Attention and Inattention for Minimalist Robot Learners

It is an increasingly popular view that much of robotics can be "solved" by brute force scaling of data, compute, and models. While scaling is certainly important to explore[1], by itself, it ignores the centrality of resource constraints in robotics such as on time, energy, compute, and training data. Good design principles entail that robots should be no more complex than necessary. My research group pursues a parallel scientific effort to understand and exploit fundamental performance-resource trade-offs. Our first line of attack has been to imbue various modules of a robot learner with the ability to selectively identify and attend to task-relevant information.

- **Representations with Object-Centric Spatial Attention:** We have developed new vision-language representations[2;3] that permit easily inferring and providing feedback to a robot on its progress ("value function") towards an image or language goal, such as "place a bowl on the dish rack". These are trained using offline reinforcement learning[4] on human videos, permitting transfer to robotic manipulation in real environments such as kitchens. In parallel, we have developed a family of pre-trained object-centric unsupervised representations that capture a scene at many granularities, permitting a downstream actor to dynamically assemble task-relevant minimal representations that enable the learner to better attend to task-relevant information amidst clutter and distribution shifts[5–9]: e.g., we can seamlessly daisy-chain individual skills trained separately to execute a complex task such as "cook an eggplant" that involves a step-by-step recipe.
- **Decision Making & Learning with Temporal Attention:** Downstream of the representation, decision making can benefit from resource-aware selective attention to key instants during task learning and execution. Attending to key future events[10;11] and spatial regions[12] during prediction and planning mitigates compounding errors, improves image goal reaching task performance, and transfers better to new robots. For real-time dynamic tasks like moving object grasping in cluttered settings, we have successfully trained meta-controllers that dynamically determine "how much planning" (horizon and compute time) to do before plan execution[13]. Applied to past experiences, temporal attention improves dynamics model and policy learning[4;14–16]: e.g., learned dynamics models in reinforcement learning work better when their training is focused on the types of experiences most likely to be experienced by the robot in its immediate future[15].
- **Attentive Sensing and Exploration:** Sensing also comes with trade-offs: sensors mediate all the environment information available to the robot, but entail resource costs. We have trained robots to strategically sense task-relevant information through active sensing and exploration[17–22]: e.g., a robot looking to identify the category of an object can strategically rotate the object in its hand[17]. We have shown how robots might self-evaluate their task progress through such *interaction*[23], to improve themselves through reinforcement learning (Best Paper Award, CORL 2022); e.g. a robot can better learn how to tighten a screw by first learning how to check whether it is tight. Once the policy is trained, the checking policy and its extra costs are no longer required. This kind of efficiency improved efficiency through mastery can also be realized in other ways: we have shown that robots can learn to operate from fewer sensory inputs[24], by cleverly exploiting access to "privileged" sensors at training time. We are now studying the foundations of sensory requirements of robot learners: for example, we have shown that fundamental limits for model-based control under partial observability also predict the difficulty and sample complexity of *learned* robotic policies[25].

Recently, in response to advances in large vision and language foundation models, we have shown that such models can automate the process of learning resource-efficient robotic policies in simulation and transferring them to real robots: this involves designing environments[26], domain randomization[27], and reward functions[28] from simple text specifications. Our methods enable challenging and dynamic behaviors, such as a quadruped walking on a yoga ball[27]. We are now pursuing its logical end point: having observed a video of a task environment, can we automatically create a simulator and train policies for various tasks?

While I have emphasized efficiency above, we also seek to address other blind spots of the "scaling" approach to robotics. Our object-centric and language-grounded representations above are shared with humans, which can enable safe and trustworthy robot learning[14;29;30]. Furthermore, we are working to progressively ease the task of teaching robots new skills, from demonstrations[14;31] to image goals[2;12], language goals[2] and task descriptions[27]. We will continue over the next several years to pursue foundational understanding while also expanding the limits of robotic capabilities.

References

[1] Open X-Embodiment Collaboration and many authors (incl. **Dinesh Jayaraman**). Open X-Embodiment: Robotic learning datasets and RT-X models. *ICRA*, 2024. URL `https://robotics-transformer-x.github.io/`.

[2] Yecheng Jason Ma, Shagun Sodhani, **Dinesh Jayaraman**, Osbert Bastani, Vikash Kumar, and Amy Zhang. VIP: Towards universal visual reward and representation via Value-Implicit Pre-Training. *ICLR (top 25 percent)*, 2023.

[3] Yecheng Jason Ma, Vikash Kumar, Amy Zhang, Osbert Bastani, and **Dinesh Jayaraman**. LIV: Language-image representations and rewards for robotic control. *ICML*, 2023.

[4] Yecheng Jason Ma, Jason Yan, **Dinesh Jayaraman**, and Osbert Bastani. How far i'll go: Offline goal-conditioned reinforcement learning via f-advantage regression. *NeurIPS*, 2022.

[5] Jianing Qian and Dinesh Jayaraman. Object representations guided by optical flow. *NeurIPS 4th Robot Learning Workshop: Self-Supervised and Lifelong Learning*, 2021.

