Deep Reinforcement Learning for Robotics:

A Survey of Real-World Successes

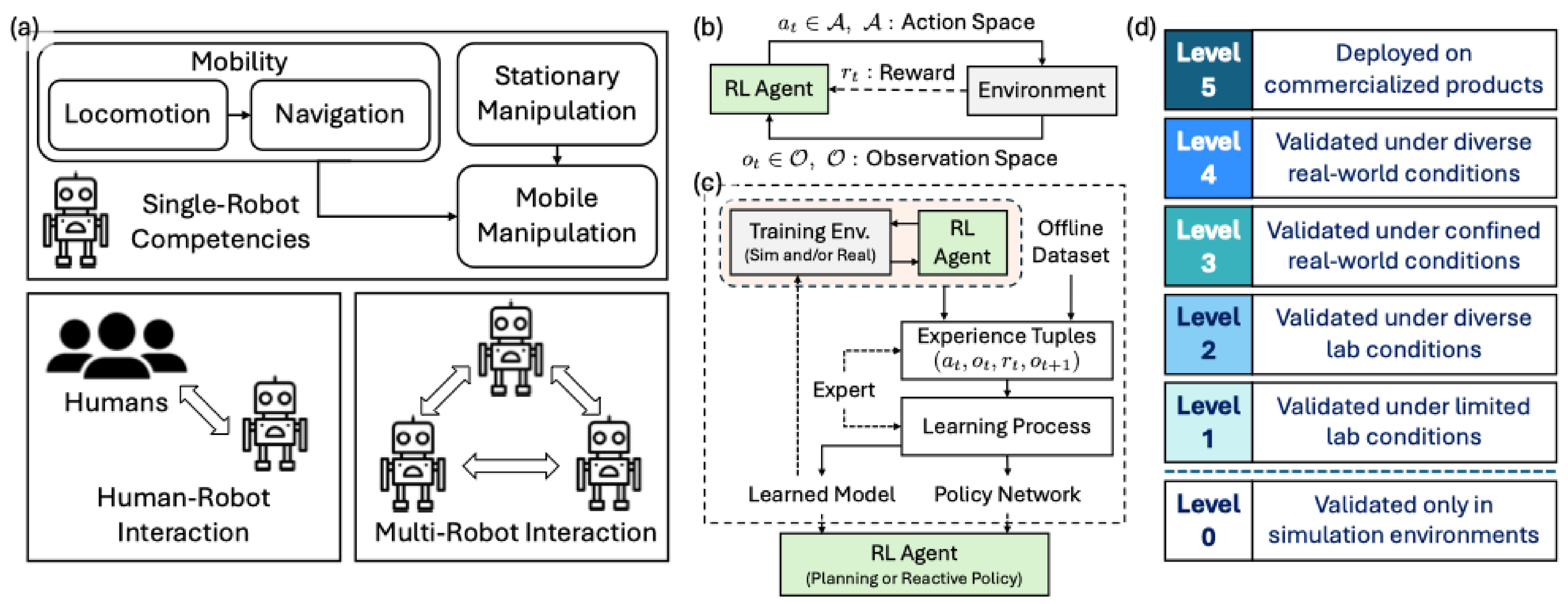
Chen Tang*[1], Ben Abbatematteo*[1], Jiaheng Hu*[1], Rohan Chandra[2], Roberto Martín-Martín[1], Peter Stone[1,3]

[1] The University of Texas at Austin [2] The University of Virginia, Charlottesville [3] Sony AI * Equal Contribution

Introduction

- DRL has achieved major successes in board games, video games, recommendation systems
- Controlling real-world robotic systems poses unique challenges
- Our survey evaluates current progress of DRL in robotics across various competencies, identifying *broadly applicable techniques*, *under-explored areas*, and *common open challenges*

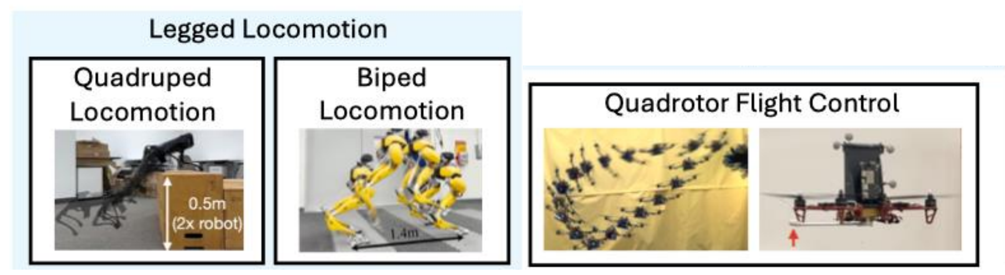
Taxonomy



(a) Robotic Competencies (b) Problem Formulation (c) Solution Method (d) Level of Success

Competency-Specific Review

Locomotion



- RL has enabled **mature quadrupedal locomotion**
  - Bipedal: dynamics are harder, higher DoF
- Key themes:
  - sim-to-real, heavy randomization, privileged info
- **Future Directions:**
  - Efficient & safe real-world learning
  - Integrating locomotion with downstream tasks, i.e., agile navigation or mobile manipulation

