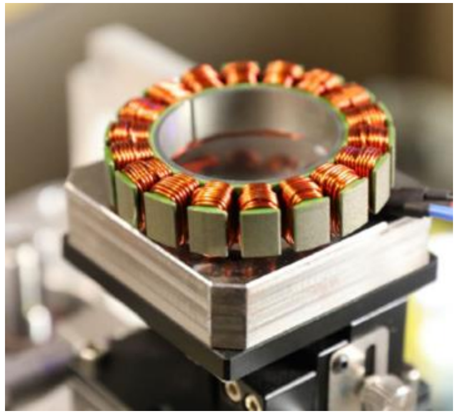
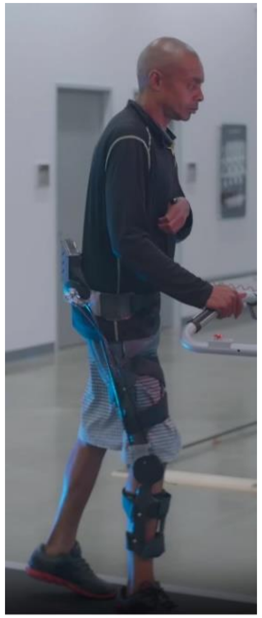
Reinforcement Learning and Control of Wearable Robots and Legged Robots







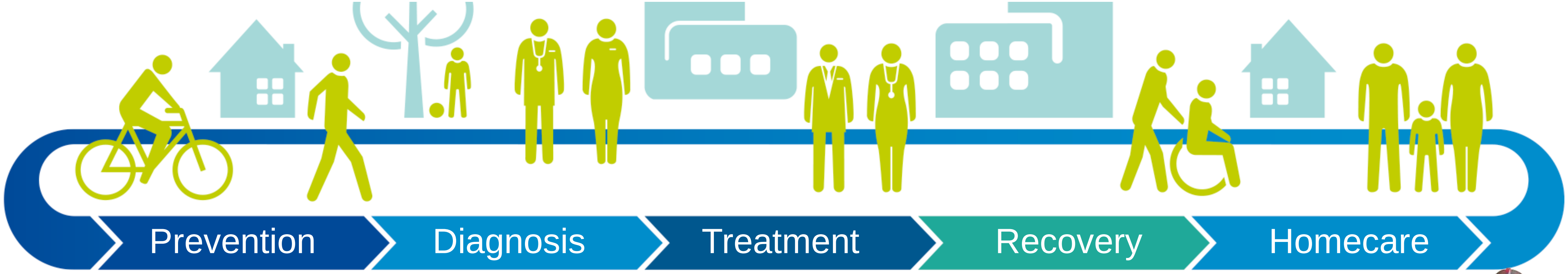

NC STATE UNIVERSITY

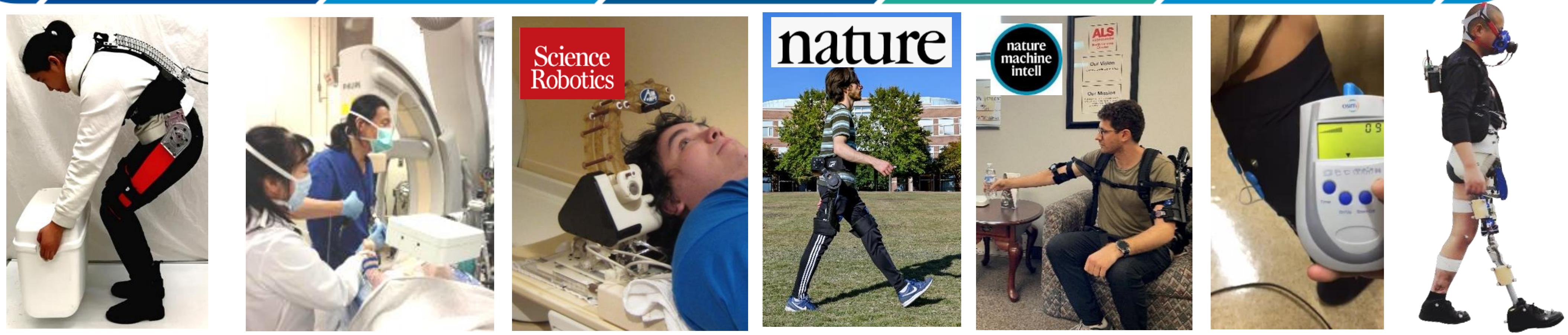
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Hao Su, PhD
Associate Professor
Director, Lab of Biomechatronics and Intelligent Robotics
https://haosu-robotics.github.io
Mechanical and Aerospace Engineering
North Carolina State University

BIRO
BIOMECHATRONICS AND INTELLIGENT ROBOTICS

Our AI-Powered Robots for Human Augmentation in Continuum of Care





Back exosuit	Ultrasound-guided steerable catheter	MRI-guided neurosurgery	Mobility exosuit	Shoulder exosuits	Transcutaneous stimulator	Robotic prosthesis

- [1] S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Silva, E. Rouse, B. Zhou, H. Yuk, X. Zhou, **H. Su**, “Experiment-free versatile optimization of exoskeleton assistance via learning-in-simulation”, Nature (minor revision)
- [2] **H. Su,** A. Di Lallo, R.R Murphy, R.H. Taylor, B.T. Garibaldi, A. Krieger, "Physical human-robot interaction for clinical care in infectious environments“, Nature Machine Intelligence, 2021
- [3] A. Gao, R. R. Murphy, W. Chen, G. Dagnino, P. Fischer, M. G. Gutierrez,..., **H. Su**, C. Wang & Yang, G. Z., “Progress in robotics for combating infectious diseases”, Science Robotics, 2021
- [4] **H. Su**, K. Kwok, ... & G.S. Fischer, "State of the Art and Future Opportunities in MRI-Guided Robot-Assisted Surgery and Interventions," Proceedings of the IEEE, 2022
- [5] X. Li, J. Zhu, M. Jiang, L. Yu, and **H. Su**, “A Hip Disarticulation Prosthesis Restores the Natural Gait and Improves Walking Efficiency”, Nature Biomedical Engineering (in prep)

Use-Inspired Basic Research for Translational Medicine

- Center of Assistive and Personal Robotics for Independent Living (APRIL)

Mechanical
Antonio Di Lallo
Postdoc
Ph.D., U of Pisa

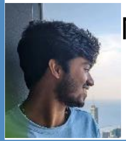
Weibo Gao
PhD student
M.S., UPitt

Yuming Yan
Postdoc
NTU

Jason Huang
PhD student
M.S., NCSU

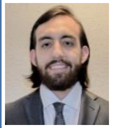
Electrical
Ivan Sanchez
Postdoc
Ph.D., CITEDI IPN

Junxi Zhu
PhD student
M.S., UMD



Nitin Srinivasan
MS student
B.S., NCSU

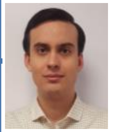
Control
Selina Zhang
PhD student
M.S., UESTC



Israel Dominguez
PhD student
M.S., CITEDI IPN

Menghan Jiang
PhD student
M.S., UPitt

Learning
Susan Luo
Postdoc
Ph.D., NJIT



Daniel Jorge
Postdoc
Ph.D., U of Seville

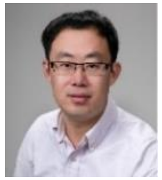
William Chen
PhD student
B.S., WFU

Collaborators

NIH | U.S. Department of Veterans Affairs | Penn University of Pennsylvania | Johns Hopkins University | Columbia University | UTHealth The University of Texas Health Science Center at Houston

Diane Damiano
Tom Bulea
Chief/Scientist
NIH Clinical Center

Ann Spungen
Professor
Bronx Veteran Affairs



Hao Su
Lab Director
NCSU

Michelle Johnson
Associate Professor
UPenn

Axel Krieger
Assistant Professor
JHU

Joel Stein
Professor, Chair
Columbia

- $10.2 M grants, NIH R01EB029765, R01EB035404, Toyota Mobility Foundation, ALS Association
- NSF: CAREER, National Robotics Initiative, Cyber-Physical Systems (CPS), Future of Work
- National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR), VA

NSF | NIH | NIDILRR National Institute on Disability, Independent Living, and Rehabilitation Research | U.S. Department of Veterans Affairs | ALS Association | Toyota mobility Foundation

Program Objectives

- Understand how to conduct research in US
  - Critical thinking, first principle mindset
- Develop skills in literature review and **innovation**
- Writing and communication with professors
  - MIT How to speak, Patrick Winston: https://www.youtube.com/watch?v=Unzc731iCUY
- Reinforcement learning (RL) for control of robots
- Ubuntu, PyTorch, IsaacGym (Nvidia), MuJoCo
- RL for legged robot control

