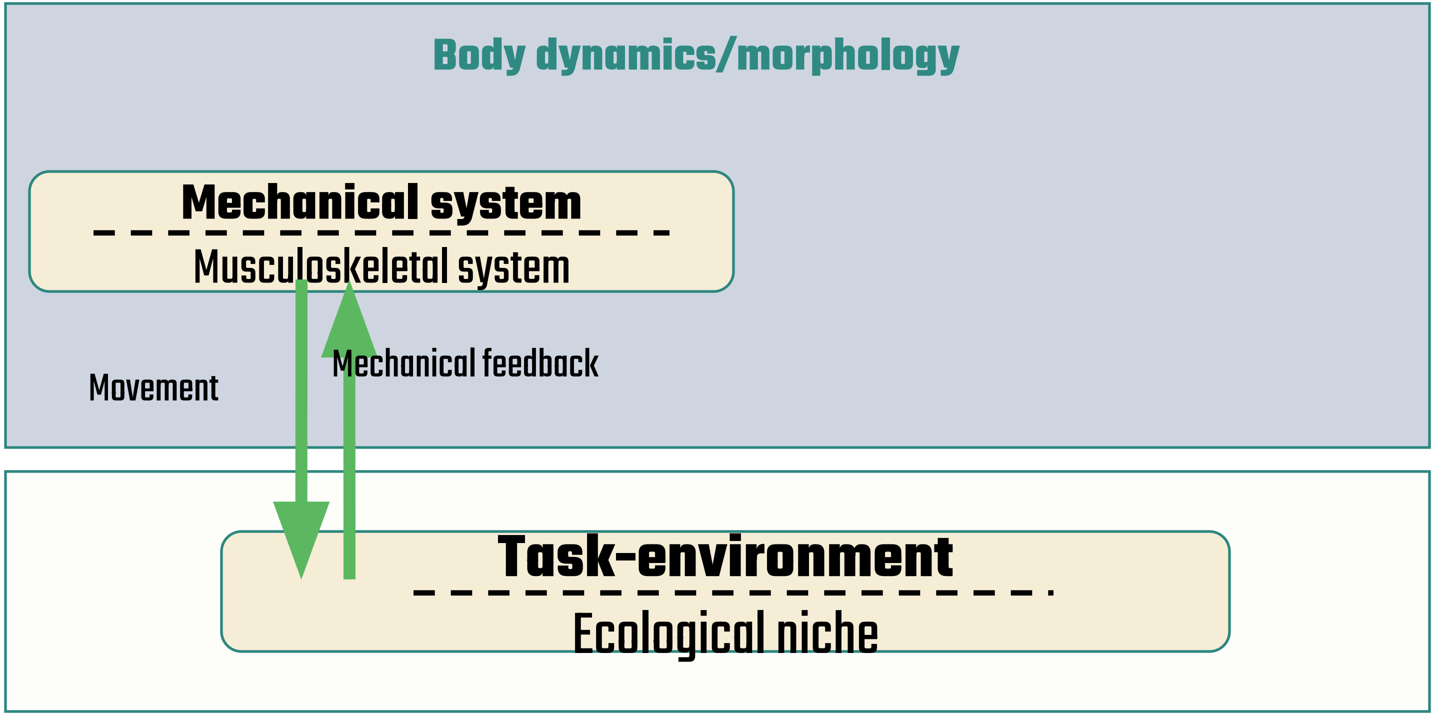
The body is part of the design problem

Morphology helps to reduce the complexity of the problem

Honda Asimo (2018)
https://youtu.be/1urL_X_vp7w

Passive Dynamic Walker – Tad McGeer (1990)
https://youtu.be/WOPED7I5Lac

Optimal control vs underactuated design



Schematics based on Pfeifer et al., Science 2007

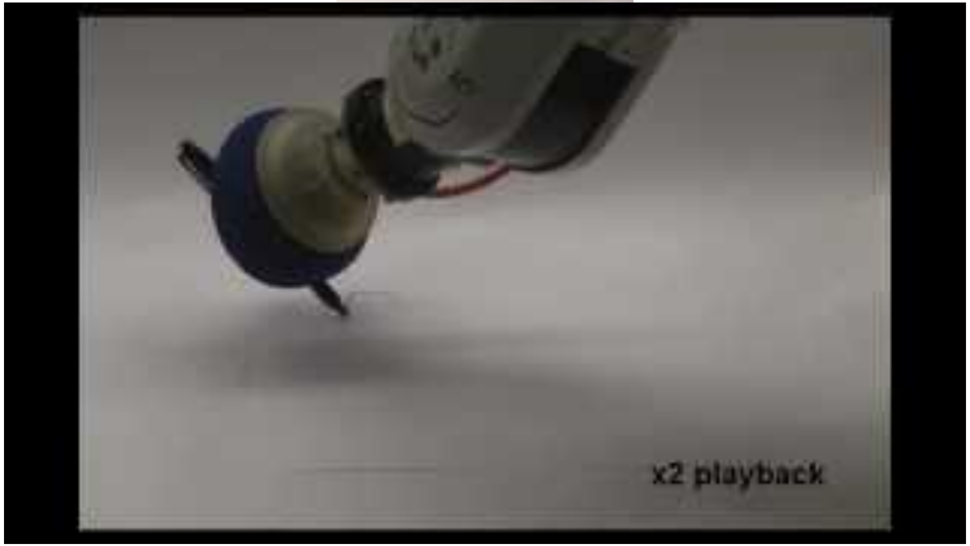
Other examples of morphology facilitating control

Self stabilization

https://youtu.be/Zt7J0dly7OM

grasping

Brown et al. 2010



https://youtu.be/ZKOI_lVDPpw

