Robots Teaching Humans: A New Communication Paradigm via Reverse Teleoperation

Blue Sky Ideas Track

Rika Antonova*
Stanford University

Ankur Handa
NVIDIA

ABSTRACT

Simulators offer a scalable platform to train robots, offering a path to creative and innovative solutions that are difficult for humans to envision a priori. We introduce a way to leverage this property, along with a new paradigm where robots discover creative solutions in simulation, then teach humans or other agents to physically execute the learned solutions via reverse teleoperation. We provide various examples ranging from learning new skills, to rehabilitation, to everyday activities, where such a system would be valuable.

KEYWORDS

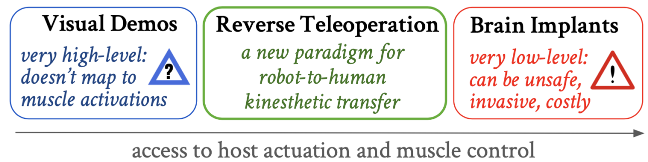
Teleoperation; Reinforcement Learning; Physical Simulation

ACM Reference Format:
Rika Antonova and Ankur Handa. 2022. Robots Teaching Humans: A New Communication Paradigm via Reverse Teleoperation: Blue Sky Ideas Track. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 5 pages.

1 MOTIVATION: COMMUNICATING NEW KINESTHETIC INSIGHTS

The 37th move of the 2nd game between the computer program AlphaGo [29] and the 18-time world champion Lee Sedol caught the human champion by surprise. The move was puzzling to most observers as well – it did not follow any known strategies. What seemed a mistake at first turned the course of the game. 'Move 37' was eventually recognized as AlphaGo's own innovation, since it could not be attributed to memorizing human strategies used during training. The successors AlphaZero [30] and MuZero [26] learned from self-play without human guidance. Go enthusiasts now recognize the potential of these programs to generate unique insights for new game strategies.

Robots – the embodied programs – could generate valuable innovative insights for solving physical tasks. They could safely attempt a myriad tries in simulation, and the resulting strategies would be applicable to the real world if the simulation-to-reality mismatch is not large. They could also learn from real-world interactions, and instantly share progress with other robots that have a similar embodiment. However, robots lack an effective way to communicate



their innovations to humans. It is unlikely that people pick up the new strategies by watching robot motions, just as the approach of learning to play the piano by watching is unlikely to succeed. What could be an effective way for the robots to teach their innovations to us? Our blue-sky idea is to develop a new paradigm that would allow robots to communicate to humans their kinesthetic insights directly via *reverse teleoperation.*

2 ROBOT-TO-HUMAN SKILL TRANSFER: POTENTIAL BENEFITS

With rapid advances in learning and simulation, running large-scale distributed optimization should soon allow robots to find novel ways of solving various tasks. They are likely to improve beyond human-level performance even when constrained by similar physical and hardware characteristics.

Instead of using this only for perfecting robotic performance, we propose to think about how this could help humans. For now, we will assume that we either train anthropomorphic agents in simulation, or train humanoid robots in reality and can solve the mapping to the human morphology. Below we summarize the benefits of transferring kinesthetic knowledge from robots to humans.

Increasing motivation and reducing time to mastery: Beginners that fail to make progress can become demotivated and disengaged. Reverse teleoperation could help execute a new skill correctly right away. A robot can then slowly attenuate the support to reduce human's reliance on hardware, balancing challenges vs success to maintain a high level of motivation even in unskilled beginners. Moreover, reverse teleoperation can ensure high quality for the full duration of the practice, reducing the amount of time wasted on rehearsing incorrectly, i.e. avoid forming 'bad habits'.

Improving ergonomics: Repetitive tasks are not going to disappear even if robots replace humans in factory work, because we need to do office work and household tasks that are difficult to fully automate. We can encode ergonomics guidelines into the reward signal directly when training simulated/robotic agents. Teaching humans to perform repetitive tasks in ergonomically optimal ways could ensure better well-being and prevent repetitive strain injuries. Even when people know about the correct approach they frequently lack the habit/discipline to follow it. Reverse teleoperation would be an easy way to help them acquire correct habits *e.g.* helping to hold the correct arm posture for typing, help keeping one's back straight during lifting, etc.