Tentative Syllabus

Week	Lecture	Content (in New York time)	Homework (due in 1 week)
1	1	Introduction to robot control with reinforcement learning (07/11, 10a - 11:30a, Room 3209, Enginnering Building III)	See Paper List
1	2	Project Overview (humanoid robot) Introduction to Ubuntu, MuJoCo, IsaacGym & Custom URDF files and visualize in MuJoCo (07/12, 10a - 11:30a, Room 3209, Enginnering Building III)	
2	3	Introduction to Reinforcement learning Cart-pole example (Date TBD)	See Homework 1 Compute Network Parameters and Compare Controller Performance
2	4	Introduction to humanoid robots (humanoid robot project) Reward function formulation (Date TBD)	See Homework 2 Reward Function Formulation
3	5	Tune reward function and train humanoid robot controller (Date TBD)	See Homework 3 Humanoid Virtual Competition
3	6	Poster feedback	Poster must be completed by July 25

GitHub: https://github.com/biomechatronics001/NCSU-GEARS-Reinforcement-Learning-Humanoid-Control

Course Logistics

- **Lecture: 10:00-11:30AM EST (Week 1)**
- **Office hour: by appointment**
- Zoom meeting link
  - https://ncsu.zoom.us/my/junxizhu
- TA
  - Junxi Zhu (jzhu35@ncsu.edu)
  - Ivan Lopez-Sanchez (ilsanche@ncsu.edu)

Prof. Hao Su WeChat

Hands-on Project

- Students will train a bipedal robot to do unique activity in simulation using **actor-critic** deep reinforcement learning techniques.
- Two students will be in one team
  - **But we need a poster from each individual**
- Please cc Dr. Su (hsu4@ncsu.edu) when you email any mentors

References for Exoskeletons

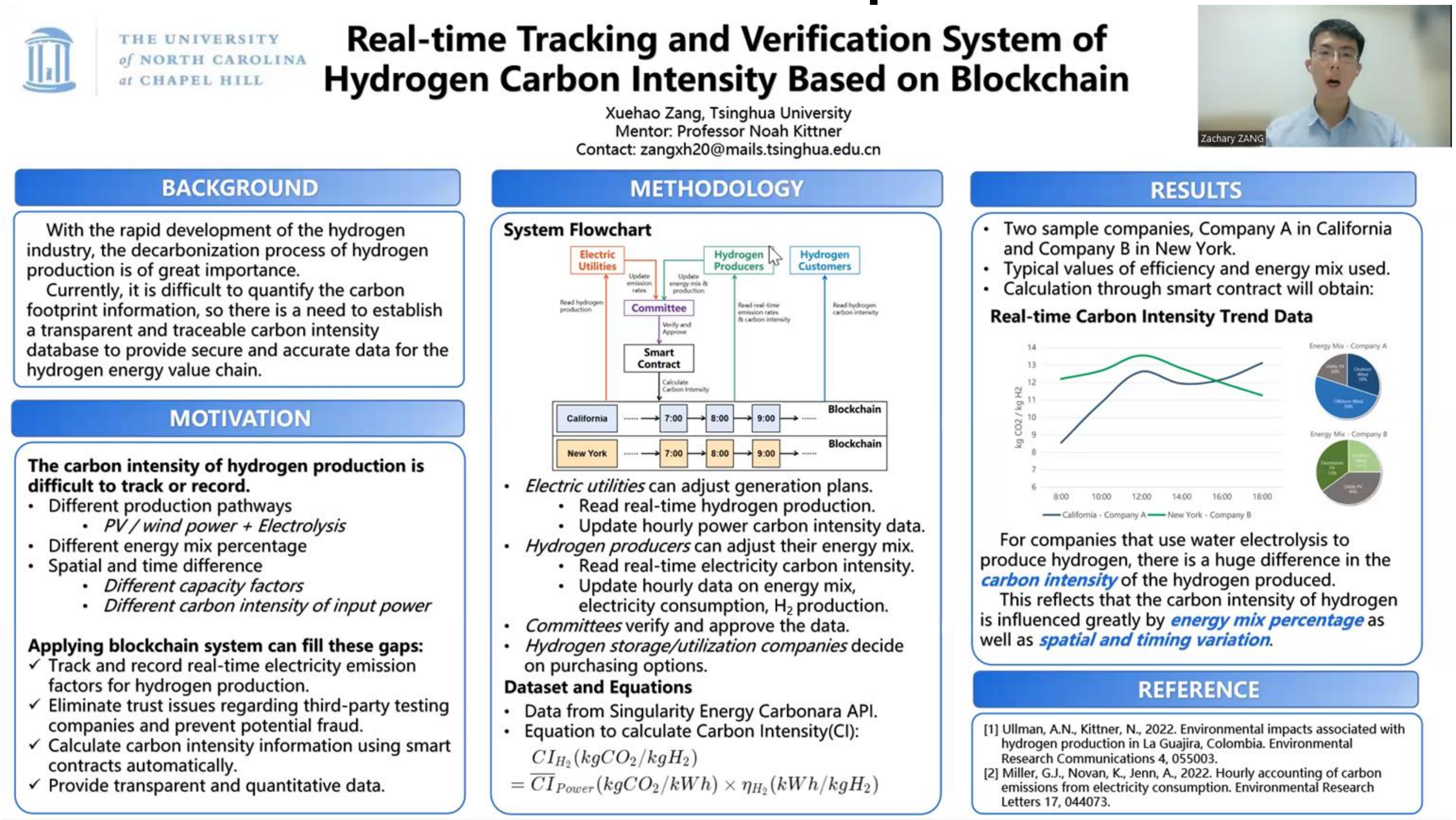
- [Delp22] Uhlrich SD, Jackson RW, Seth A, Kolesar JA, Delp SL. Muscle coordination retraining inspired by musculoskeletal simulations reduces knee contact force. Scientific Reports. 2022 Jul 7;12(1):1-3.
- [Collins21] Poggensee KL, Collins SH. How adaptation, training, and customization contribute to benefits from exoskeleton assistance. Science Robotics. 2021 Sep 29;6(58):eabf1078.
- [Collins20] Witte KA, Fiers P, Sheets-Singer AL, Collins SH. Improving the energy economy of human running with powered and unpowered ankle exoskeleton assistance. Science Robotics. 2020 Mar 25;5(40):eaay9108.
- [Huang22] Huang TH, Zhang S, Yu S, Su H. Modeling and Stiffness-Based Continuous Torque Control of Lightweight Quasi-Direct-Drive Knee Exoskeletons for Versatile Walking Assistance. IEEE Transactions on Robotics. 2022

- Conor Walsh: Harvard
  - https://www.youtube.com/watch?v=tCBoiepuLto
- Steve Collins: Stanford
  - https://www.youtube.com/watch?v=gZJMMXAf1Pc

References for Reinforcement Learning

- **[Sutton92]** Sutton RS, Barto AG, Williams RJ. Reinforcement learning is direct adaptive optimal control. IEEE control systems magazine. 1992 Apr;12(2):19-22.
- **[Hutter19 Science Robotics]** Hwangbo J, Lee J, Dosovitskiy A, Bellicoso D, Tsounis V, Koltun V, Hutter M. Learning agile and dynamic motor skills for legged robots. Science Robotics. 2019 Jan 16;4(26):eaau5872.
  - https://www.youtube.com/watch?v=6WLqK2X1RAc
- **[Hutter20 Science Robotics]** Lee J, Hwangbo J, Wellhausen L, Koltun V, Hutter M. Learning quadrupedal locomotion over challenging terrain. Science robotics (SR). 2020 Oct 21;5(47):eabc5986.
- Keynote speech **Marco Hutter ICRA 2022**
  - https://www.youtube.com/watch?v=abdLIFOzdRo
- **[Luo21]** S. Luo, G. Androwis, S. Adamovich, H. Su, E. Nunez and X. Zhou (2021) Reinforcement Learning and Control of a Lower Extremity Exoskeleton for Squat Assistance. Front. Robot. AI 8:702845.
- **[Luo22]** Luo S, Androwis G, Adamovich S, Nunez E, Su H, Zhou X. Robust Walking Control of a Lower Limb Rehabilitation Exoskeleton Coupled with a Musculoskeletal Model via Deep Reinforcement Learning, 2022
- **[Jeon23 ICRA]** S. H. Jeon, S. Heim, C. Khazoom, and S. Kim, "Benchmarking Potential Based Rewards for Learning Humanoid Locomotion," in 2023 IEEE International Conference on Robotics and Automation (ICRA), London, United Kingdom, May 2023, pp. 9204–9210.
- **[Davide23 Nature]** E. Kaufmann, L. Bauersfeld, A. Loquercio, M. Müller, V. Koltun, and D. Scaramuzza, "Champion-level drone racing using deep reinforcement learning," Nature, vol. 620, no. 7976, pp. 982–987, Aug. 2023.
- **[Our 2024 Nature]** S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, X. Zhou, and H. Su, "Experiment-free exoskeleton assistance via learning in simulation," Nature, vol. 630, no. 8016, pp. 353–359, Jun. 2024.