Efficient control strategies

Motor synergies/motion primitives

Modelling hands with synergies A. Bicchi *et al.* 3155

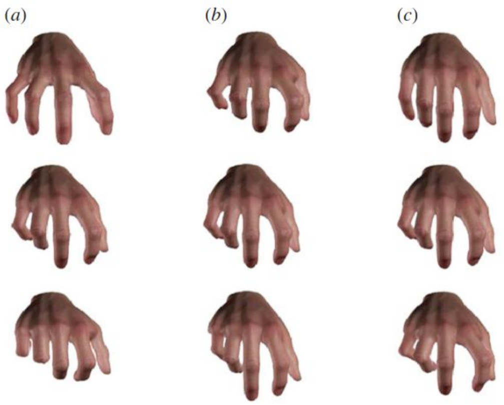
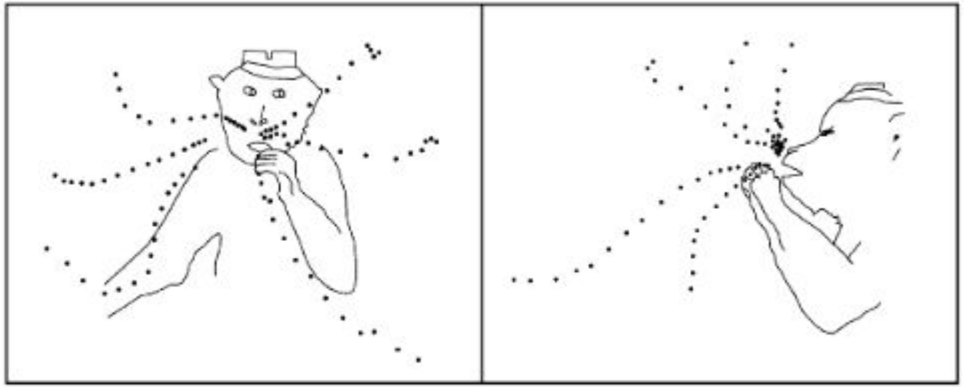


Figure 1. The three first synergies of the human hand: (*a*) the first, (*b*) second and (*c*) third synergy. Rows (top to bottom) correspond to negative, null (average) and positive intensities.



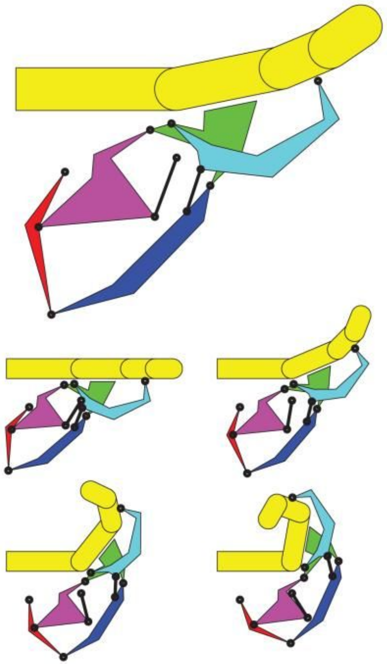
Complex Movements Evoked by Microstimulation of Precentral Cortex, Graziano et al. 2002