*Corresponding author: rika.antonova@stanford.edu
Rika is supported by the National Science Foundation grant No.2030859 to the Computing Research Association for the CIFellows Project.
** The authors would like to thank Sam Devlin (Microsoft Research, Cambridge, UK) for comments and suggestions, and also acknowledge the anonymous reviewers for their constructive feedback.

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

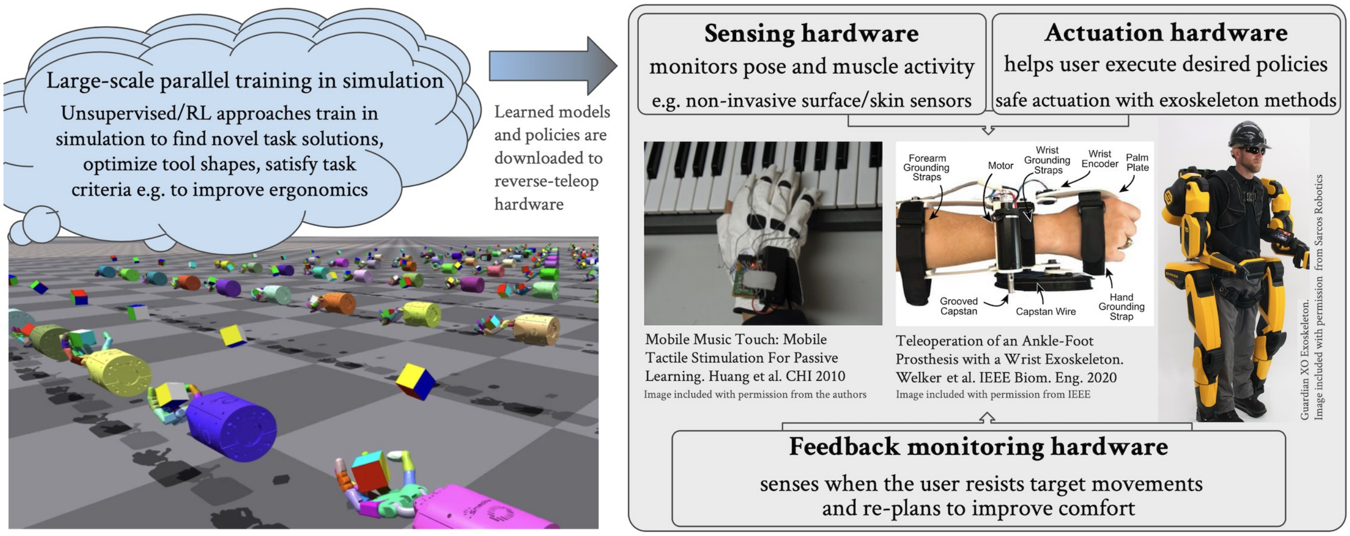


Figure 1: Overview of the proposed reverse teleoperation architecture: policies are learned in simulation and then run on hardware that activates a person's motions with feedback-driven control.

New ways of solving old tasks: High jump has been a popular sport since the 19th century, but the optimality of the 'Fosbury flop' to go over the bar backwards has not been evident to all, until Dick Fosbury won the 1968 Olympics [36]. The skills and tool use strategies we have for many of our everyday and professional tasks could be suboptimal. Daring exploration strategies during large-scale simulation training could let us overcome these local optima.

Designing new tools & teaching humans how to use them: Distributed large-scale training in simulation could help create new tools and optimize their shapes. Shape optimization using physics (aerodynamics) simulation has been employed in automotive and aerospace industries [18, 19], and more recently in robotics [4, 15]. We could search for new tools, optimize tool shapes, then quickly teach humans to use these tools. We can learn basic use of a novel tool first (e.g. with [34]), then attempt risky exploration strategies when optimizing in simulation or with a simplified robust robot hardware. This would ease innovation, even for risky tasks, without injuries (e.g. test new designs of a nail holder for hammering). This could also automate design of custom tools for people with disabilities and impairments, e.g. for the elderly. Overall, this could improve efficiency and safety of the tools.

