

# We Choose to Go to Space: Agent-driven Human and Multi-Robot Collaboration in Microgravity

Miao Xin<sup>1\*</sup>, Zhongrui You<sup>2†</sup>, Zihan Zhang<sup>3†</sup>, Taoran Jiang<sup>4†</sup>, Tingjia Xu<sup>5†</sup>, Haotian Liang<sup>4†</sup>, Guojing Ge<sup>1</sup>, Yuchen Ji<sup>6†</sup>, Shentong Mo<sup>7</sup>, Jian Cheng<sup>1</sup>

<sup>1</sup>Institute of Automation Chinese Academy of Sciences <sup>2</sup>Beijing University of Posts and Telecommunications <sup>3</sup>TongJi University <sup>4</sup>Beihang University <sup>5</sup>Beijing Normal University <sup>6</sup>Nanjing University of Aeronautics and Astronautics <sup>7</sup>Carnegie Mellon University

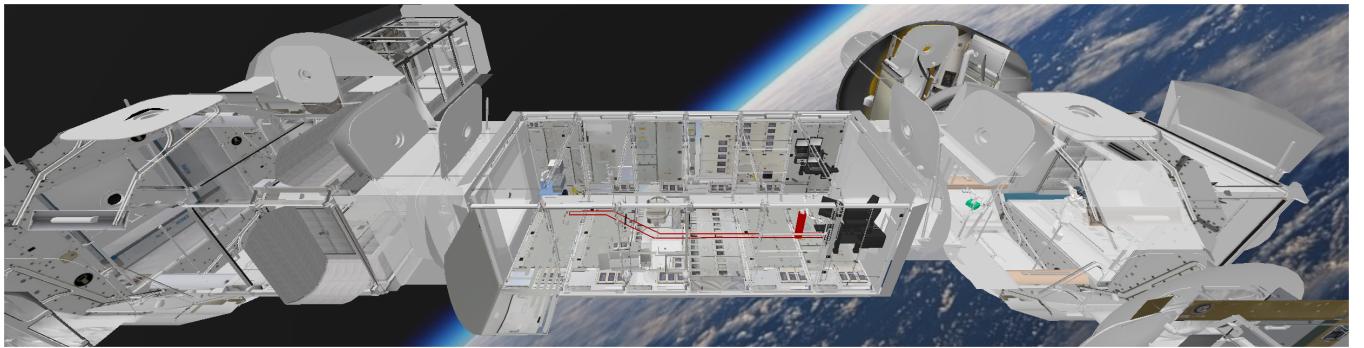


Figure 1: **SpaceAgents-1** simulates the human and multi-robot collaboration in microgravity under the control of AI Agents. In this figure, an astronaut is transporting cargoes across multiple modules with a free-flying robot, a rail-type robot and a dexterous robot working together.

## Abstract

We present **SpaceAgents-1**, a system for learning human and multi-robot collaboration (HMRC) strategies under microgravity conditions. Future space exploration requires humans to work together with robots. However, acquiring proficient robot skills and adept collaboration under microgravity conditions poses significant challenges within ground laboratories. To address this issue, we develop a microgravity simulation environment and present three typical configurations of intra-cabin robots. We propose a hierarchical heterogeneous multi-agent collaboration architecture: guided by foundation models, a Decision-Making Agent serves as a task planner for human-robot collaboration, while individual Skill-Expert Agents manage the embodied control of robots. This mechanism empowers the SpaceAgents-1 system to execute a range of intricate long-horizon HMRC tasks.

## 1 Introduction

We believe that human beings will eventually go into further space and to the outer planets. In the current maintenance

of space stations [Li *et al.*, 2022] and future space exploration [Ackerman *et al.*, 2022], a great deal of missions need to be carried out by both humans and robots. This makes embodied collaborations between machine intelligence and human intelligence essential. However, due to the particularity of microgravity, counterintuitive phenomena are difficult to obtain via experiments on Earth. Therefore, highly realistic simulation of microgravity and learning of HMRC is crucial.