“12 synergies account for 80% for variation” during hand usage and grasping

Kinematic synergies of hand grasps: a comprehensive study on a large publicly available dataset

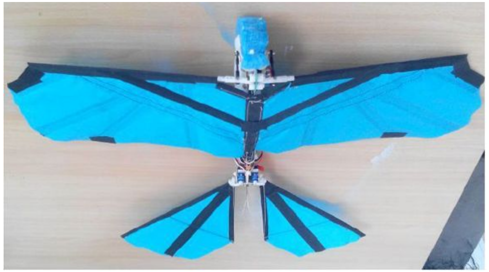
Néstor J. Jarque-Bou

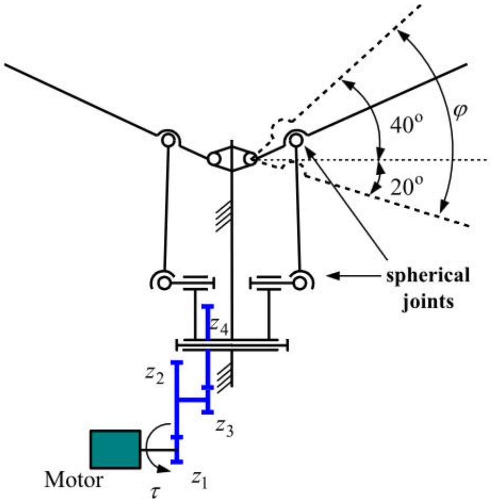
Underactuated designs



Underactuated finger

[Wolbrecht 2011]





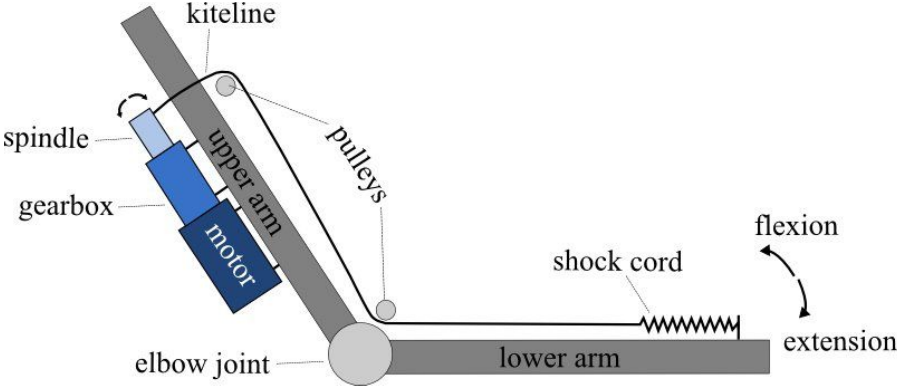
Underactuated flapping robot

[Sun at al. 2022]

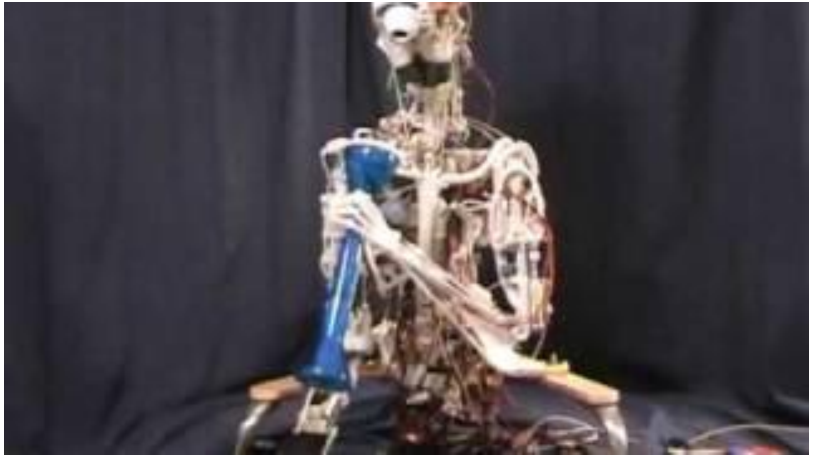
Swimming, flapping, walking, grasping, etc..