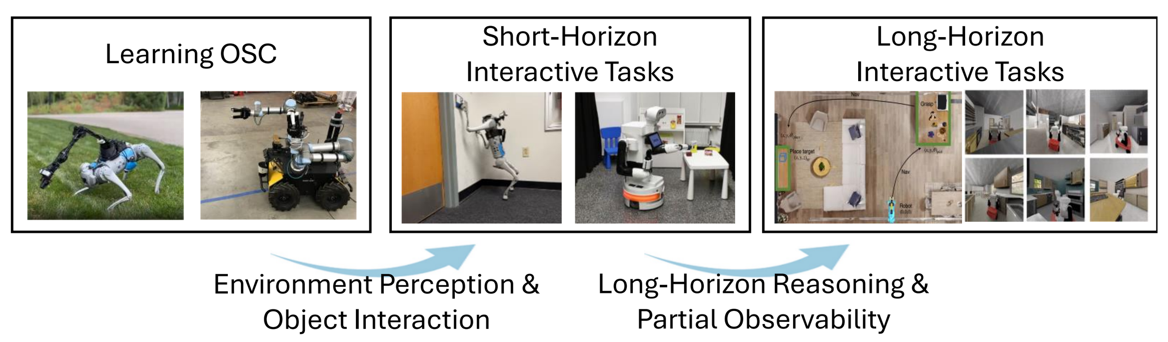
Navigation

- For indoor nav, end-to-end RL excels in simulation
- But, most successful real-world systems are **modular**
- Offline RL has shown promise for outdoor navigation
- **Highlight**: human-level drone racing
- **Future Directions:**
  - How much of the navigation stack should we learn?
  - Effectively jointly learn navigation & locomotion
  - Safety critical applications (e.g., autonomous driving)

Stationary Manipulation

- RL is more successful on more **constrained tasks**, enumerable a priori
  - E.g., grasping, in-hand manipulation, non-prehensile
  - Allows for zero-shot sim-to-real & dense reward
- **Future Directions:**
  - Integrating priors from classical robotics e.g., symmetry, geometry, collision-avoidance
  - Learning from human videos
  - Scaling to open-world manipulation

Mobile Manipulation

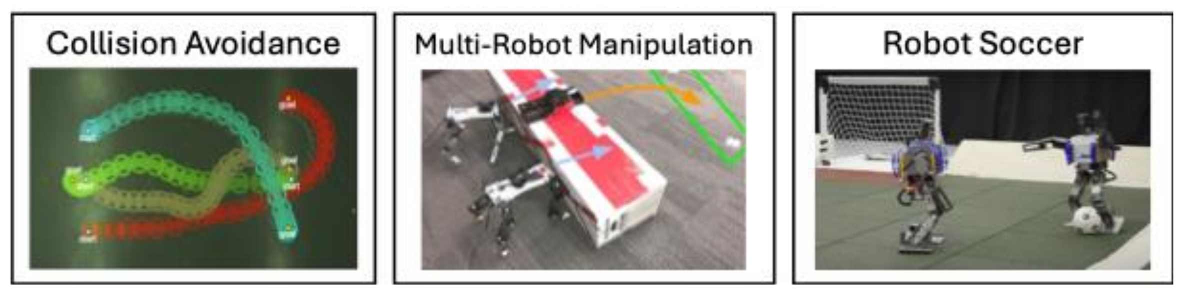


- **Some initial successes**, especially in short-horizon tasks, often sim-to-real
- Action space is critical, diverse morphologies
- **Future Directions:**
  - Multi-tasking
  - Long-term memory
  - Safe exploration

Human-Robot Interaction

- **Fewer successes** than "single-robot" competencies
- Hard to collect human-like data
  - Non-Markovian, limited rationality, expensive
- **Future Directions:**
  - Enable real-world learning alongside humans
  - Develop realistic human behavior simulation

Multi-Robot Interaction



- **Limited successes** in cooperative "homogeneous" settings
  - E.g., collision-avoidance
- Challenges in complexity & scalability
- **Future Directions:**
  - Communication between agents
  - Convergence & stability
  - General, non-cooperative settings

To appear in the 2025 Annual Review of Control, Robotics, and Autonomous Systems

General Trends

- **More mature domains:**
  - Quadrupedal locomotion, some navigation & manipulation tasks
- **Less mature domains:**
  - MoMA, HRI, Multi-robot
- **Mature solutions are commonly sim-to-real**
  - E.g., locomotion, grasping, in-hand manipulation
  - Dense, engineered reward functions
  - On-policy is feasible

Open Challenges

- **Stability** & **sample-efficiency** of RL algorithms
- **Real-world learning**
  - Gathering data: safe exploration, reward design, environment resets, sample efficiency
  - Hardware design for learning-based systems
  - Transfer, multi-task, meta- and lifelong learning
- **Long-horizon tasks**
  - What skills should the robot learn?
  - How should they be combined?
- **Principled approaches for RL systems**
  - E.g., reward design, action space choices
  - Integration with classical model-based tools
- **Benchmarking**: standard platforms and problems
- **Leveraging Foundation Models**
  - Path toward generalization, language-conditioning
  - Meta applications: reward design, simulation task and asset creation

AAAI-25 **Full Paper:**