Surgeries: Just as DaVinci robot [12] helped pioneer teleoperation for complex surgical operations, we could imagine future robots training surgeons to do delicate surgeries with known or new strategies. General surgical training takes at least five years after finishing medical school. Speeding up the process of correct muscle memory formation for the surgeons would yield quicker training. Making reverse teleoperation systems accessible to people throughout the world could be crucial to ensuring enough highly skilled surgeons, in both developed and developing countries.

Rehabilitation: Patients with various functional limb actuation or grasp pathologies can benefit from having a glove [23, 31] or full-body exoskeleton to enable learning to move and interact in the real world with various strategies that are tailored to their situation. Importantly, the robot could also help rehabilitation patients follow the correct progression of rehabilitation exercises, and facilitate faster recovery by providing a personalized schedules based on their current recovery progress.

Arts, Sports, and Fun: Lastly, we envision a wide range of further scenarios in our daily lives where reverse teleoperation can help to speed up muscle memory formation, e.g. playing piano [31], using chopsticks, touch typing, mastering novel painting techniques, forming correct habits for yoga poses and other physical exercises. Reverse teleoperation can also help mastering difficult fun skills, such as juggling and balancing for slacklining.

3 BACKGROUND AND PRIOR WORK

3.1 Teleoperation

The standard teleoperation paradigm in robotics is comprised of a hardware system and a user interface that allow a human to relay the desired motions to a robotic system. In our paradigm the direction is reversed: robot hardware is used to teach a human new skills by physically moving the limbs and inducing the desired muscle activations. Nonetheless, some of the conventional teleoperation hardware and interfaces can be useful for implementing our proposed paradigm as well. Below we review various modern and historical teleoperation systems, and outline the relevant concepts.

Teleoperation has been most frequently used as an intuitive interface for humans to control robotic systems. For example, the developers of HaptX [7] show that their teleoperation system can be used to perform extremely fine manipulations using Shadow Hand hardware system [27] by relaying human finger movements over to the robot hand. Their system also provides haptic feedback and therefore the teleoperator can get a sense of the force needed to manipulate various objects. Advanced teleoperation systems can help with critical tasks, such as telesurgery. For example, it has been estimated that da Vinci surgical systems have already helped to complete more than 7 million surgical procedures [12].

Complex teleoperation hardware is usually expensive, but recently [6] introduced DexPilot — a low-cost teleoperation system that is also able to perform fine manipulations using a low-cost

Image included with permission from magazineart.org Image included with permission from Sarcos Robotics

Figure 2: Left: A 'remote doctor' concept on the cover of the 1924 "Radio News" magazine; source: magazineart.org [17]. Right: A recent powered full-body exoskeleton suite that allows users to lift heavy objects; source: sarcos.com.

Allegro hand [37]. The system operates directly with RGB cameras; it does not use haptic feedback as of now, but could benefit from it. Telexistence [32] is a company that designs robots to help with physical work in supermarkets and also aims to make their system accessible in terms of cost.

Teleoperation has also been used as a tool in shared autonomy with a goal of augmenting the capabilities of a robotic system. In such settings, a human is still always in the loop to ensure that a task is completed reliably, providing ingenuity for tackling difficult problems and correcting unwanted behaviours even in setups where a robot has partial autonomy. For example, a relevant application is space exploration. SpaceJustin robot developed by the German Aerospace Center (DLR) has been used for teleoperation to rehearse scenarios such as robots on a different planet being teleoperated by humans. An astronaut on board the International Space Station has remotely operated a humanoid robot to inspect and repair a solar farm on a simulated Mars environment set up in Munich [11]. Such setups are called *supervised autonomy* — a concept somewhere between full autonomy and direct teleoperation.

