

Giuseppe Rauso

Incremental Learning for Robotic Manipulation and Attention-Based Structured Task Execution

Tutor: prof. Alberto Finzi

Cycle: XXXIX

co-Tutor: prof. Vincenzo Lippiello

Year: Second

My background

- M.Sc. Degree in **Computer Science** (July 2023) @ DIETI – University of Naples Federico II
 - Thesis: *An Approach to Learning Robotic Manipulation Tasks based on Incremental Demonstration in a Virtual Environment*
- Research group: **PRISMA Lab** – DIETI Unina
- ITEE Ph.D. started in November 2023
 - PNRR Partenariato Esteso PE15, Agenzia Spaziale Italiana, Tematica 15, SpaceltUp – Research Theme: *Learning methods for exploration robotics. Started in 2023*

Summary of study activities

	Courses	Seminars	Research	Tutorship
First Year	23	6.6	35	0
Second Year	11	5.7	44	0.16
Expected	10 – 20 (30-70 total)	5 – 10 (10-30 total)	10 – 45 (80-140 total)	0 – 1.6 (0-4.8 total)

- Courses attended
 - How To Boost Your PhD
 - Planning and Navigation
- Conferences attended
 - 23rd International Conference of the Italian Association for Artificial Intelligence. AIxIA 2024 and AIRO 2024, Bolzano
 - 12th AIRO 2025 co-located with the Italian Institute of Robotics and Intelligent Machines (I-RIM 3D 2025), Rome

Research area

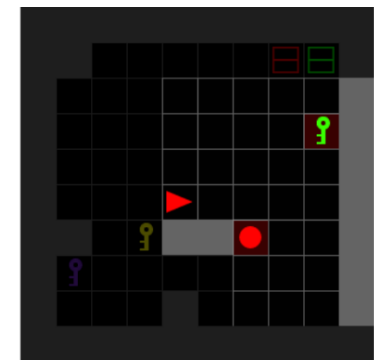
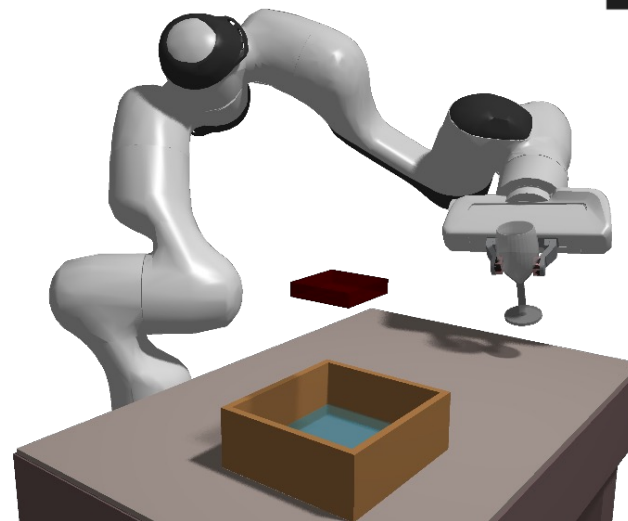
Learning of complex robotic tasks

Learning structured/complex robotic tasks, e.g., assembly and disassembly tasks or human-robot collaboration

Towards modular and extensible learning by integrating learning from demonstration and learned skill composition