Poster Example



Real-time Tracking and Verification System of Hydrogen Carbon Intensity Based on Blockchain

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Xuehao Zang, Tsinghua University
Mentor: Professor Noah Kittner
Contact: zangxh20@mails.tsinghua.edu.cn

Zachary ZANG

BACKGROUND

With the rapid development of the hydrogen industry, the decarbonization process of hydrogen production is of great importance.

Currently, it is difficult to quantify the carbon footprint information, so there is a need to establish a transparent and traceable carbon intensity database to provide secure and accurate data for the hydrogen energy value chain.

MOTIVATION

The carbon intensity of hydrogen production is difficult to track or record.

- Different production pathways
  - *PV / wind power + Electrolysis*
- Different energy mix percentage
- Spatial and time difference
  - *Different capacity factors*
  - *Different carbon intensity of input power*

Applying blockchain system can fill these gaps:

✓ Track and record real-time electricity emission factors for hydrogen production.
✓ Eliminate trust issues regarding third-party testing companies and prevent potential fraud.
✓ Calculate carbon intensity information using smart contracts automatically.
✓ Provide transparent and quantitative data.

METHODOLOGY

System Flowchart

- *Electric utilities* can adjust generation plans.
  - Read real-time hydrogen production.
  - Update hourly power carbon intensity data.
- *Hydrogen producers* can adjust their energy mix.
  - Read real-time electricity carbon intensity.
  - Update hourly data on energy mix, electricity consumption, H_2 production.
- *Committees* verify and approve the data.
- *Hydrogen storage/utilization companies* decide on purchasing options.

Dataset and Equations

- Data from Singularity Energy Carbonara API.
- Equation to calculate Carbon Intensity(CI):

$$CI_{H_2}(kgCO_2/kgH_2)$$
$$= \overline{CI}_{Power}(kgCO_2/kWh) \times \eta_{H_2}(kWh/kgH_2)$$

RESULTS

- Two sample companies, Company A in California and Company B in New York.
- Typical values of efficiency and energy mix used.
- Calculation through smart contract will obtain:

Real-time Carbon Intensity Trend Data

For companies that use water electrolysis to produce hydrogen, there is a huge difference in the ***carbon intensity*** of the hydrogen produced.

This reflects that the carbon intensity of hydrogen is influenced greatly by ***energy mix percentage*** as well as ***spatial and timing variation***.

REFERENCE

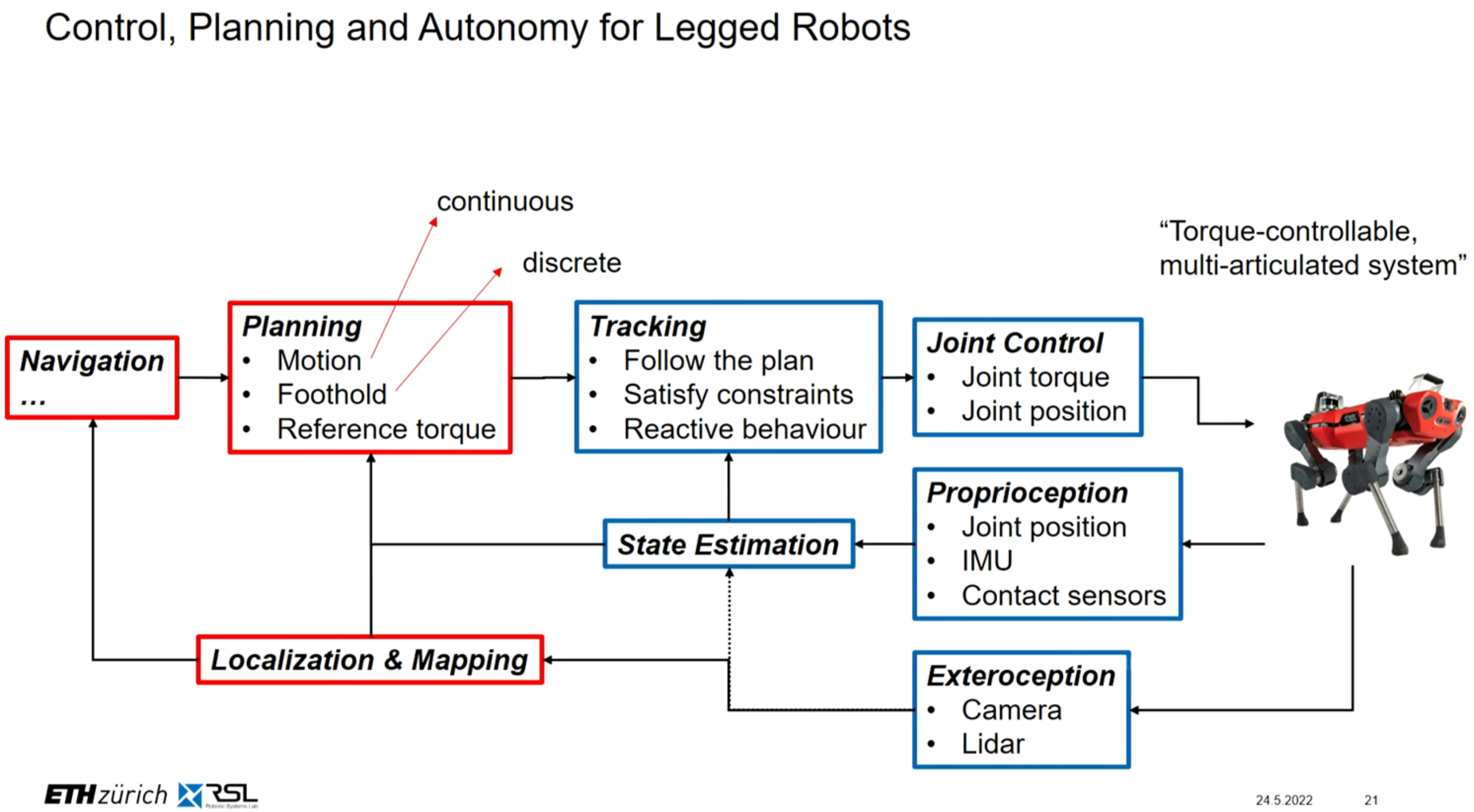
[1] Ullman, A.N., Kittner, N., 2022. Environmental impacts associated with hydrogen production in La Guajira, Colombia. Environmental Research Communications 4, 055003.

[2] Miller, G.J., Novan, K., Jenn, A., 2022. Hourly accounting of carbon emissions from electricity consumption. Environmental Research Letters 17, 044073.

https://gti.ncsu.edu/2022/14/31/gears-virtual-poster-presentation-winners-summer-2022/