Teleoperation has also been a topic of interest in popular science fiction stories *e.g.* the 1942 short story "Waldo" by Robert Heinlein features a man who invents and teleoperates a device using his hand and fingers [8]. Below is a short extract from the story:

> *Waldo put his arms into the primary pair before him; all three pairs, including the secondary pair before the machine, came to life. Waldo flexed and extended his fingers gently; the two pairs of waldoes in the screen followed in exact, simultaneous parallelism.*

Such has been the popularity of this story that many real remote manipulators developed later also came to be called waldoes. See [24] for a history of teleoperators, exoskeletons and industrial robots.

3.2 Teaching Humans

Prior research in passive learning provides evidence that our paradigm could indeed bring tangible benefits for teaching humans. One example is a prototype of a music instruction system with fingerless gloves and vibrators on each finger that activates based on which finger is used to play a musical note [10]. Such passive haptic gloves can teach the user to build "muscle memory" to play piano. This system can also be used for passive haptic rehabilitation: helping people with partial spinal cord injury improve sensation and dexterity in their affected hands.

In contrast with the above prior work, our idea is to go further than re-playing pre-recorded signals. We aim to first let reinforcement learning (RL) agents figure out good strategies for solving a task without human input, then use policies from such agents to teach humans. For this, relevant prior work includes approaches in the area of AI for teaching/instruction. One example is a framework where an RL agent instructs students by suggesting actions the students can take as they learn and aims to find optimal moments for giving suggestions [33]. Another relevant approach is to define a parameterized space of instructional policies and search this space to identify an optimum [16]. This approach was extended to make the teaching policy adapt to the current student performance and also automatically identify which activities are beneficial for furthering the student learning progress [1]. Building on these ideas, we envision a tutoring system comprised of a set of RL agents, which learn to solve the given task autonomously (without the human input); we select among these agents based not only on how well an agent's policy solves the given task, but also based on how easy it is to teach the human to perform the task using each policy.

4 REVERSE TELEOPERATION: HARDWARE AND ALGORITHMS

To enable the robot-to-human transfer we need a safe and effective reverse teleoperation system. Figure 1 outlines the necessary components. We start by setting up a task in simulation, specifying a general RL objective/reward function and the constraints that the learned policy must satisfy to be safe. We propose to keep objectives relatively high level, *e.g.* sparse rewards as opposed to dense, not too specific (no reward tweaking or demonstrations). This should allow creative solutions to emerge and avoid biasing RL to only seek solutions that humans can envision. It has been demonstrated that interesting behaviors emerge with sufficiently diversified large-scale training [2, 28]. These solutions may range from exploring novel shapes/designs (e.g. new tools) to finding new trajectories that solve the task, covering the a range of aspects to explore in simulation about the environment structure and the task. With recent progress in GPU-accelerated simulation, it is possible to train more than 10K environments in parallel on one GPU, even for advanced cases like ShadowHand [20].

Parallelizing across GPUs would enable scaling to 1 million parallel environments. We will aim to train a 'main' policy for an average human profile/size, then fine-tune it to accommodate humans of various ages, customizing for different levels of flexibility and other physical characteristics.

The next stage is a mechanism to actuate a person to follow trajectories proposed by the robot. The learned policies can be downloaded and run on the hardware that is worn by the user. This hardware would be composed of a sensing module to keep track of the human pose and muscle activities, and an actuation module that safely executes the next actions to move human limbs. When users resist the suggested movements, the controller would re-plan and adapt to ensure comfort and safety. Highly sensitive force-torque sensors that can detect such resistance reliably along with impedance control can enable this [13]. If user data is shared via

Figure 3: Exopulse Mollii Suit; source: exopulse.com [21].

a central repository, reverse teleoperation systems could improve user preference models based on feedback from similar users.