Approach

- Incremental Learning
- Learning from Demonstration
- Attentional mechanisms



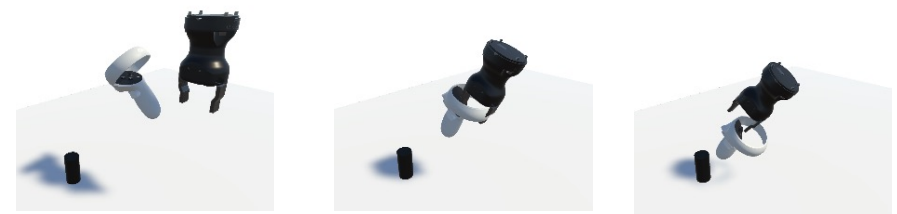
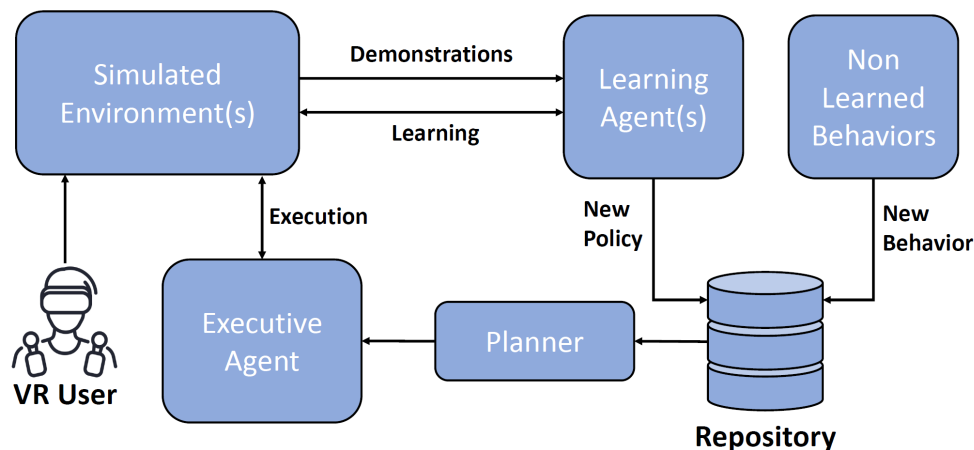
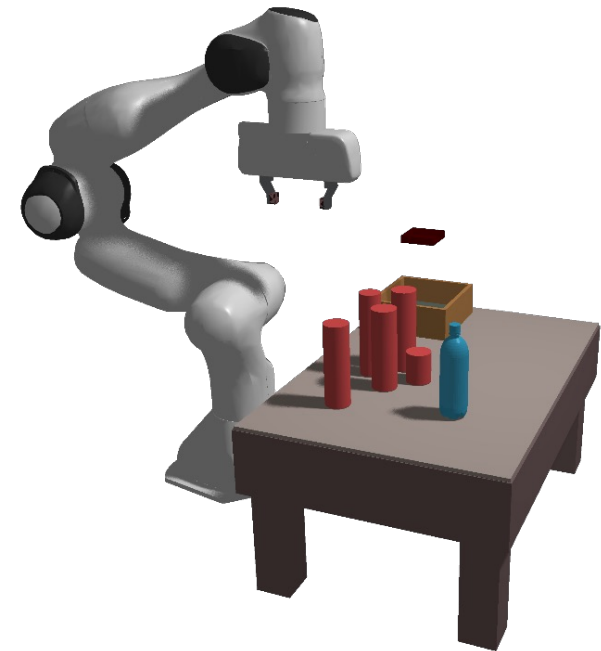
pick up the red ball and pick up the green key

Research field of interest

- Learning robotic manipulation tasks from *demonstrations*
 - Reinforcement Learning and Imitation Learning methods for robotics
 - Incremental Learning
 - Simulation and Interaction in Virtual Reality
- *Attentional mechanisms* and *executive control* for robot task learning
 - Multi-modal attention
 - Task-oriented learning and execution
 - Task attention
- Single-agent and multi-agent *robotic exploration*
 - Single-agent and multi-agent reinforcement learning for exploration
 - Multi-agent learning from demonstration

Research activity

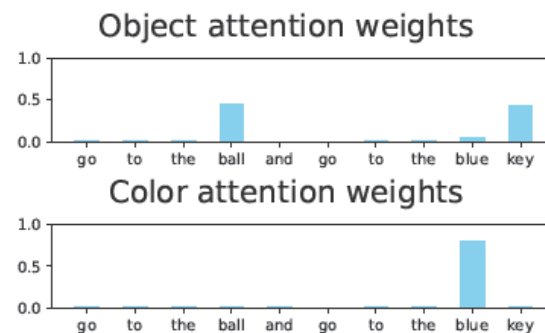
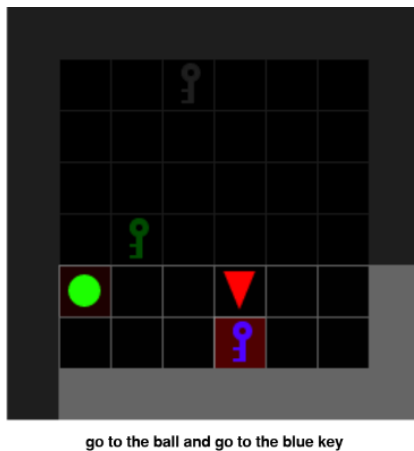
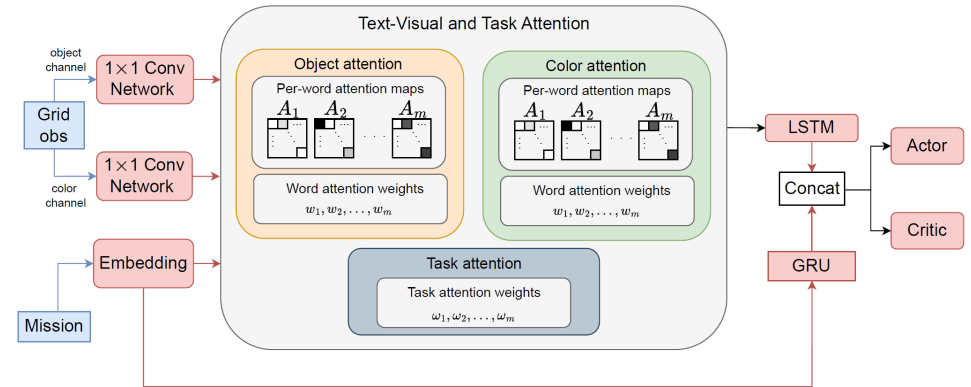
- *Combining Virtual Demonstrations and Incremental Skill Composition for Robotic Manipulation*
 - Incremental learning of primitive actions
 - VR demonstrations
 - HTN planning using learned and predefined skills