Musculoskeletal actuation

Elastic actuation

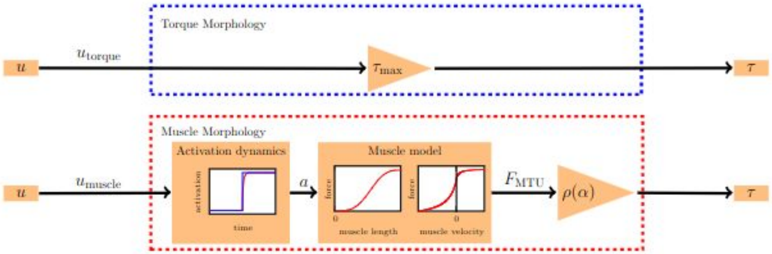


ECCEROBOT
http://eccerobot.org/home/robot/actuatorsubsystem.html

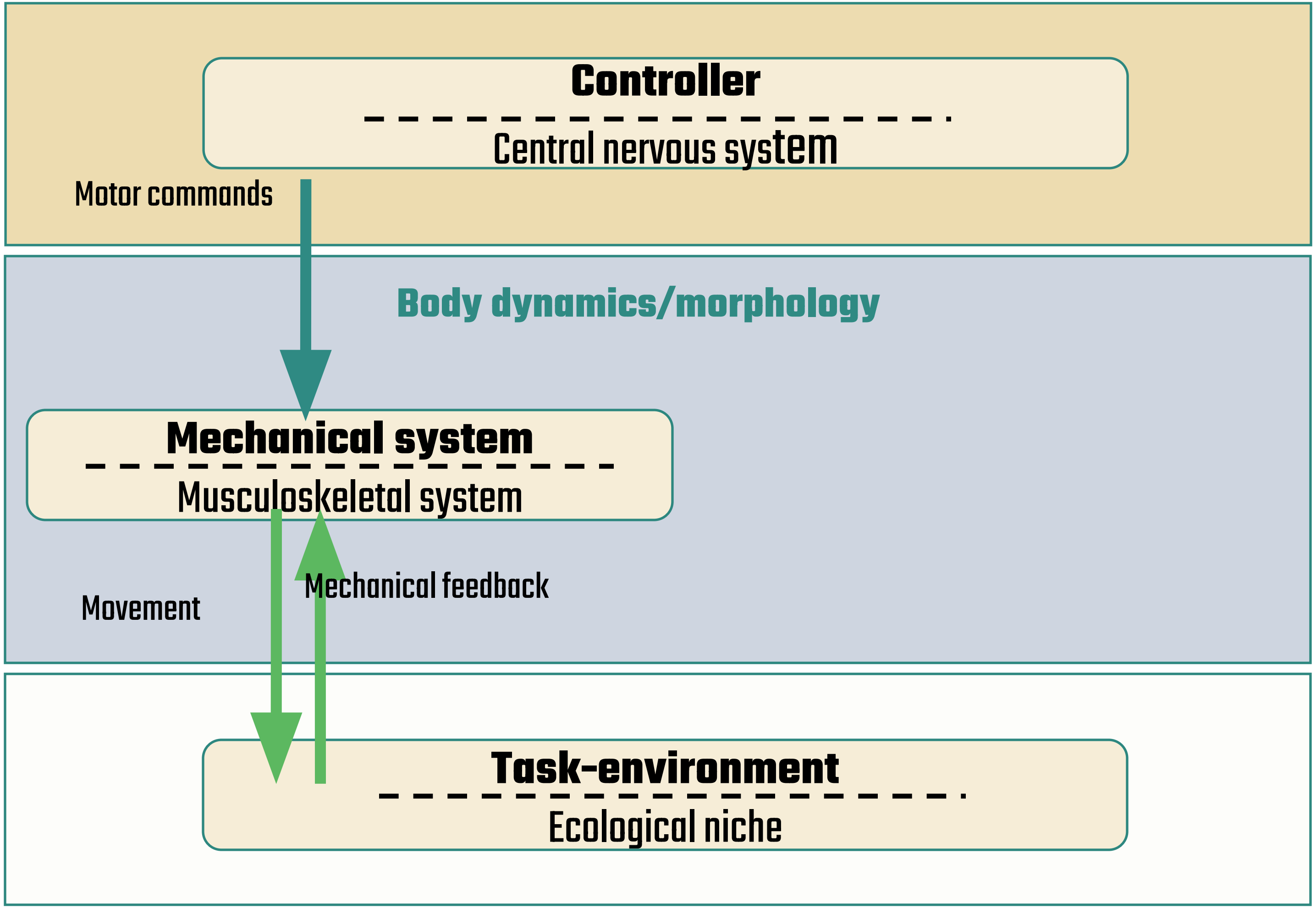


https://www.youtube.com/watch?v=cI9H4FoA0b4

Muscle models



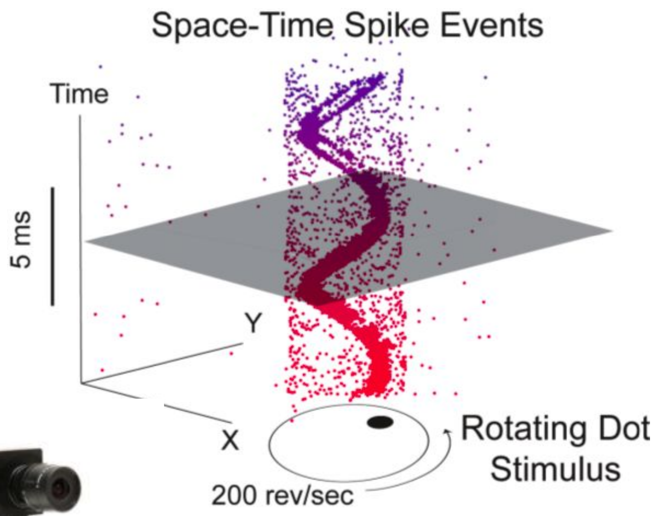
Wochner et al. 2023
Learning with Muscles: Benefits for Data-Efficiency and Robustness in Anthropomorphic Tasks,


Controller
Central nervous system
Motor commands
Body dynamics/morphology
Mechanical system
Musculoskeletal system
Movement
Mechanical feedback
Task-environment
Ecological niche