4.1 Exoskeletons

Exopulse Mollii Suit [21] is a full-body suit for neuromodulation. It includes 58 embedded electrodes positioned to stimulate 40 muscles in various parts of the body through low frequency electro stimulation. Mollii Suit is used to relax muscle spasms, activate muscles to increase blood circulation and prevent atrophy [22]. It can also alleviate tremors from Parkinson's disease. We envision extending a suite like this to activate fine muscles for reverse teleoperation. Experimenting with various electrode types that could support more targeted muscle activation could allow to extend this suite beyond its current applications. The fact that the suite is a medical device offers a safe starting point. A safe full-body exoskeleton has been recently introduced by Sarcos (rightmost in Figure 1). It provides power-assist for users who need to lift heavy objects, apply large torques to turn industrial valves, etc. We envision that this full-body exoskeleton could be combined with work on rehabilitation exoskeletons for limbs (e.g. [5]) to create a powerful yet agile system for human actuation. Work on 'teleoperating' users' own ankle prosthesis (middle in Figure 1) provides insights for how to sense user's fine motions, and could give insights on how users react to a part connected to their body being actuated [35].

4.2 Simulation and Reinforcement Learning

The recent success of RL has spurred massive interest in simulators to train robots. Today, high quality image rendering is tractable and physics simulators and renderers run on GPU gathering 100-1000x more experience compared to CPUs [20, 25]. Moreover, there is progress on simulating muscles and tendons [14]. As the scene complexity increases, so does the simulation time. Nonetheless, we anticipate that the growth of available computational resources will continue, enabling faster and higher fidelity simulation in the near future. In the meantime, we could employ approaches that function even in the presence of a moderate simulation-to-reality gap [9]. We also expect to see further improvements in RL algorithms. For example, RL could be combined with hierarchical reasoning, symbolic planning, model-based optimisation and other learning/robotics approaches to create data efficient hybrid methods. Hence, even if some tasks would be beyond the simulation fidelity and RL capabilities initially, we can count on improvements in the near future.

While we can expect to acquire enough computational resources for training flexible and imaginative RL agents, the need to teach a human being creates the pressure to make the overall teaching approach data-efficient. Hence, we envision a tutoring system that selects among various RL policies based not only on the task reward, but also based on how easy it is to teach the human using each policy. Algorithm 1 gives the outline. We start by setting up the desired task in simulation. Then, we train N RL agents in parallel, and for each policy π_n record mean episode reward w_{π_n} after training. We then sample a policy π from N trained policies using weights w_{π_n}. After that, we run reverse teleoperation using π and periodically lighten the forces exerted by the exoskeleton to see whether the human keeps following the trajectory that π would follow. We can treat these periods of motion as observations of whether the human has made progress towards learning to mimic policy π. We can therefore use Bayesian knowledge tracing (BKT) [3] to compute the probability that the human has reached the state 'trained' for policy π. We then evaluate how well the human can complete the task without exoskeleton forces and update the weight for π accordingly, then re-sample the next policy to try for instruction.

Algorithm 1: Adaptive Reverse Teleoperation Tutor

Set up target task in sim. and train N parallel RL agents
Init policy weights: $w_{\pi_n} \leftarrow reward(\pi_n|task)$ for $n = 1..N$
$\mathcal{B} \leftarrow$ default Bayesian knowledge tracing (BKT) params
$O \leftarrow \{\}$; $t \leftarrow 0$
while $t < human\_time\_budget$ **do**
    $\pi \leftarrow$ sample RL policy (using w_{π_n} weights)
    $obs \leftarrow \{\}$
    for $k = 1..K$ **do**
        Run reverse teleop to train the human to mimic π
        $o_k \leftarrow$ observe teaching progress: lighten exoskeleton forces, see if human actions and π diverge
        $obs \leftarrow obs \cup o_k$
        if $P(state_{human} = trained|\mathcal{B}, obs) > thesh.$ **then**
            break
    $w_{\pi_n} \leftarrow \alpha \cdot w_{\pi_n} + (1 - \alpha) \cdot reward(\pi_{human}|task)$
    $O \leftarrow O \cup obs$; (re)-train BKT model using O
    $t \leftarrow t + elapsed\_time$

5 CONCLUSION

We hope that our paradigm of reverse teleoperation will offer a way to teach users "muscle memory" for various everyday tasks, help design new tools and teach humans to use them, improve medical surgery and rehabilitation. We also imagine that deployment of such a system at scale would offer a chance to continually improve it, leveraging user data that would be uploaded regularly to a common repository. Any user can therefore benefit from others that use the system across the world, yielding a direct consumer-to-consumer model. Furthermore, as more users deploy this system for various innovative use cases, the repository would benefit from an ever increasing category of skills, leading to reduced adaptation time and faster learning for new users. Ultimately, this offers a new digital layer for humans to improve learning of new skills, where algorithmic creativity in simulation, together with kinesthetic movements in the real world, provide a fundamentally new way to transform the teaching/learning process for physical skills.