G. Rauso et al., "Incremental Learning from Virtual Demonstrations and Task Composition for Robotic Manipulation", RAS, under revision

Research activity

- *Task-Aware Multimodal Attention in Language-Guided Robotics*
 - Combined text-visual attention models
 - Task attention models

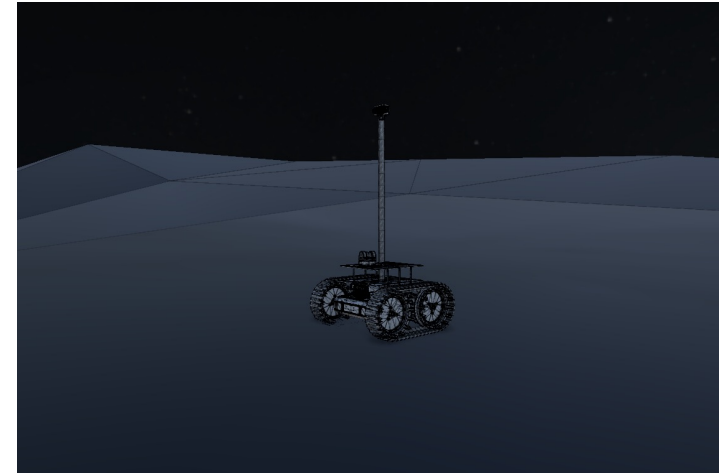


- Improved grounding of mission words
- Channel-wise attention calculation (objects and colors)

G. Rauso et al., "Task-Aware Multimodal Attention for Language-Guided Learning and Execution in Robotics", In preparation for submission to IEEE Trans. on Cog. and Devel. Systems

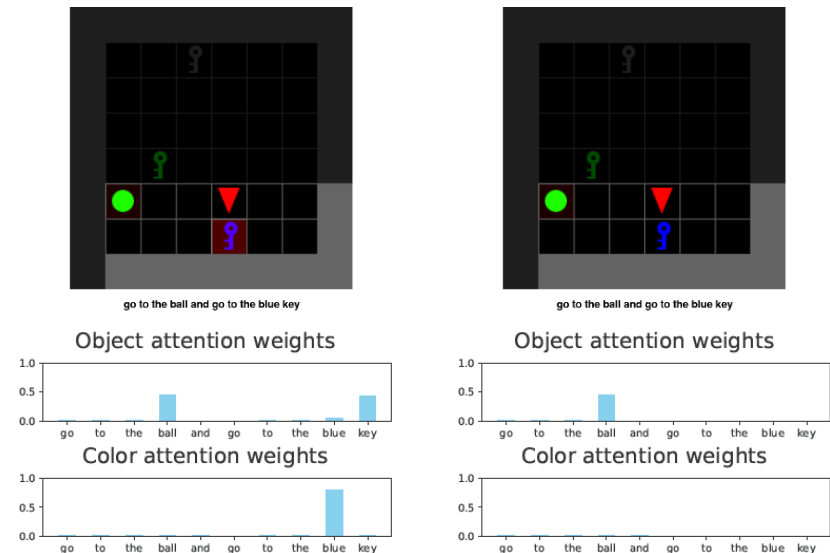
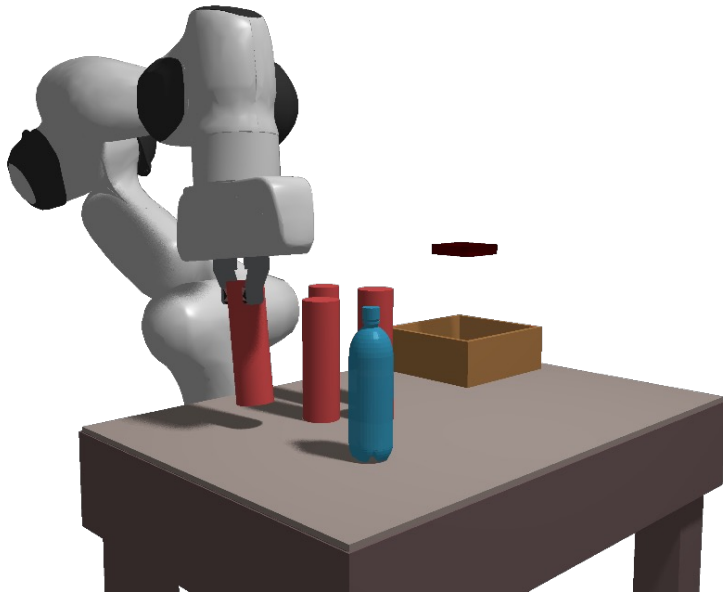
Research activity