Marco Hutter ICRA 2022

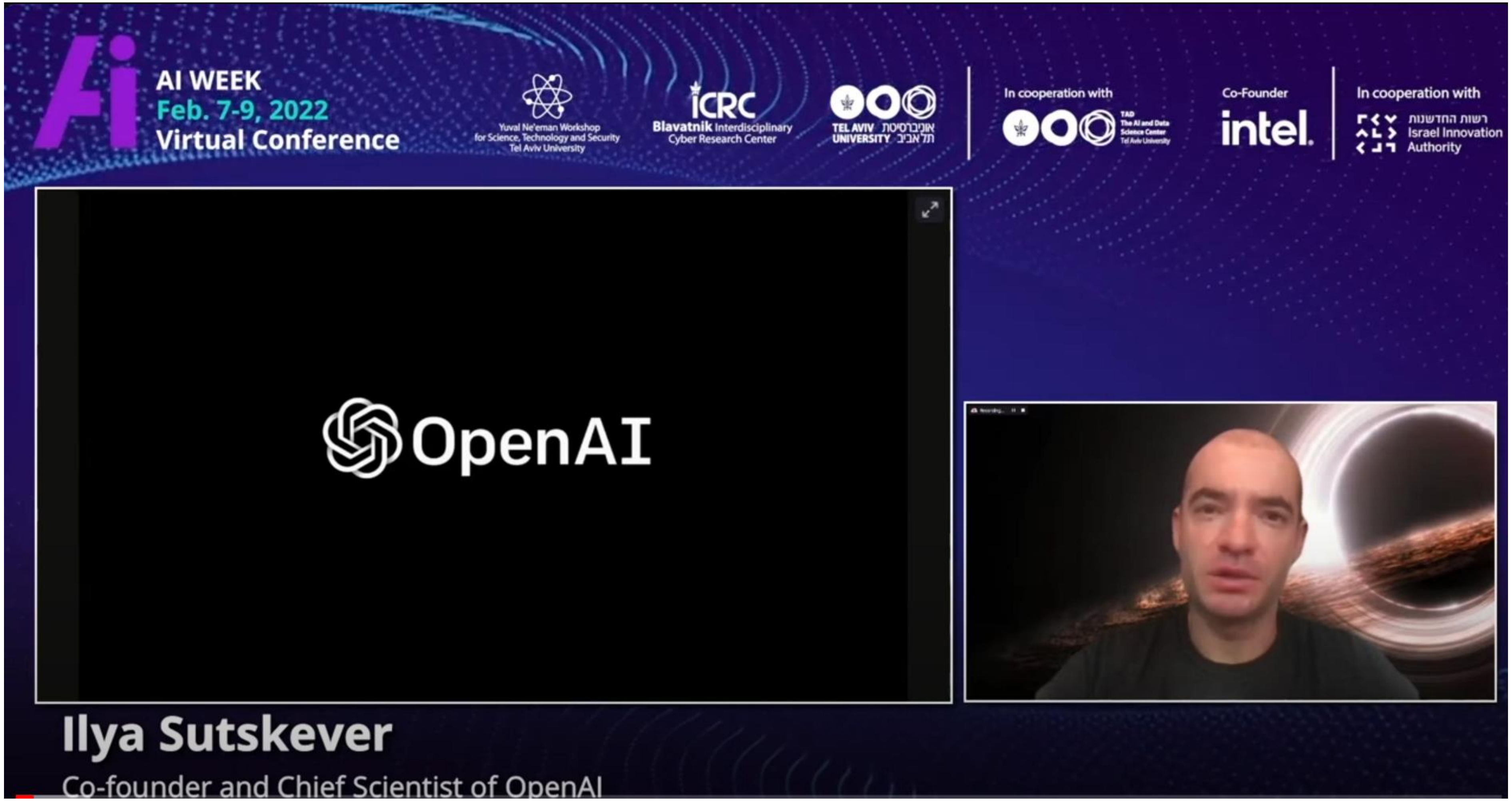
- https://www.youtube.com/watch?v=abdLIFOzdRo



Ilya Sutskever on Deep Learning

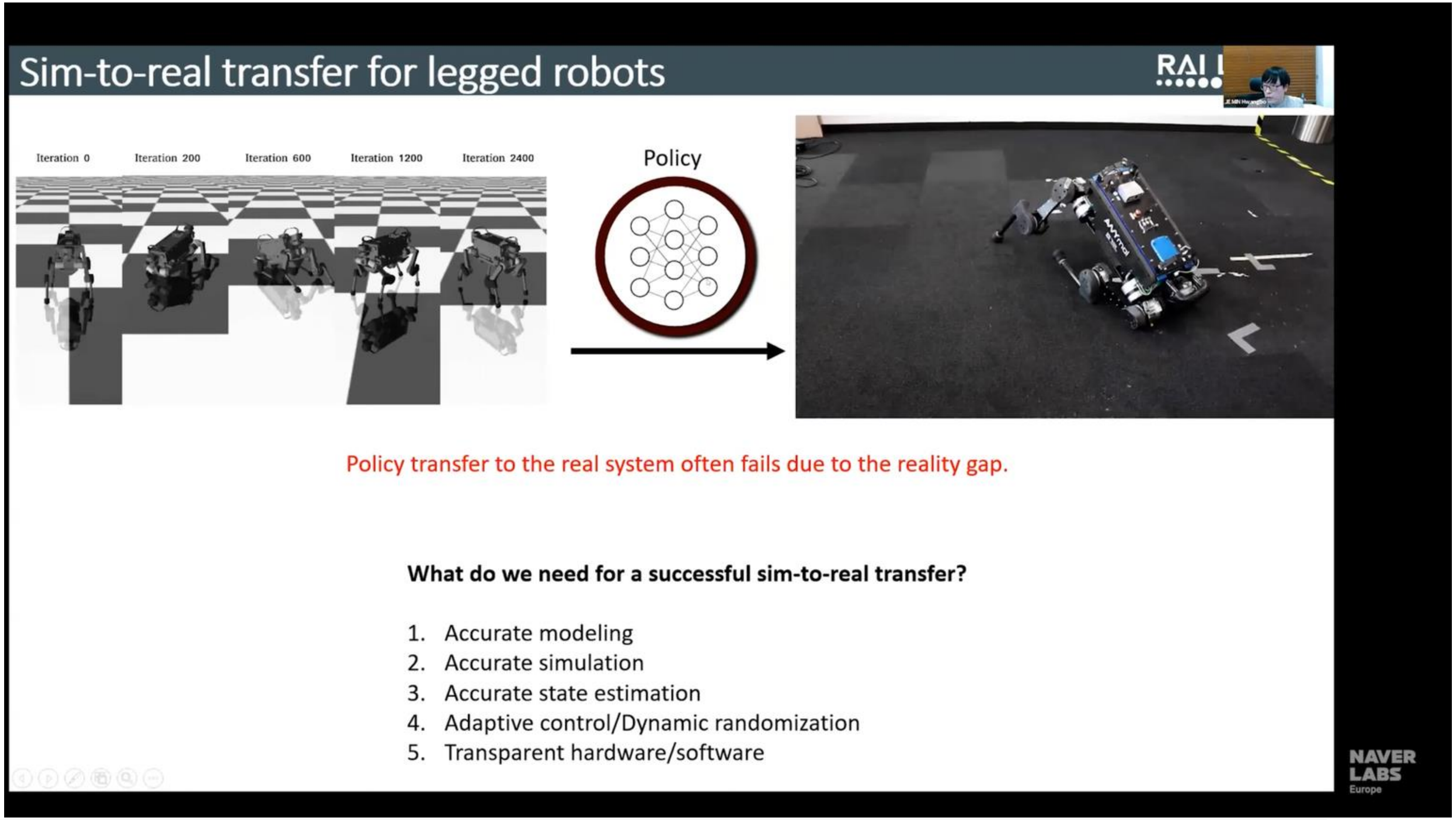
AI Week 2022

- https://www.youtube.com/watch?v=OAm6zyR_c8k



Control of Legged Robots using Reinforcement Learning, Jemin Hwangbo, KAIST, PhD at ETH

- https://www.youtube.com/watch?v=6WLqK2X1RAc



Our Experiment-free Learning-in-simulation Method

- Recently published in Nature: S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, X. Zhou, and H. Su, “Experiment-free exoskeleton assistance via learning in simulation,” Nature, vol. 630, no. 8016, pp. 353–359, Jun. 2024.
- Reduced 24.3%, 13.1% and 15.4% metabolic rate for walking, running and stair climbing

nature > articles > article

Article | Published: 12 June 2024

Experiment-free exoskeleton assistance via learning in simulation

Shuzhen Luo, Menghan Jiang, Sainan Zhang, Junxi Zhu, Shuangyue Yu, Israel Dominguez Silva, Tian Wang, Elliott Rouse, Bolei Zhou, Hyunwoo Yuk, Xianlian Zhou & Hao Su

Nature **630**, 353–359 (2024) | Cite this article

Barriers for Autonomous Control of Assistive Robots

- The real problem: people are extremely hard to predict

Can robot understand human intent for **autonomous** control of **multiple locomotion activities**?