REFERENCES

[1] Rika Antonova, Joe Runde, Min Hyung Lee, and Emma Brunskill. 2016. Automatically learning to teach to the learning objectives. In *Proceedings of the third (2016) ACM conference on learning@ scale*. 317–320.

[2] Bowen Baker, Ingmar Kanitscheider, Todor Markov, Yi Wu, Glenn Powell, Bob McGrew, and Igor Mordatch. 2020. Emergent Tool Use From Multi-Agent Autocurricula. *International Conference on Learning Representations (ICLR)* (2020).

[3] Albert T Corbett and John R Anderson. 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction* 4, 4 (1994), 253–278.

[4] Raphael Deimel, Patrick Irmisch, Vincent Wall, and Oliver Brock. 2017. Automated co-design of soft hand morphology and control strategy for grasping. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 1213–1218.

[5] Patrick W Franks, Nicholas A Bianco, Gwendolyn M Bryan, Jennifer L Hicks, Scott L Delp, and Steven H Collins. 2020. Testing simulated assistance strategies on a hip-knee-ankle exoskeleton: a case study. In *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 700–707.

[6] Ankur Handa, Karl Van Wyk, Wei Yang, Jacky Liang, Yu-Wei Chao, Qian Wan, Stan Birchfield, Nathan D. Ratliff, and Dieter Fox. 2020. DexPilot: Vision Based Teleoperation of Dexterous Robotic Hand-Arm System. In *International Conference on Robotics and Automation (ICRA)*. IEEE.

[7] HaptX. 2021. HaptX Gloves. https://haptx.com.

[8] Robert A. Heinlein. 2021. Waldo (short story). https://en.wikipedia.org/wiki/Waldo_(short_story).

[9] Sebastian Höfer, Kostas Bekris, Ankur Handa, Juan Camilo Gamboa, Florian Golemo, Melissa Mozifian, Chris Atkeson, Dieter Fox, Ken Goldberg, John Leonard, et al. 2020. Perspectives on sim2real transfer for robotics: A summary of the R: SS 2020 workshop. *arXiv preprint arXiv:2012.03806* (2020).

[10] Kevin Huang, Thad Starner, Ellen Do, Gil Weinberg, Daniel Kohlsdorf, Claas Ahlrichs, and Ruediger Leibrandt. 2010. Mobile Music Touch: Mobile Tactile Stimulation for Passive Learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.

[11] IEEE Spectrum. 2017. ISS Astronauts Operating Remote Robots Show Future of Planetary Exploration. https://spectrum.ieee.org/automaton/robotics/space-robots/iss-astronauts-operating-remote-robots-show-future-of-planetary-exploration.

[12] Intuitive Surgical. 2021. Da Vinci Sytems. https://www.davincisurgery.com.

[13] Arvid QL Keemink, Herman van der Kooij, and Arno HA Stienen. 2018. Admittance control for physical human–robot interaction. *The International Journal of Robotics Research* (2018).

[14] Seunghwan Lee, Moonseok Park, Kyoungmin Lee, and Jehee Lee. 2019. Scalable muscle-actuated human simulation and control. *ACM Transactions On Graphics (TOG)* (2019).

[15] Thomas Liao, Grant Wang, Brian Yang, Rene Lee, Kristofer Pister, Sergey Levine, and Roberto Calandra. 2019. Data-efficient learning of morphology and controller for a microrobot. In *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2488–2494.

