

# DayDreamer: An algorithm to quickly teach robots new behaviors in the real world

July 27 2022, by Ingrid Fadelli

---



After learning to walk in 1 hour, we start applying external perturbations to the quadruped robot. While being fragile initially, the robot learns to adapt to withstand pushes or quickly roll back on its feet within 10 minutes of continued learning. Credit: Wu et al.

Training robots to complete tasks in the real-world can be a very time-consuming process, which involves building a fast and efficient simulator, performing numerous trials on it, and then transferring the

behaviors learned during these trials to the real world. In many cases, however, the performance achieved in simulations does not match the one attained in the real-world, due to unpredictable changes in the environment or task.

Researchers at the University of California, Berkeley (UC Berkeley) have recently developed DayDreamer, a tool that could be used to train robots to complete [real-world](#) tasks more effectively. Their approach, introduced in a paper pre-published on arXiv, is based on learning models of the world that allow robots to predict the outcomes of their movements and actions, reducing the need for extensive trial and error training in the real-world.

"We wanted to build robots that continuously learn directly in the real world, without having to create a simulation environment," Danijar Hafner, one of the researchers who carried out the study, told TechXplore. "We had only learned world models of video games before, so it was super exciting to see that the same algorithm allows robots to quickly learn in the real world, too!"

Using their approach, the researchers were able to efficiently and quickly teach robots to perform specific behaviors in the real world. For instance, they trained a robotic dog to roll off its back, stand up and walk in just one hour.

After it was trained, the team started pushing the robot and found that, within 10 minutes, it was also able to withstand pushes or quickly roll back on its feet. The team also tested their tool on robotic arms, training them to pick up objects and place them in specific places, without telling them where the objects were initially located.

"We saw the robots adapt to changes in lighting conditions, such as shadows moving with the sun over the course of a day," Hafner said.

"Besides learning quickly and continuously in the real world, the same algorithm without any changes worked well across the four different robots and tasks. Thus, we think that world models and online adaptation will play a big role in robotics going forward."

Computational models based on reinforcement learning can teach robots behaviors over time, by giving them rewards for desirable behavior, such as good object grasping strategies or moving at a suitable velocity. Typically, these models are trained through a lengthy trial and error process, using both simulations that can be sped up and experiments in the real world.

On the other hand, Dreamer, the algorithm developed by Hafner and his colleagues, builds a world [model](#) based on its past "experiences." This world model can then be used to teach robots new behaviors based on "imagined" interactions. This significantly reduces the need for trials in real-world environment, thus substantially speeding up the training process.

"Directly predicting future sensory inputs would be too slow and expensive, especially when large inputs like camera images are involved," Hafner said. "The world model first learns to encode its sensory inputs at each time step (motor angles, accelerometer measurements, camera images, etc.) into a compact representation. Given a representation and a motor command, it then learns to predict the resulting representation at the next time step."

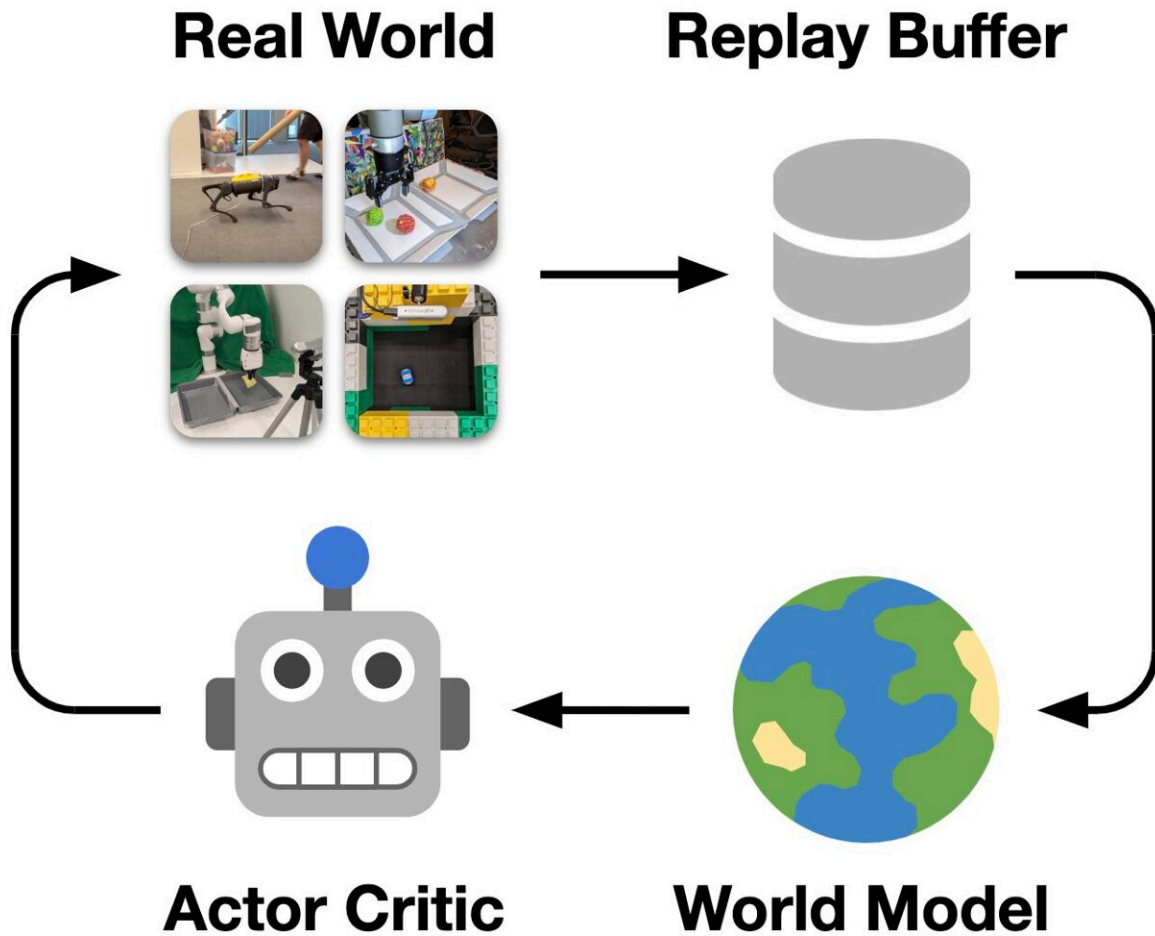
The world model produced by Dreamer allows robots to "imagine" future representations instead of processing raw sensory inputs. This in turn allows the model to plan thousands of [action sequences](#) in parallel, using a single graphics processing unit (GPU). These "imagined" sequences help to quickly improve the robots' performance on specific tasks.