[6] Jianing Qian, Anastasios Panagopoulos, and **Dinesh Jayaraman**. Discovering deformable keypoint pyramids. *ECCV*, 2022.

[7] Junyao Shi*, Jianing Qian*, Yecheng Jason Ma, and **Dinesh Jayaraman**. Composing pre-trained object-centric representations for robotics from "what" and "where" foundation models. *ICRA*, 2024. URL `https://sites.google.com/view/pocr`.

[8] Jianing Qian, Anastasios Panagopoulos, and **Dinesh Jayaraman**. Recasting generic pretrained vision transformers as object-centric scene encoders for manipulation policies. *ICRA*, 2024. URL `https://sites.google.com/view/robot-soft/`.

[9] Task-oriented hierarchical object decomposition for visuomotor control. *CORL (under review)*, 2024.

[10] Dinesh Jayaraman, Frederik Ebert, Alexei A Efros, and Sergey Levine. Time-agnostic prediction: Predicting predictable video frames. *ICLR*, 2019.

[11] Karl Pertsch, Oleh Rybkin, Frederik Ebert, Dinesh Jayaraman, Chelsea Finn, and Sergey Levine. Long-horizon visual planning with goal-conditioned hierarchical predictors. *NeurIPS*, 2020.

[12] Edward S. Hu, Kun Huang, Oleh Rybkin, and **Dinesh Jayaraman**. Know thyself: Transferable visuomotor control through robot-awareness. *ICLR*, 2022.

[13] Yinsen Jia, Jingxi Xu, **Dinesh Jayaraman**, and Shuran Song. Learning a meta-controller for dynamic grasping. *CASE*, 2024.

[14] Chuan Wen, Jierui Lin, Jianing Qian, Yang Gao, and Dinesh Jayaraman. Keyframe-focused visual imitation learning. *ICML*, 2021.

[15] Yecheng Jason Ma, Kausik Sivakumar, Jason Yen, Osbert Bastani, and **Dinesh Jayaraman**. Learning policy-aware models for model-based reinforcement learning via transition occupancy matching. *L4DC*, 2023.

[16] Kaustubh Sridhar, Souradeep Dutta, **Dinesh Jayaraman**, James Weimer, and Insup Lee. Memory-consistent neural networks for imitation learning. *ICLR*, 2024.

[17] Dinesh Jayaraman and Kristen Grauman. Look-ahead before you leap: end-to-end active recognition by forecasting the effect of motion. *ECCV*, 2016.

[18] Dinesh Jayaraman and Kristen Grauman. Learning to look around: Intelligently exploring unseen environments for unknown tasks. *CVPR*, 2018.

[19] Santhosh K Ramakrishnan*, Dinesh Jayaraman*, and Kristen Grauman. Emergence of exploratory look-around behaviors through active observation completion. *Science Robotics*, 2019.

[20] Santhosh K Ramakrishnan, Dinesh Jayaraman, and Kristen Grauman. An exploration of embodied visual exploration. *IJCV*, 2021.

[21] Sriram Narayanan, **Dinesh Jayaraman**, and Manmohan Chandraker. Long-hot: A modular hierarchical approach for long-horizon object transport. *ICRA*, 2024.

[22] Edward Hu, Richard Chang, Oleh Rybkin, and **Dinesh Jayaraman**. Planning goals for exploration. *ICLR (top 25 percent) and Best Workshop Paper at CORL 2022 Robot Adaptation Workshop*, 2023.

[23] Kun Huang, Edward Hu, and **Dinesh Jayaraman**. Training robots to evaluate robots: Example-based interactive reward functions for policy learning. *CORL*, 2022.

[24] Edward Hu, James Springer, Oleh Rybkin, and **Dinesh Jayaraman**. Privileged sensing scaffolds reinforcement learning. *ICLR*, 2024.

[25] Jingxi Xu, Bruce Lee, Nikolai Matni, and Dinesh Jayaraman. How are learned perception-based controllers impacted by the limits of robust control? *L4DC*, 2021.

[26] Environment curriculum generation via large language models. *CORL (under review)*, 2024.

[27] Yecheng Jason Ma, William Liang, Hungju Wang, Sam Wang, Yuke Zhu, Linxi Fan, Osbert Bastani, and **Dinesh Jayaraman**. Dreureka: Language model guided sim-to-real transfer. *RSS*, 2024.

[28] Yecheng Jason Ma, William Liang, Guanzhi Wang, De-An Huang, Osbert Bastani, **Dinesh Jayaraman**, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Eureka: Human-level reward design via coding large language models. *ICLR*, 2024.

[29] Jesse Zhang, Brian Cheung, Chelsea Finn, Sergey Levine, and Dinesh Jayaraman. Cautious adaptation for reinforcement learning in safety-critical settings. *ICML*, 2020.

[30] Yecheng Jason Ma, Dinesh Jayaraman, and Osbert Bastani. Conservative offline distributional reinforcement learning. *NeurIPS*, 2021.

[31] Neha Das, Sarah Bechtle, Todor Davchev, Dinesh Jayaraman, Akshara Rai, and Franziska Meier. Model-based inverse reinforcement learning from visual demonstrations. *CORL*, 2020.