[16] Robert V Lindsey, Michael C Mozer, William J Huggins, and Harold Pashler. 2013. Optimizing instructional policies. In *Advances in neural information processing systems (NeurIPS)*. 2778–2786.

[17] Magazineart.com. 2021. Magazine cover of the 1924-04 edition of "Radio News". http://www.magazineart.org/piwigo/picture.php?/10291/category/155.

[18] MIT Technology Review. 2018. The US military wants AI to dream up weird new helicopters. https://www.technologyreview.com/2018/03/19/144587/the-us-military-wants-ai-to-dream-up-weird-new-helicopters/. Accessed: 2021-05-30.

[19] Frédérique Muyl, Laurent Dumas, and Vincent Herbert. 2004. Hybrid method for aerodynamic shape optimization in automotive industry. *Computers & Fluids* 33, 5-6 (2004), 849–858.

[20] NVIDIA. 2021. IsaacGym. https://developer.nvidia.com/isaac-gym.

[21] Ottobock. 2021. Exopulse Mollii Suit. https://www.exopulse.com.

[22] Susanne Palmcrantz, Gaia Valentina Pennati, Hanna Bergling, and Jörgen Borg. 2020. Feasibility and potential effects of using the electro-dress Mollii on spasticity and functioning in chronic stroke. *Journal of NeuroEngineering and Rehabilitation* 17, 1 (2020), 1–10.

[23] Panagiotis Polygerinos, Zheng Wang, Kevin C Galloway, Robert J Wood, and Conor J Walsh. 2015. Soft robotic glove for combined assistance and at-home rehabilitation. *Robotics and Autonomous Systems* (2015).

[24] Reuben Hoggett. 2016. Teleoperation History. http://cyberneticzoo.com/early-teleoperators.

[25] Manolis Savva, Abhishek Kadian, Oleksandr Maksymets, Yili Zhao, Erik Wijmans, Bhavana Jain, Julian Straub, Jia Liu, Vladlen Koltun, Jitendra Malik, , Devi Parikh, and Dhruv Batra. 2019. Habitat: A platform for embodied ai research. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 9339–9347.

[26] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. 2020. Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* 588, 7839 (2020), 604–609.

[27] Shadow Robot Company. 2021. Shadow Hand. https://www.shadowrobot.com/dexterous-hand-series.

[28] Archit Sharma, Michael Ahn, Sergey Levine, Vikash Kumar, Karol Hausman, and Shixiang Gu. 2020. Emergent real-world robotic skills via unsupervised off-policy reinforcement learning. *arXiv preprint arXiv:2004.12974* (2020).

[29] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 7587 (2016), 484–489.

[30] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. 2017. Mastering the game of go without human knowledge. *Nature* 550, 7676 (2017), 354–359.

[31] Thad Starner. 2017. Passive Haptic Learning: Learn to Type or Play Piano Without Attention Using Wearables. https://gvu.gatech.edu/research/projects/passive-haptic-learning-learn-type-or-play-piano-without-attention-using-wearables.

[32] Telexistence Inc. 2021. Telexistence. https://tx-inc.com.

[33] Lisa Torrey and Matthew Taylor. 2013. Teaching on a budget: Agents advising agents in reinforcement learning. In *Proceedings of the International Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*. 1053–1060.

[34] Marc A Toussaint, Kelsey Rebecca Allen, Kevin A Smith, and Joshua B Tenenbaum. 2018. Differentiable physics and stable modes for tool-use and manipulation planning. In *Robotics: Science and Systems (RSS)*.

[35] Cara G. Welker, Vincent L. Chiu, Alexandra S. Voloshina, Steven H. Collins, and Allison M. Okamura. 2021. Teleoperation of an Ankle-Foot Prosthesis With a Wrist Exoskeleton. *IEEE Transactions on Biomedical Engineering* 68, 5 (2021), 1714–1725.

[36] Wikimedia. 2021. High jump. https://en.wikipedia.org/wiki/High_jump.

[37] Wonik Robotics. 2021. Allegro Hand. https://www.wonikrobotics.com/research-robot-hand.