"The use of latent features in reinforcement learning has been studied extensively in the context of representation learning; the idea being that one can create a compact representation of large sensory inputs (camera images, depth scans), thereby reducing model size and perhaps reducing the training time required," Alejandro Escontrela, another researcher involved in the study, told TechXplore. "However, representation learning techniques still require that the robot interact with the real world or a simulator for a long time to learn a task. Dreamer instead allows the robot to learn from imagined interaction by using its learned representations as an accurate and hyper efficient 'simulator.' This enables the robot to perform a huge amount of training within the learned world model."

While training robots, Dreamer continuously collects new experiences and uses them to enhance its world model, thus improving the robots' behavior. Their method allowed the researchers to train a quadruped robot to walk and adapt to specific environmental stimuli in only one hour, without using a simulator, which had never been achieved before.

"In the future, we imagine that this technology will enable users to teach robots many new skills directly in the real world, removing the need to design simulators for each task," Hafner said. "It also opens the door for building robots that adapt to hardware failures, such as being able to walk despite a broken motor in one of the legs."

In their initial tests, Hafner, Escontrela, Philip Wu and their colleagues also used their method to train a robot to pick up objects and place them in specific places. This task, which is performed by human workers in warehouses and assembly lines every day, can be difficult for robots to complete, particularly when the position of the objects they are expected to pick up is unknown.



Dreamer follows a simple pipeline for online learning on physical robots, without the need for simulators. Interaction with the real world is added to the replay buffer that stores all past experiences. A world model learns on sequences taken from the replay buffer at random. The behavior learns from predictions of the world model using an "actor critic" algorithm. The current behavior is used to interact with the world to collect new experiences, closing the loop. Credit: Wu et al.

"Another difficulty associated with this task is that we cannot give intermediate feedback or reward to the robot until it has actually grasped something, so there is a lot for the robot to explore without intermediate

guidance," Hafner said. "In 10 hours of fully autonomous operation, the robot trained using Dreamer approached the performance of human teleoperators. This result suggests world models as a promising approach for automating stations in warehouses and assembly lines."

In their experiments, the researchers successfully used the Dreamer algorithm to train four morphologically different robots on various tasks. Training these robots using conventional reinforcement learning typically requires substantial manual tuning, performed well across tasks without additional tuning.

"Based on our results, we are expecting that more robotics teams will start using and improving Dreamer to solve more challenging robotics problems," Hafner said. "Having a [reinforcement learning](#) algorithm that works out of the box gives teams more time to focus on building the robot hardware and on specifying the tasks they want to automate with the world model."

The algorithm can easily be applied to robots and its code will soon be [open source](#). This means that other teams will soon be able to use it to train their own robots using world models.

Hafner, Escontrela, Wu and their colleagues would now like to conduct new experiments, equipping a quadruped robot with a camera so that it can learn not only to walk, but also to identify nearby objects. This should allow the robot to tackle more complex tasks, for instance avoiding obstacles, identifying objects of interest in its environment or walking next to a human user.

"An open challenge in robotics is how users can intuitively specify tasks for robots," Hafner added. "In our work, we implemented the reward signals that the [robot](#) optimizes as Python functions, but ultimately it would be nice to teach robots from human preferences by directly telling

them when they did something right or wrong. This could happen by pressing a button to give a reward or even by equipping the robots with an understanding of human language."

So far, the team only used their algorithm to train robots on [specific tasks](#), which were clearly defined at the beginning of their experiments. In the future, however, they would also like to train robots to explore their environment without tackling a clearly defined task.

"A promising direction would be to train the robots to explore their surroundings in the absence of a task through artificial curiosity, and then later adapt to solve tasks specified by users even faster," Hafner added.

**More information:** Philipp Wu et al, DayDreamer: world models for physical robot learning. arXiv:2206.14176v1 [cs.RO], [arxiv.org/abs/2206.14176](https://arxiv.org/abs/2206.14176)

© 2022 Science X Network

Citation: DayDreamer: An algorithm to quickly teach robots new behaviors in the real world (2022, July 27) retrieved 4 May 2024 from <https://techxplore.com/news/2022-07-daydreamer-algorithm-quickly-robots-behaviors.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.