- Intensive human testing (1h)

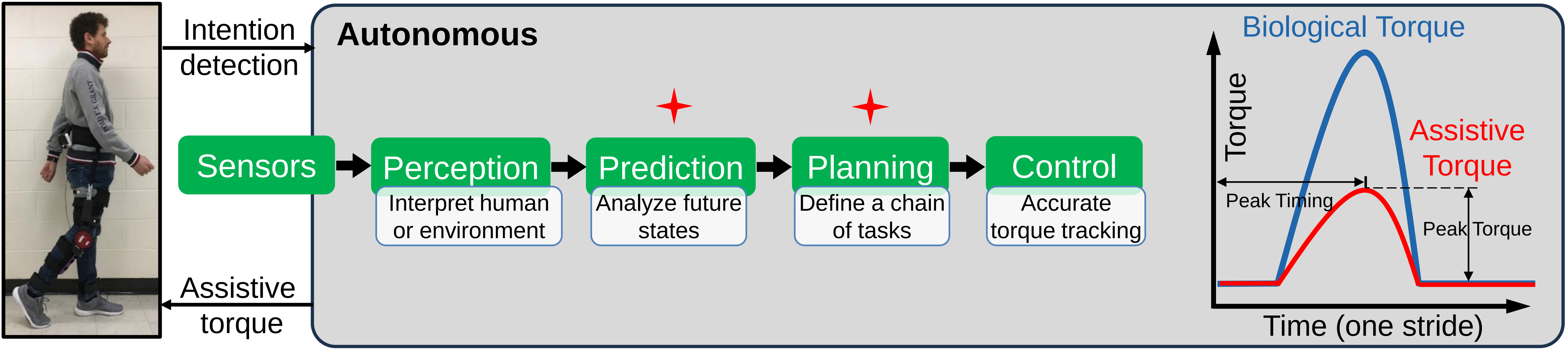
J. Zhang, P. Fiers, K.A. Witte, R.W. Jackson, K.L. Poggensee, C.G. Atkeson, S.H. Collins. Human-in-the-loop optimization of exoskeleton assistance during walking, Science, 2017

- Manual control to detect multimodal locomotion intention

L.J. Hargrove, A.J. Young, A.M. Simon, N.P. Fey, R.D. Lipschutz, S.B. Finucane, E.G. Halsne, K.A. Ingraham and T.A. Kuiken, Intuitive control of a powered prosthetic leg during ambulation: a randomized clinical trial. JAMA, 2015

Autonomous Control of Assistive Robots

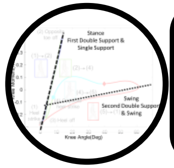
- State of the Art
  - performing a series of prescribed actions (predefined torque profile) based on internal states or user inputs
- Our goal: **autonomous**
  - **decision making** based on perception of human, environment, and internal states



S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, X. Zhou, and H. Su, “Experiment-free exoskeleton assistance via learning in simulation,” Nature, vol. 630, no. 8016, pp. 353–359, Jun. 2024.

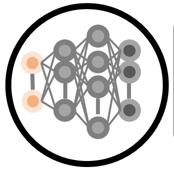
Autonomous Decision Making to Assist Multimodal Locomotion

- Leverage reinforcement learning to harness power of digital twin to eliminate human tests



Dynamics models + muscle models for explainable AI

+

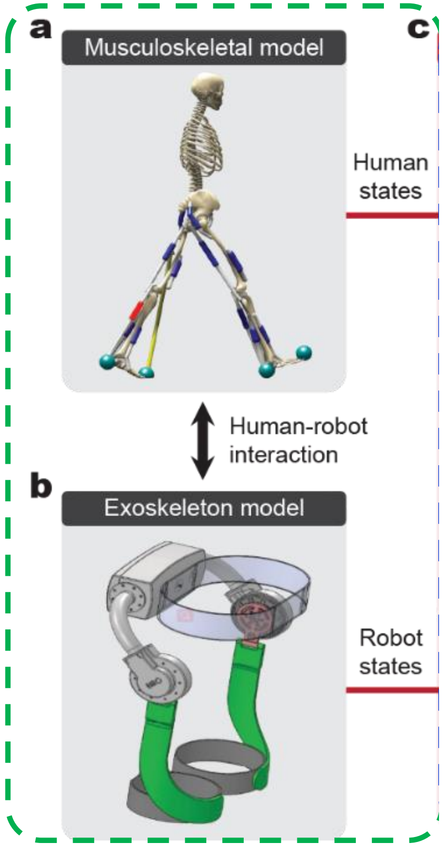


Large datasets to assist diverse locomotion behaviors

ChatGPT

Reinforcement Learning
(non-physical system)

Model-based Approach

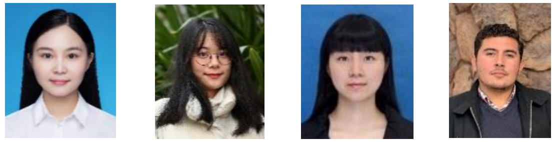


S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, X. Zhou, and H. Su, “Experiment-free exoskeleton assistance via learning in simulation,” Nature, vol. 630, no. 8016, pp. 353–359, Jun. 2024.

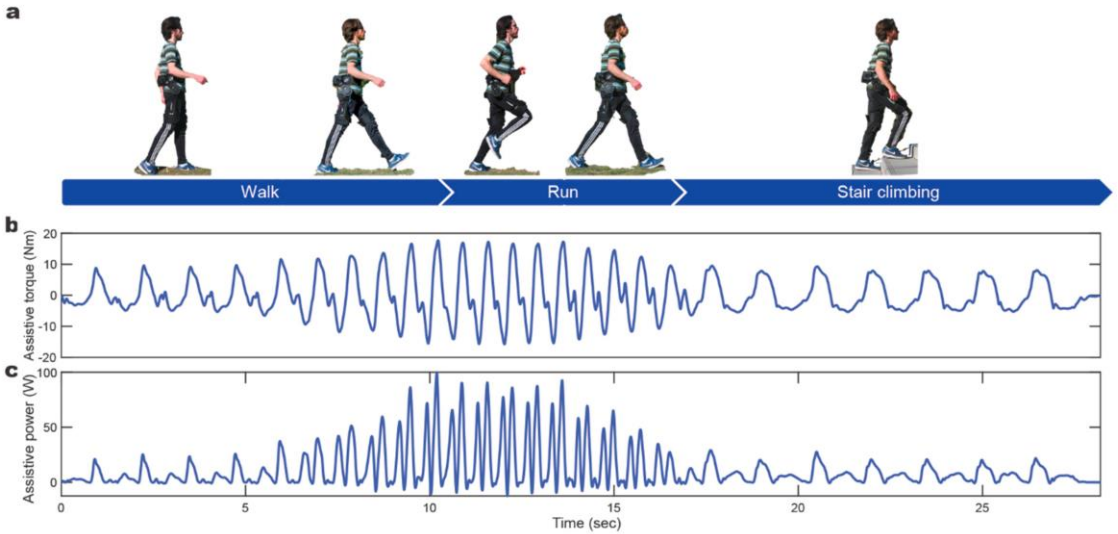
Continuous Assistive Torque for Versatile Activities

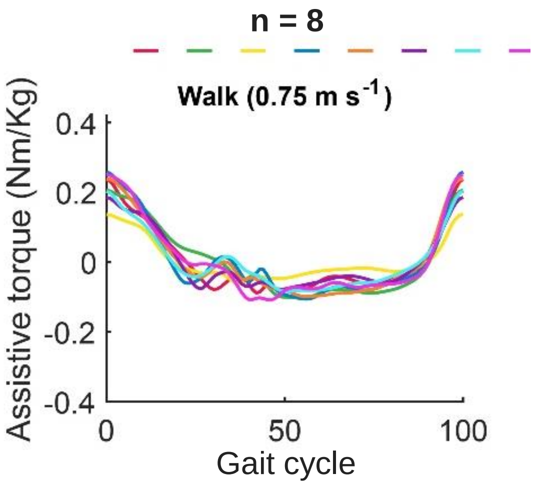
- Efficient (end-to-end): trained controller **does not rely on gait detection** or require any control parameter tuning
- Versatile: robot learn **individualized** controller to assist walking, running, stair-climbing

Sensors → Perception (Interpret human or environment) → Prediction (Analyze future states) → Planning (Define a chain of tasks) → Control (Accurate torque tracking)