Neuromorphic/event based sensing

Only changes matter

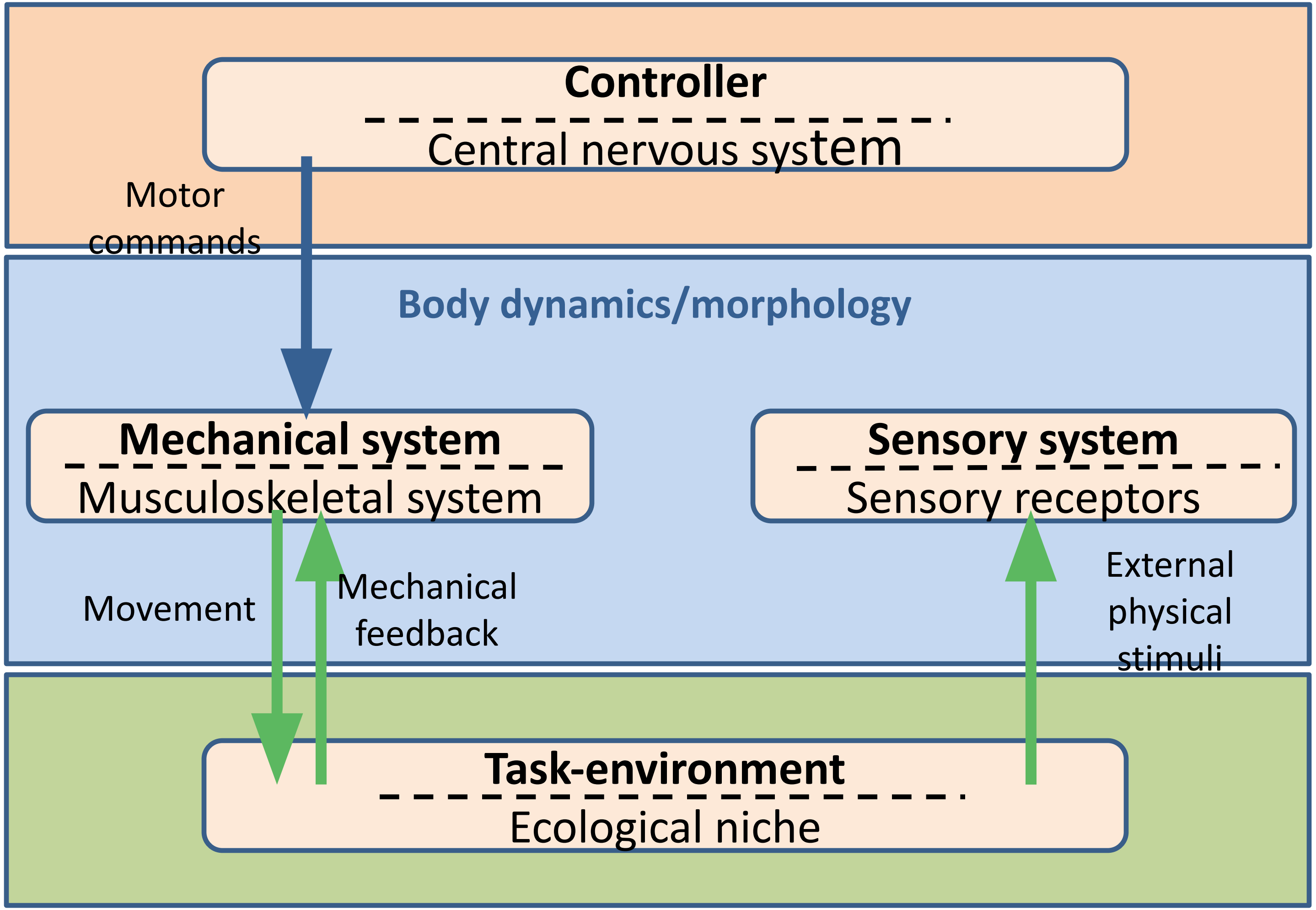
Dynamic Vision Sensor



Mueggler at al. 2015

Davide Scaramuzza (ETH Zurich)
https://www.youtube.com/watch?v=6Sn9-M7qXLk

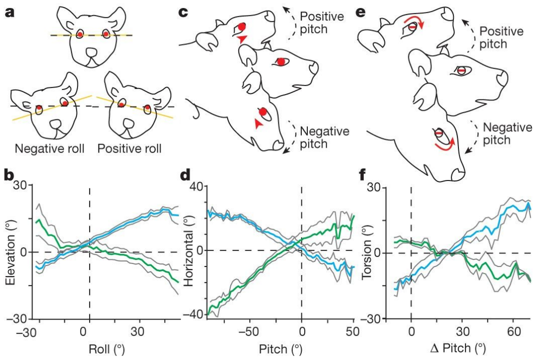
Event camera simulator:
https://rpg.ifi.uzh.ch/esim.html



Controller
Central nervous system
Motor commands
