
MATLAB EXPO 2019

Presenter:
Ali Marjaninejad
Todays Robots

Expectation

Reality

BOSTON DYNAMICS

2005
BIG DOG

2012
LS3

2015
SPOT

2017
SPOT MINI

2009
PETMAN

2013
RHEX

2016
WILDCAT

2018
HANDLE

2018
SPOT MINI (+ARM)

www.eset.com
The answer might be in the physical structure!
Other limitations

• No model of the plant

  • A precise model of the system is not available in many scenarios

  • Even when there is a model, it will lack many details such as skin effects

  • Changes in the system

  • Contact dynamics

• No model of environment

  • Is only available for simulations or lab environment (even then, it will be with great simplifications)

  • Will not be applicable for unpredictable scenarios such as natural disasters or exploration missions

https://news.usc.edu/69355/perfecting-a-fully-functioning-prosthetic-hand/
Other limitations (continued)

- **Minimal dependency on real-time feedback**
  - Real-time feedback is not available in many scenarios including biological systems
  - Systems that heavily rely on error-correction are prone to instability and can consume lots of power

- **Data/time efficiency**
  - Data/time limitations in physical world are strict
  - Opportunity Cost
  - Evolutionary pressure

Hoffman et al., 2008
Youtube.com/Alltime10s
Problem statement

• Producing autonomous functional movements in a tendon-driven system

• With limited experience

• Without any prior model or simulation of the system or the environment

• Without any real-time feedback
How did we solve this?

- 3 tendons
- 2 DoFs
- Back-drivable motors

Darío Urbina-Meléndez
How did we solve this?

• Two-level control structure (Hierarchical learning)

  • Lower-level
    • Create an initial inverse model using data collected from motor babbling

  • Higher-level
    • Explore a reduced set of task parameters via reinforcement learning
    • Refine the inverse model (lower-level) with every each attempt
• **G2P: Motor Babbling** (lower-level controller)
• **G2P: Reinforcement Learning** (Higher-level controller)
• G2P: Reinforcement Learning (Higher-level controller)
• **G2P: Reinforcement Learning** (Higher-level controller)
Autonomous fusion in a tendon–drive actuation system

Ali Marjaninejad
Darío Urbina-Meléndez
Brian A. Cohn
Francisco J. Valero-Cuevas
Results
Results
Results

Good-enough gets you a long way!

Colors represent the independent reinforcement runs, and match with the figure above.

- First attempt to break above the reward threshold
- Attempt which yielded the highest reward
- Attempt

Polygons show the enclosing shape for all attempts of a given replicate that yielded an above-threshold reward.
Results
What is next?
What is the added value by MATLAB to this project?

- Common among many academic disciplines
- Flawless inter-toolbox communications
- Reproducibility
- Excellent support
Acknowledgements

Darío Urbina-Meléndez

Francisco Valero-Cuevas

Brian Cohn
Acknowledgements
See, Feel, Act: Hierarchical Learning for Complex Manipulation Skills with Multi-sensory Fusion
Nima Fazeli et. al. 2019
Dexterous Manipulation with Deep Reinforcement Learning:

https://sites.google.com/view/deeprl-handmanipulation
ROBEL: RObotics BEnchmarks for Learning with low-cost robots

ROBEL's open source platforms are modular, easy to build and extend

D'Claw

D'Kitty

https://sites.google.com/view/roboticsbenchmarks
Learning Dexterous Manipulation Policies from Experience and Imitation

Vikash Kumar*, Abhishek Gupta^, Emanuel Todorov*, Sergey Livine^

*University of Washington, Seattle  ^University of California, Berkeley

International Journal of Robotics Research

Thank you!
Supplementary slides
Trajectories

One possible time history of feasible command signals

Motor 1

Motor 2

Motor 3

Kinematic trajectory
Trajectories
Table 1 | Pseudo code for the RL

while $R < \text{Reward\_threshold}$
  
  $f\_\text{bar} = \text{Uniform\_distribution}([0.15, 1])$
  
  $R = \text{execute}(F\_\text{bar})$
  
end

$F\_\text{best} = F\_\text{bar}$

$R\_\text{best} = R$

for $i = 1$ to $15$
  
  $F\_\text{bar} = \text{Normal\_distribution}(F\_\text{best}, \text{sigma}.*\text{Identity}(10))$
  
  $F\_\text{bar} = \max(\min(F\_\text{bar}, f\_M), f\_m)$
  
  $R = \text{execute}(F\_\text{bar})$
  
if $R > R\_\text{best}$
  
  $R\_\text{best} = R$
  
  $F\_\text{best} = F\_\text{bar}$
  
  $\text{sigma} = (a - R\_\text{best})/b$
  
end

end
• **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them

• **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process

- Robustness to delays and noise in sensory signal
- Robustness to unmodeled dynamics
- Minimal reliance on feedback
- Generalizable to different designs
- Enhances both performance and learning

- Minimalistic approach (joint angle readings only)
- Tendon-driven (2-DoF 3-tendons)
• **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them

  • **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process
• **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them

• **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process
• **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them

• **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process
Results:

Simple Kinematic Feedback Enhances Autonomous Learning in Bio-Inspired Tendon-Driven Systems

Physical System Demonstrations
Physical system results:
Physical system results:
Simulation results:
Simulation results:
Results (cntd.):

- Enhanced accuracy in all experiments

- Robust to delays

\[\text{mean error (rads)}\]

open-loop

close-loop

\[\text{cyclical (sim) point-to-point (sim) cycle period (sim) cyclical (phys) point-to-point (phys) cycle period (phys) with contact refinements (w/ shorter babbling)}\]