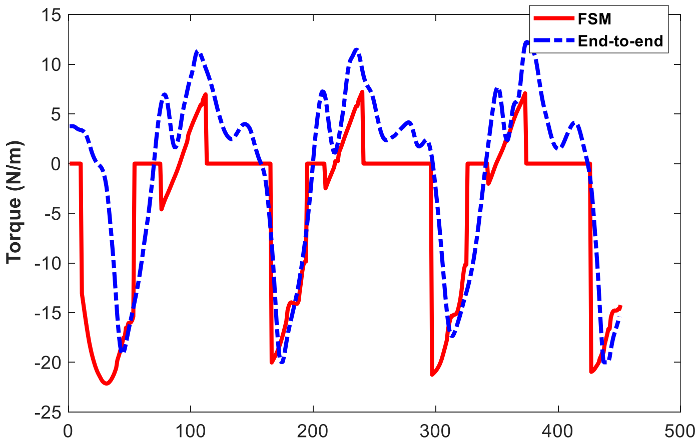


Susan Luo | Megan JIang | Sainan Zhang | Ivan Sanchez





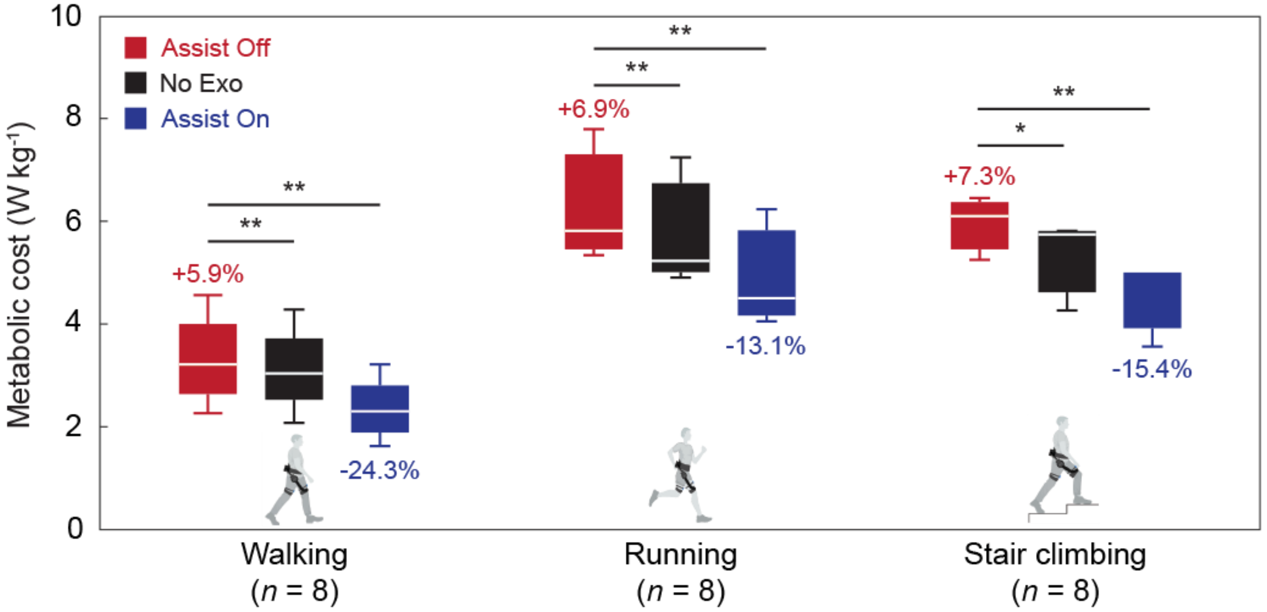
- Finite-State Machine (FSM): jerky
- Our method: continuous and synergistic



S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, X. Zhou, and H. Su, “Experiment-free exoskeleton assistance via learning in simulation,” Nature, vol. 630, no. 8016, pp. 353–359, Jun. 2024.

Our Controller: Significant Saving of Human Energetics

- Controller trained in the simulation works in the real world!
- Large reduction of metabolic cost : walk (24.3%), run (13.1%), stair climb (15.4%)
  - Equivalent to reducing body mass by 25kg, 14kg, 16kg (athletic performance)



S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, X. Zhou, and H. Su, "Experiment-free exoskeleton assistance via learning in simulation," Nature, vol. 630, no. 8016, pp. 353–359, Jun. 2024.

Computational Model for Digital Evaluation of Assistive Robots

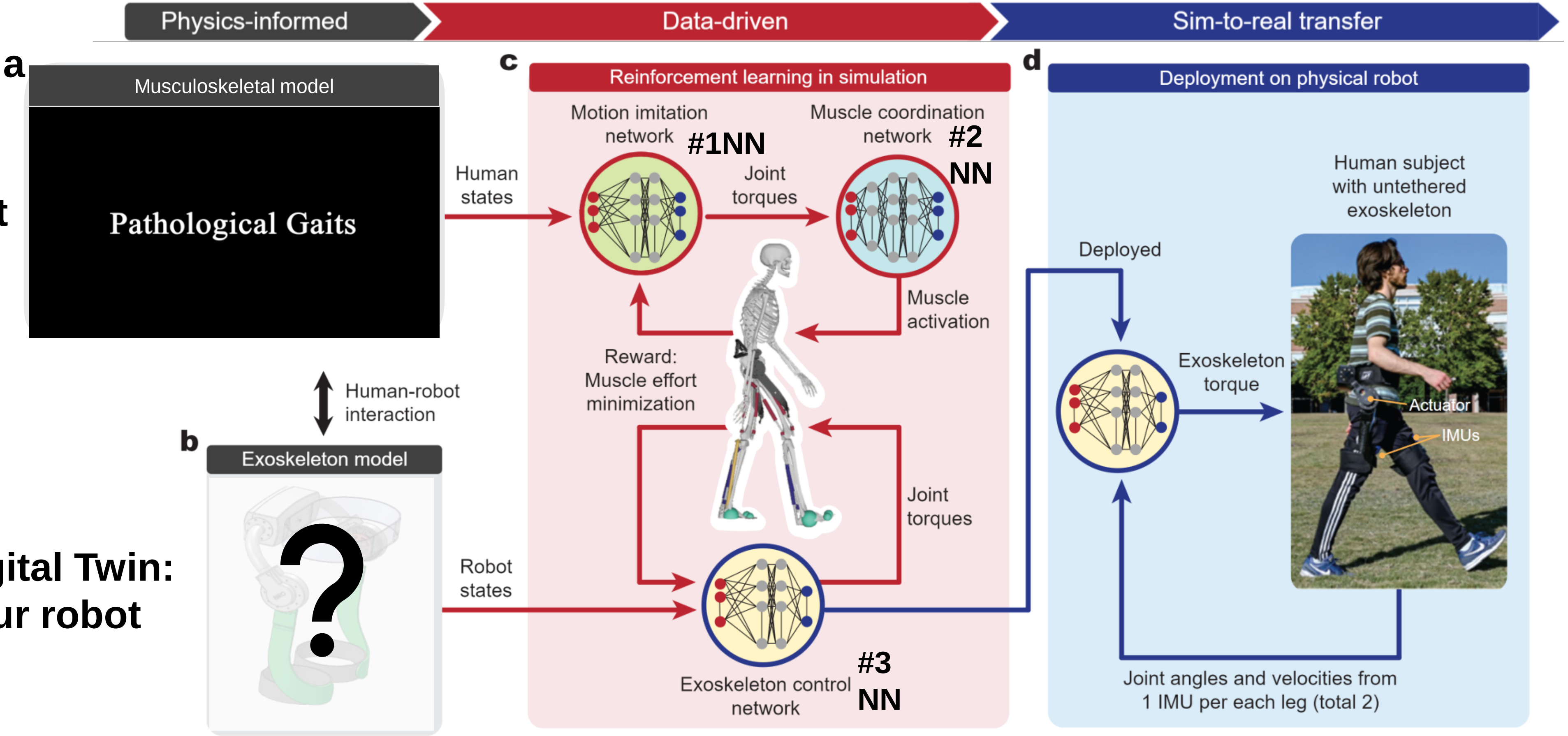
- Digital twin of human and assistive devices
- Foot drop and vaulting gait of stroke, couch gait of cerebral palsy
- Is it possible to perform digital clinical trial of assistive devices?


Ivan Sanchez


Daniel Jorge

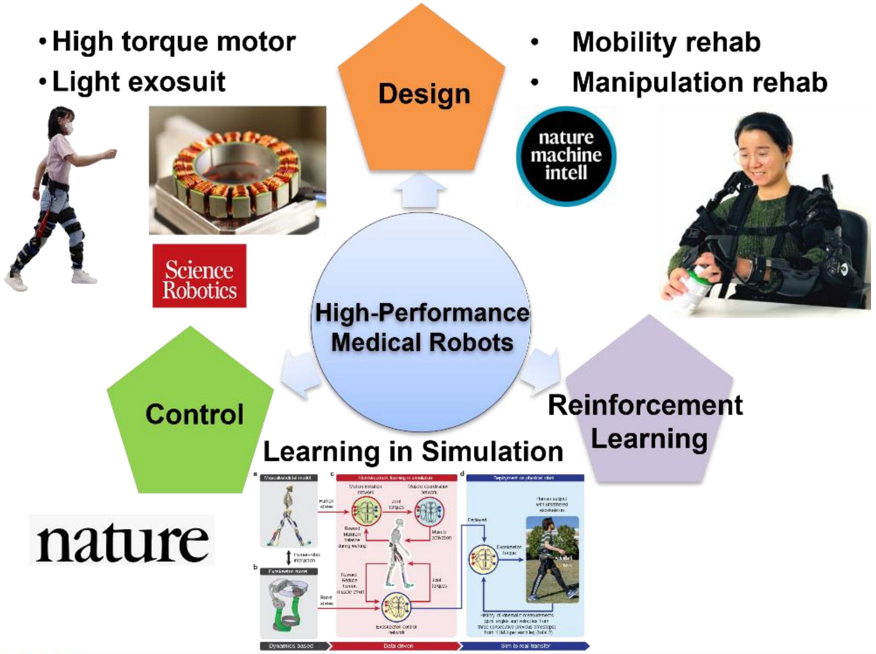
Digital Twin: Pathological Gait

Digital Twin: Your robot



Summary: AI-Powered Soft Robots for Human Augmentation and Rehabilitation

- Contribution to **Engineering**
  - Bio-inspired design of electromagnetic actuator creates comfort-centered soft exosuits
- Contribution to **Science**
  - Physics-informed and data-driven reinforcement learning provides autonomous control to augment human performance (13%-24% saving of human energy)
- Contribution to **Medicine**
  - Promising solutions for pain-relief of knee osteoarthritis and gait restoration of neurorehabilitation