Body dynamics/morphology
Mechanical system
Musculoskeletal system
Sensory system
Sensory receptors
Movement
Mechanical feedback
External physical stimuli
Task-environment
Ecological niche

Coupling sensing and actuation *Reducing complexity*

Stabilization:



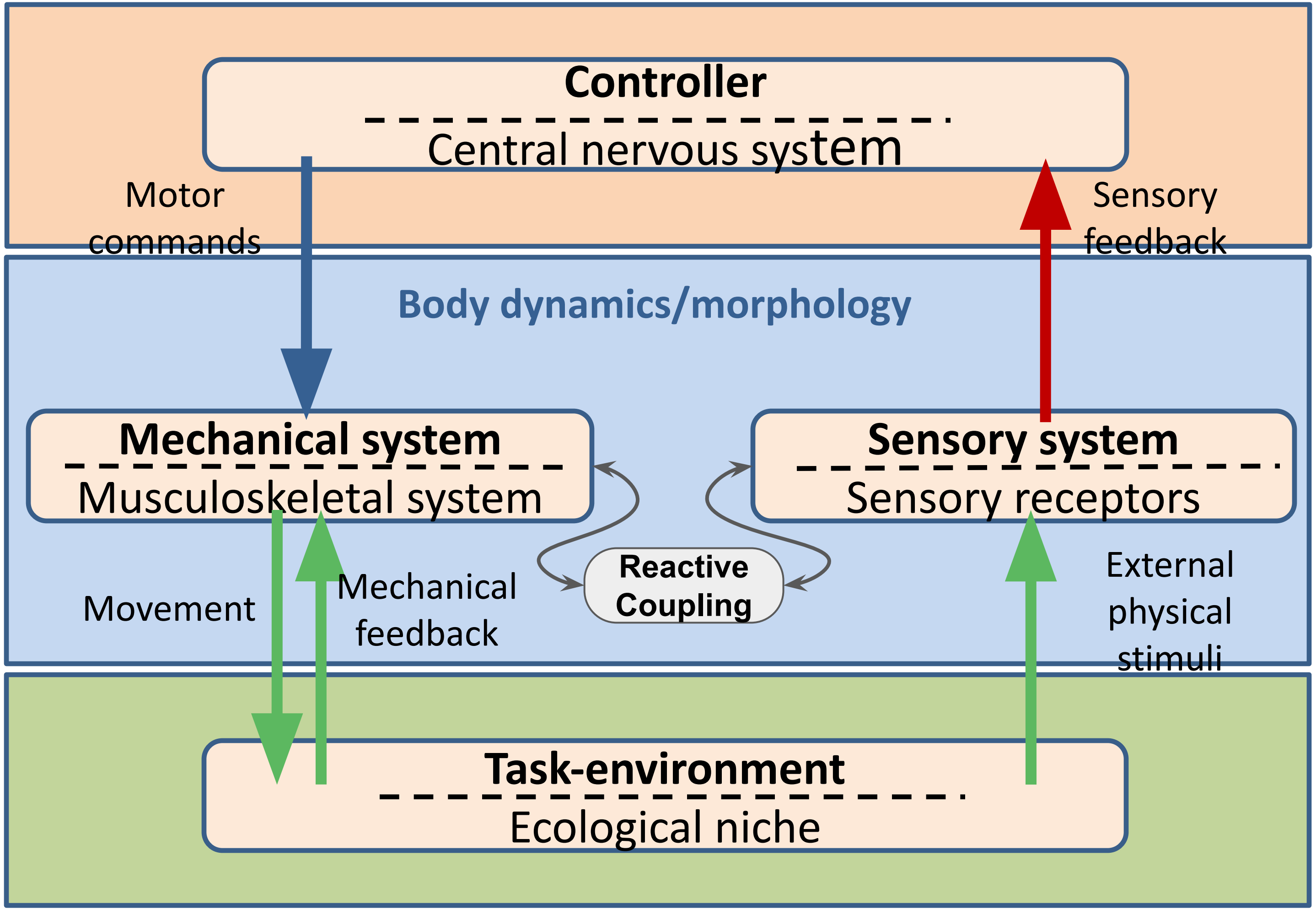
Rats maintain an overhead binocular field at the expense of constant fusion