- *Learning Lunar Exploration Tasks*
 - Reinforcement learning for exploration
 - Simulated lunar environment in Unity
- Sim-to-Real Transfer of Grid-Based Policies for Rover Exploration
 - Abstract grid-based environment representation
 - Reaching and retrieval with attentional policies



Research results

- Reinforcement learning can optimize rover behavior in environments with strict physical and energy constraints
- Multimodal attention mechanisms improve performance and enable grounding between instructions and observed features



- Learning from demonstration enables acquisition of primitive manipulation skills, reducing training time and improving performance
- Combining learning and planning with learned and predefined skills allows reuse of policies for complex long-horizon tasks

Research products

[P1]	G. Rauso, R. Caccavale, A. Finzi <i>Combined Text-Visual Attention Models for Robot Task Learning and Execution</i> 23rd International Conference of the Italian Association for Artificial Intelligence – AIxIA 2024, Bolzano AIxIA – Advances in Artificial Intelligence, Lecture Notes in Computer Science, volume 15450, Published 2025
[P2]	G. Rauso, R. Caccavale, V. Lippiello, A. Finzi <i>Integrating Text-Visual and Task Attention for Language-Guided Robot Learning</i> 11th Italian Workshop on Artificial Intelligence and Robotics co-located with AIxIA 2024 – AIRO 2024, Bolzano CEUR Workshop Proceedings, vol 3956, Published 2025
[P3]	G. Rauso, R. Caccavale, A. Finzi <i>Incremental Learning from Virtual Demonstrations and Task Composition for Robotic Manipulation</i> Robotics and Autonomous Systems (RAS) - Journal Submitted, currently in second review round (first round: minor revision)
[P4]	G. Rauso, R. Caccavale, A. Finzi <i>Task-Aware Multimodal Attention for Language-Guided Learning and Execution in Robotics</i> 39th Annual AAAI Conference on Artificial Intelligence - AAAI 2026, Singapore Not accepted. It is currently being extended for submission to the IEEE Transactions on Cognitive and Developmental Systems.

Research products

[P5]	R. Caccavale and S. Ciaravino, A. Finzi, V. Lippiello, G. Rauso <i>A Heterogeneous Multi-robot Framework for Cooperative Lunar Exploration</i> 28th AIDAA International Congress and the 10th CEAS Aerospace Europe Conference - CEAS-AIDAA, Turin Accepted
[P6]	G. Rauso, R. Caccavale, A. Finzi <i>A Unified Framework for Incremental Skill Acquisition and Symbolic Task Composition in Robotic Manipulation</i> 12th Italian Workshop on Artificial Intelligence and Robotics co-located with the Italian Institute of Robotics and Intelligent Machines (I-RIM 3D 2025) – AIRO 2025, Rome Accepted and presented
[P7]	A. Mora, G. Rauso, R. Caccavale, A. Finzi, R. Barber <i>Grounded Intent Validation: A Visual-Semantic Framework for Task Feasibility Assessment in Assistive Robotics</i> 2026 IEEE International Conference on Robotics & Automation - ICRA 2026, Vienna Submitted
[P8]	G. Rauso, R. Caccavale, A. Finzi <i>Combining Text-Visual and Task Attention for Language-Guided Robot Learning</i> IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots - SIMPAR 2025, Palermo Poster presentation (not included in the proceedings)

Research products

[P9]	Contribution to the development of deliverable D8.2.2 – Report on robotic systems and associated technologies & TRL assessment plan for the <i>SpaceItUp</i> project
[P10]	Prototype software: Extended Unity and CoppeliaSim simulations for robotic manipulation, enabling demonstration recording, end-effector-only grippers, and sim-to-sim policy transfer

Future Work

- Combine multimodal attentional mechanisms with learning of robotic manipulation and exploration tasks from demonstrations
- Leverage attentional grounding in complex environments with realistic visual inputs to guide the robot toward relevant scene elements
- Explore integration of foundation models (LLMs, VLMs, VLAs) to bridge language-based task definitions, human perception, and execution of learned or predefined robotic behaviors

Thank you for your
attention