Neurorehabilitation

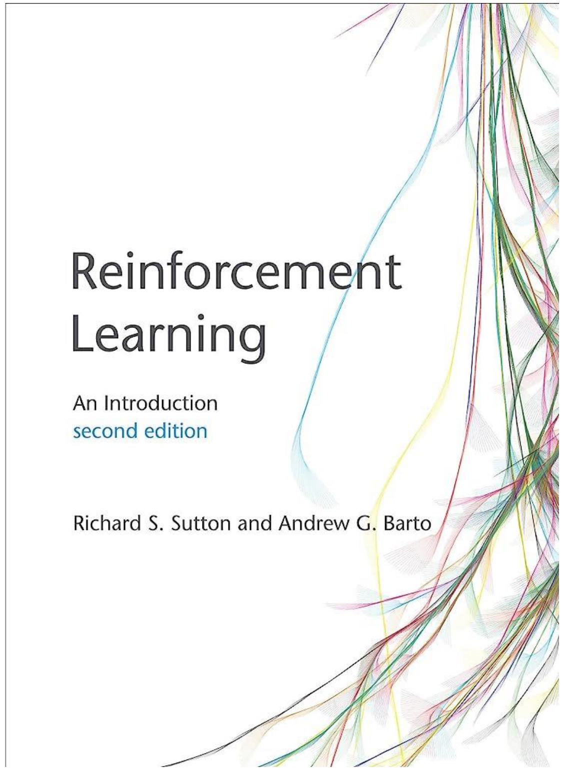
Surgery

Orthopedics

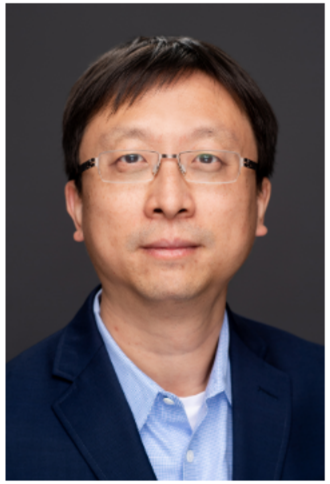
Regenerative medicine

Richard Sutton

- Reinforcement Learning, second edition: An Introduction (Adaptive Computation and Machine Learning series)



Yi Ma

- Recorded video of **An Overview of Reinforcement Learning and Optimal Control** (with **Slides**), February 17, 2021.
- 

Professor Yi Ma

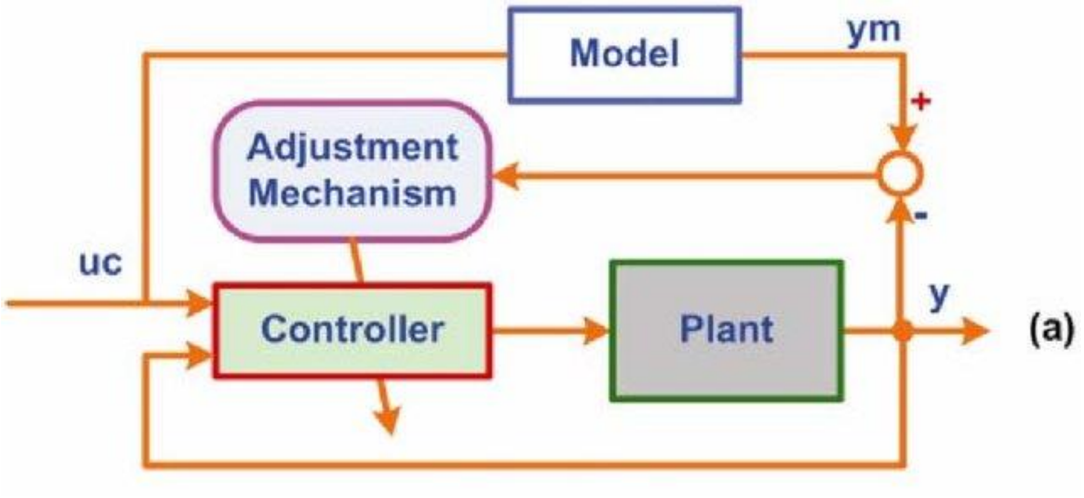
Electrical Engineering and Computer Sciences
University of California, Berkeley

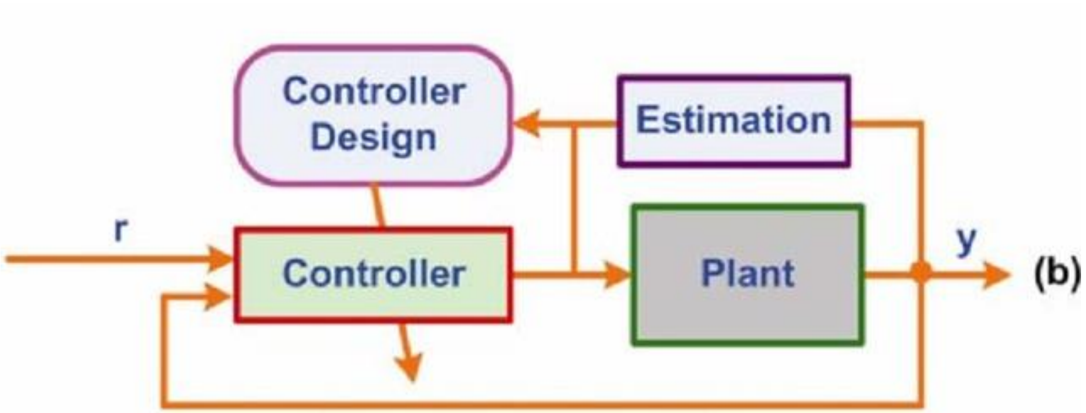
Office: *EECS Department, 333C Cory Hall # 1770*
University of California, Berkeley, CA 94720-1770, USA.
Email: *yima@eecs.berkeley.edu*
Phone: *+1 (510) 664 4565*

Other Affiliations
Berkeley Artificial Intelligence Research (BAIR)
Berkeley Center for Augmented Cognition (CRC)
Berkeley FHL Vive Center for Enhanced Reality
Tsinghua-Berkeley Shenzhen Institute
Institute of Data Science, Hong Kong University