Wallace et al. 2013

Try on yourself:

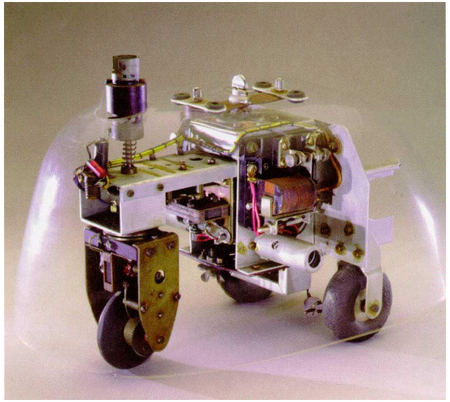
Eye convergence + smooth pursuit + head turn compensation.

Note: Infant develop smooth pursuit only after few months of development.


Controller
Central nervous system
Motor commands
Sensory feedback
Body dynamics/morphology
Mechanical system
Musculoskeletal system
Sensory system
Sensory receptors
Reactive Coupling
Movement
Mechanical feedback
External physical stimuli
Task-environment
Ecological niche

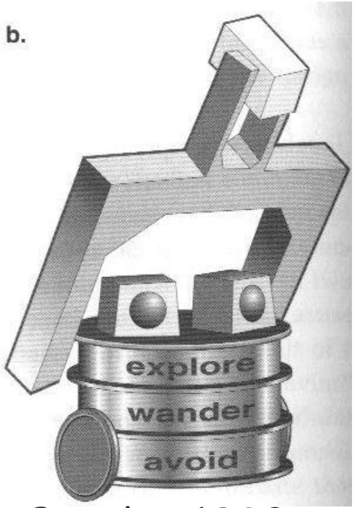
What about Behavior

Emergent from simple sensorimotor loops



Grey Walter Turtle, 1940s

V. Breitenberg, 1980s



R. Brooks, 1980s subsumption architecture

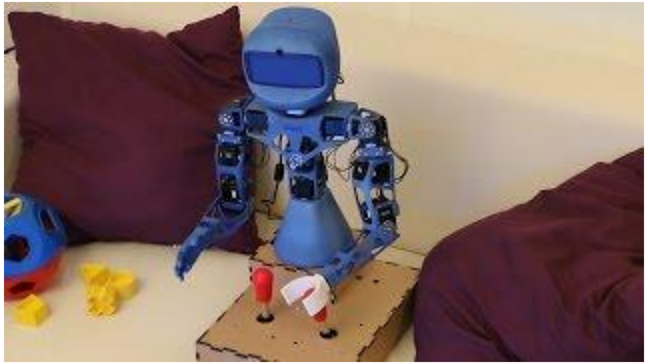
Grey Walter Tortoise

Reactive agent: Exploration - Interaction

https://youtu.be/lLULRlmXkKo

Learning/Adaptation

Intrinsic goals (open - ended) vs extrinsic teaching (end-to-end)



