Summary: ”Crossing the Reality Gap: A Survey on Sim-to-Real Transferability of Robot Controllers in Reinforcement Learning” Salvato et. al.

1st Ian Sinclair
Ritchie School of Engineering and Computer Science
University of Denver
ian.sinclair@du.edu

Abstract*—Advanced robotic behavior is difficult to actualize using regular kinematics and control laws. Either from real-world models being intractable, or the behavior not fitting well into known control schemas. This is a prominent reason for the success of Reinforcement learning (RL) as a way of designing robotic systems. In general, RL models fit a map between states and actions through a series of trial-and-error attempts. This resolves the need for extremely accurate kinematic models and complex laws of behavior. However, it does require many-many successive attempts for the agent to learn, which by time constraints is usually intractable in the real world. This is where simulation techniques that can run faster than real time becomes pivotal. Now, completing $10,000$ trials that would take a physical robot days to complete can be accomplished digitally in a few hours. However, there remains a question, how can we be sure that the behavior the robot learns in simulation can be transferred over to the real world? What if there is something in the physical environment that was not included in the simulation? Are RL models still valid or robust out of simulation? A paper by Salvato el. al. ”Crossing the Reality Gap: A Survey on Sim-to-Real Transferability of Robot Controllers in Reinforcement Learning” [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] approaches some of these challenges, by discussing the *reality gap* in RL models which is a ’mismatch’ on the performance of the control law in simulation vs the real world. Here we provide a summary of the paper by Salvato et. al. particularly highlighting their analysis on three main approaches to reduce the reality gap and add robustness to simulated models.

I. The Need for Simulation

On a high level, Reinforcement learning uncovers a map between states and actions; unfortunately, as problems become more complicated, usually the state and action spaces grow in size or complexity. This typically requires more and more trials to be completed before a ’good’ policy can be found. Also, in each of these trials, there is a potential for the robot to perform unsafe actions like hurtling itself into walls.... And so, [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] describes two main motivations for training in a simulated environment, 1) time: trials can be completed faster than real-time. And 2) safety: the robot can damage itself all it wants in simulation and can be instantly reset. Then, after the agent has learned a good behavior it can be transferred to the real-world robot.

II. A Fundamental Challenge in Simulation (Reality Gap)

According to [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] the more representative a simulation is to the actual robot’s environment the better the policy can be transferred to the real world. However, more accurate simulations usually come at a higher computation cost, hindering one of the most useful attributes of using a simulation (time). And so, less accurate models that run much faster are usually preferred. However, this may cause a phenomenon in which the behavior learned in simulation degrades when used in the physical environment. [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] describes this as the *Reality Gap*, or the difference between performance in simulation and the real world. And so there remains a question, can a fast simulation be used while maintaining a small reality gap?

III. The *State-of-the-art* Solutions to the Reality Gap

[Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] accumulates modern literature on ways to reduce the sim-to-real reality gap focusing exclusively on papers that handle explicitly the reality gap problem, are based in robot control, and employ RL techniques. Salvato et. al. identify three major approaches to deal with the reality gap, 1) domain randomization (DR), 2) adversarial RL (ARL), and 3) transfer learning (TL). All of which represent difference *state of the art* approaches to lowering the reality gap.

IV. Domain Randomization

Domain randomization deals with the concept of unforeseen factors appearing in the physical environment that were not present in simulation. These factors can throw the robot off course, making the actions that preformed ideally in simulation less effective and could potentially cause the robot to fail its task. To compensate for this, random perturbations are added to each action/state in the simulation. [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] Highlights the effect of this, by noting that during a training trial, the agent will select an action that is supposed to cause some known behavior, however, now the action is corrupted by some small disturbance that randomly throws the agent off course. Then, over the course of the training cycle, the policy gets adjusted to be more robust to these small disturbances. Then, even without known what factors will perturb the robot in the real world, its has already learned to correct for them in simulation.

V. Adversarial RL (This is my favorite type of RL)

This type of learning strategy pits two RL models, a protagonist and antagonist, against each other. The protagonist attempts to maximize its reward or rather, learns the best policy to solve a task, while the antagonist has control over the environment or over perturbations on the protagonists actions, and tries to minimize the protagonists rewards. Generally, the antagonists finds the perturbations that the protagonist is most sensitive too as a means of throwing it off course. And after many attempts the protagonist can learn to be more robust to challenging disturbances.

VI. Transfer Learning

[Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] notes that the previous two methods seek to improve the controller in simulation, under the assumption that a more robust controller should have a lower reality gap. While transfer learning uses two learning phases, the first generates a control law in simulation, the second on the physical robot. This allows the simulation to learn a basic concept of the correct policy and ideally remove behaviors that can cause damage to the robot. Then, the second learning phase is shorter and learns to overcome real-world disturbances.

VII. Quality of the Taxonomy

[Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] primarily focuses their analysis on papers that tackle reality gap using RL models. This narrows the scope of the review, and allows them to be very specific on the techniques benefiting sim-to-real transfer using RL algorithms. However, this also leaves a potential for missed overlap from reality gap solutions from non-RL methods. For example, solutions derived from regular control techniques that may also lower the reality gap of RL models. Otherwise, [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] references a strong combinations of sim-to-sim and sim-to-real experiments along with many variations on each algorithm to justify the effectiveness of each of the three main categories. I also appreciate that they transferred the notation of each paper into one formalism, so the theory behind each can be more easily compared.

VIII. Conclusion

Reinforcement learning allows control laws to be created with a laxer kinematic and environmental model by means of many trial and error attempts. However, this can be slow and dangerous if done in the real world. This is where simulation becomes pivital. However, [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] notes that there is likely a reality gap between simulation and the real world which may cause the robot to function poorly compared to its simulated agent. [Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] creates a taxonomy of three methods to overcome with reality gap, being domain randomization, adversarial RL, and transfer learning. Each of which tries of add robustness to the robots control policy meaning models can be trained in less complex simulations and maintain their performance in the real world.

References

[Salvato et al.(2021)Salvato, Fenu, Medvet, and Pellegrino] E. Salvato, G. Fenu, E. Medvet, and F. A. Pellegrino. Crossing the reality gap: A survey on sim-to-real transferability of robot controllers in reinforcement learning. *IEEE Access*, 9:153171–153187, 2021. doi: 10.1109/ACCESS.2021.3126658.