Two Types of Adaptive Control

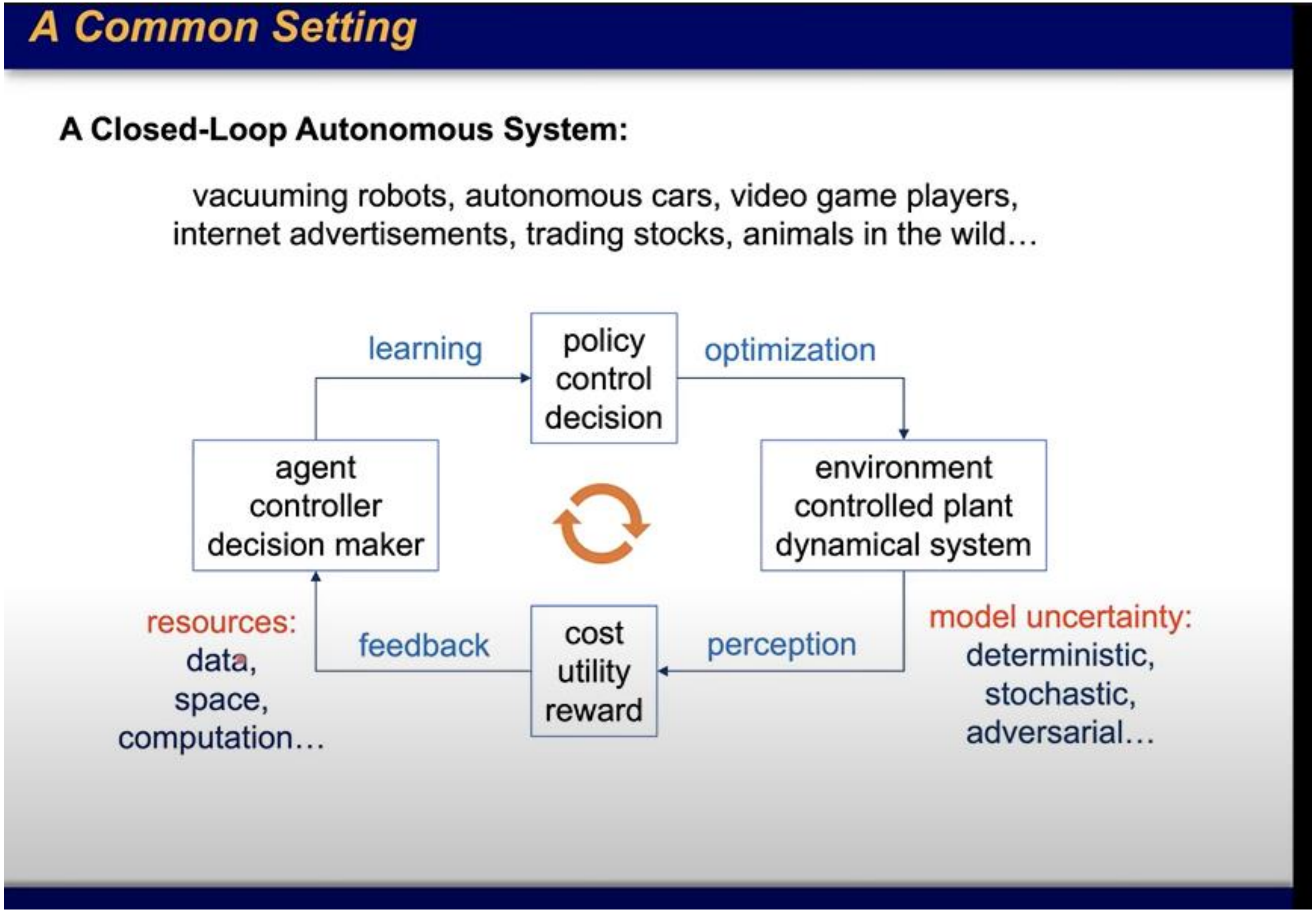
- Direct adaptive control scheme
  - Direct adaptive **optimal** control: RL
- Indirect adaptive control scheme

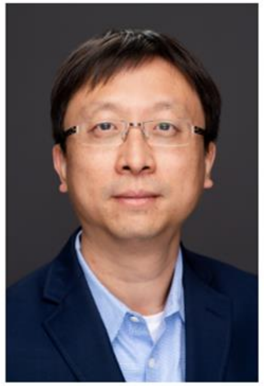




[Sutton92] Sutton RS, Barto AG, Williams RJ. Reinforcement learning is direct adaptive optimal control. IEEE control systems magazine. 1992 Apr;12(2):19-22.

Overview of Reinforcement Learning and Optimal Control





Overview of Reinforcement Learning and Optimal Control

From Principle to Computation!

What to Compute, and How?

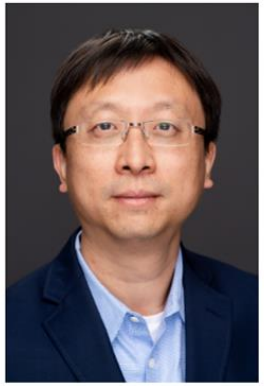
	OC/DP	AI/RL
Optimal value function:	$J^*(x)$,	$V^\star(s)$
Optimal Q-function:	$Q^*(x,u)$,	$Q^\star(s,a)$
Optimal control/policy:	$u^*(x)$,	$\pi^\star(a \mid s)$ (or $u^\star(y)$, $\pi^\star(a \mid o)$)
System/model identification:	$f^*(x,u)$,	$p^\star(s_{t+1} \mid s_t, a_t)$

Closed-form versus numerical solution (simulation & optimization)

LQR: $J^*(x_k) = \min_{u_k} \left[x_k^T Q x_k + u_k^T R u_k + J^*(A x_{k+1} + B u_k) \right]$

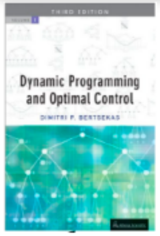
The Riccati equation (Kalman Filter '60):

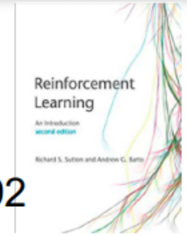
$$K_k = -(\bar{R} + \bar{B}^\intercal V_{k+1} \bar{B})^{-1} \bar{B}^\intercal V_{k+1} \bar{A} \tag{48}$$
$$V_k = \bar{Q} + \bar{A}^\intercal V_{k+1} \bar{A} - \bar{A}^\intercal V_{k+1} \bar{B} (\bar{R} + \bar{B}^\intercal V_{k+1} \bar{B})^{-1} \bar{B}^\intercal V_{k+1} \bar{A} \tag{49}$$



Overview of Reinforcement Learning and Optimal Control

Control versus Learning





OC/DP

- LQR
- Parallel parking
- Chained form systems
- Mechanical systems…

Conditions & Assumptions

- clear model class/uncertainty
- clear cost function
- low to moderate dimension
- continuous state/time…

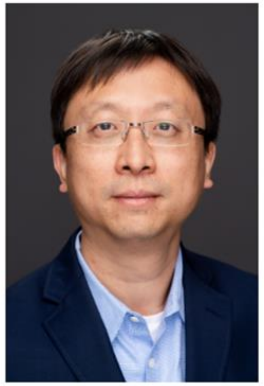
AI/RL

- Backgammon: Tesauro, 1992
- Chess: Deep Blue, 1997
- Go: Alpha Go, 2017
- Video games, robots…

Conditions & Assumptions

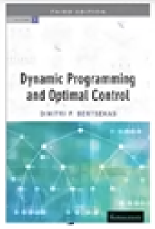
- unknown models (but can sample)
- uncertain, long-horizon return
- large-scale, high-dimensional
- discrete state/time…

Solutions that work for a broad class of problems v.s. a few (important) instances



Overview of Reinforcement Learning and Optimal Control

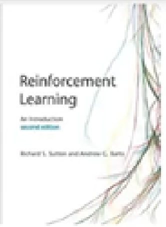
Terminology: State Space Model



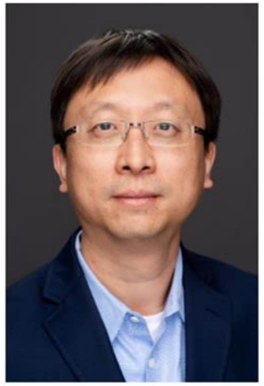
OC/DP

environment
controlled plant
dynamical system

AI/RL



OC/DP	AI/RL
State and Control space: $\mathcal{S}, \mathcal{U}$	State and Action space: $\mathcal{S}, \mathcal{A}$
State: $x_k \in \mathcal{S}, k = 0, 1, \ldots$	State: $s_t \in \mathcal{S}, t = 0, 1, \ldots$
Control: $u_k \in \mathcal{U}, k = 0, 1, \ldots$	Action: $a_t \in \mathcal{A}, t = 0, 1, \ldots$
Dynamical System:	MDP Transition (or simulation):
$x_{k+1} = f(x_k, u_k)$ $x_{k+1} = f(x_k, u_k, \underbrace{w_k}_{\text{stochastic}})$	$\mathcal{T}_{ijk} = p(s_{t+1} = i \mid s_t = j, a_t = k)$
Output/observation (feature):	Observation (feature):
$y_k = h(x_k, u_k) + \underbrace{n_k}_{\text{stochastic}}$	$p(o_t \mid s_t)$



Homework 1: Cart-pole Control with Reinforcement Learning

- This example uses OpenAI Gym
- Compare the performance of the controller when changing some features of its architecture
- Learn how to create graphical representations of a neural network architecture
- Develop a critical perspective on the features of a neural network architecture and how it impacts the performance

- The details of this activity will be available for you in the GitHub repository at the Cart-Pole Example section.

Tentative Program Deadline

- Homework are generally due in 1 week
- July 18: Homework 1 Due
- July 24: Homework 2 Due
- July 24: Homework 3 Due
- July 25: Poster Rough Draft Due
- July 26: Poster final version