Forestier et al. 2017

Task agnostic
Curriculum learning
(hierarchical)

Kalashnikov 2021

End-to-end learning
Huge datasets
Sim2real or Robot farms

New methods - Embodied AI

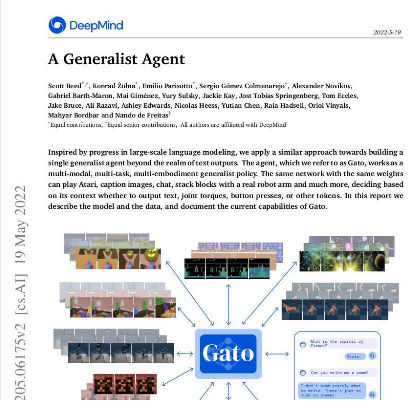
Unsupervised reinforcement-learning

Eysenbach et al. 2018 Diversity is all you need: https://sites.google.com/view/diayn/

Generate policies, e.g. of **diverse** motion primitives from random exploration. Pure information-theoric objectives.

Foundations models

GATO



DeepMind 2022-5-19

A Generalist Agent

Scott Reed, Konrad Żołna, Emilio Parisotto, Sergio Gómez Colmenarejo, Alexander Novikov, Gabriel Barth-Maron, Mai Giménez, Yury Sulsky, Jackie Kay, Jost Tobias Springenberg, Tom Eccles, Jake Bruce, Ali Razavi, Ashley Edwards, Nicolas Heess, Yutian Chen, Raia Hadsell, Oriol Vinyals, Mahyar Bordbar and Nando de Freitas

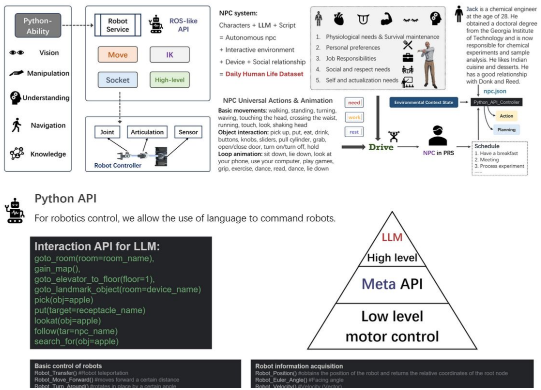
Equal contributions, Equal senior contributions, All authors are affiliated with DeepMind

Inspired by progress in large-scale language modeling, we apply a similar approach towards building a single generalist agent beyond the realm of text outputs. The agent, which we refer to as Gato, works as a multi-modal, multi-task, multi-embodiment generalist policy. The same network with the same weights can play Atari, caption images, chat, stack blocks with a real robot arm and much more, deciding based on its context whether to output text, joint torques, button presses, or other tokens. In this report we describe the model and the data, and document the current capabilities of Gato.

205.06175v2 [cs.AI] 19 May 2022

From LLM to low level

PRS challenge



Embodied AI - the new trend

Evolution: 2020-2024

[Embodied AI workshop]

2023

Foundation Models: Large pretrained models such as CLIP, ViLD and PaLI which enable few-shot and zero-shot performance on novel tasks.

Generalist Agents: Single learning methods for multiple tasks, such as RT-1, which enable models trained on one task to be expanded to novel tasks.

Sim to Real Transfer: Techniques which enable models trained in simulation to be deployed in the real world.

2024

Embodied Mobile Manipulation Many interesting embodied tasks combine manipulation and navigation to solve problems that cannot be done with either manipulation or navigation alone.

Generative AI for Embodied AI Topics such as generative AI for simulation, generative AI for data generation, and generative AI for policies (e.g., diffusion policies and world models)

Language Model Planning When we go somewhere to do something we do it for a purpose. Language model planning uses large language models (LLMs), vision-language models (VLMs), and multimodal foundation models to turn arbitrary language commands into plans and sequences for action.

Question?
Can foundation models, and LLM, solve the AI embodiment problem in robots?

AMECA robot

Yes
We can solve embodiment with engineering and general models, by going from high to low levels of interaction. In a modular way with foundation models and exploit one-shot learning strategies. **Standard** approach.

No
We first need to learn through experience and interaction of the body with the environment. We need co-design of body and brain architectures and then tune with experience. **Idiosyncratic** approach.