In this work, we demonstrate the *SpaceAgent-1* system, designed to learn robot skills in microgravity, and explore the HMRC to accomplish various tasks inside the cabin. The system is driven by AI Agents [Durante *et al.*, 2024]. We designed a hierarchical multi-agent architecture to achieve collaborative control and embodied execution. A foundation model [OpenAI, 2023] provides the agents with multiple abilities such as planning, memory, action, and critique, allowing them to exhibit the potential to align with human skills. We also develop a platform, *SpaceSim*, for microgravity physics simulation and provide space robot simulators.

In summary, we explore a potential technical route that coordinates HMRC [Lippi *et al.*, 2022] in microgravity employing AI Agents and provide a simulation platform for verification, presenting a new endeavor for future research.

## 2 SpaceSim: Microgravity Simulation

The SpaceSim platform consists of three parts: *i.* microgravity simulator, *ii.* space robots, and *iii.* real2sim human-robot interaction (HRI).

\*Corresponding author (e-mail: miao.xin@ia.ac.cn). †Work done while they were Research Interns with CASIA.

**Microgravity simulation.** The 3D environment provides highly realistic physical simulations of microgravity and various phenomena it causes. The simulation engine of SpaceSim is built on SAPIEN [Xiang *et al.*, 2020], a part-based and physics-rich simulation environment that allows us to train robot skills based on reinforcement learning (RL) through the Gym interface.

**Intra-cabin space robots.** SpaceSim provides three types of robots (Figure 2). (1) **Free-flying robot (F)** can fly aerodynamically inside the space station, similar to *Astrobee* [Bualat *et al.*, 2018] robot but with limited cross-cabin dragging capability. (2) **Rail-type robot (R)** can move along the sliding rail and grasp objects with the robot arms, which is suitable for stable handling tasks within one module. (3) **Dexterous robot (D)** can perform delicate operational tasks, such as picking objects, or opening boxes.

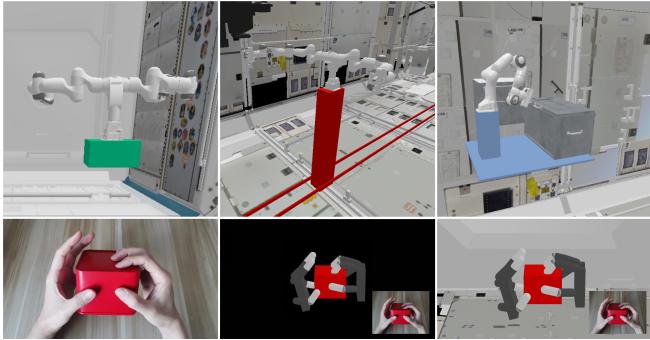


Figure 2: Robots in SpaceSim (up) and HRI simulation (bottom).

**Real2sim HRI interface.** With the hand-object interaction algorithm [Zhou *et al.*, 2020], **Human (H)** collaborators map real-world hand manipulations to the SpaceSim environment through visual teleoperation. Compared to using specialized equipment [Glauser *et al.*, 2019], computer vision-based hand-object motion capture has lower costs, enabling large-scale data collection.

### 3 SpaceAgents: Hybrid and Embodied Multi-agent Collaboration

#### 3.1 Hierarchical Architecture

HMRC typically performs in human-supervisor and robot-executor paradigm, or equal partnership paradigm [Xi *et al.*, 2023]. In contrast to both of these, we adopt a hierarchical mode. The **Decision-Making Agent (DMA)** controls the collaboration between various **Skill-Expert Agents (SEAs)** and humans, while each embodied SEA controls its corresponding robot. DMA focuses on global collaborative planning, but it is not embodied. In contrast, SEAs can accomplish embodied execution, but are more focused on executing specific skills. The advantage of this design lies in the decoupling of policies. When a certain SEA needs to acquire new skills, the system does not need to be updated as a whole.

#### 3.2 Working Flow

The Agent consists of a *Planner*, an *Actor* and a *Discriminator*. Figure 3 sketches the working flow of the system.

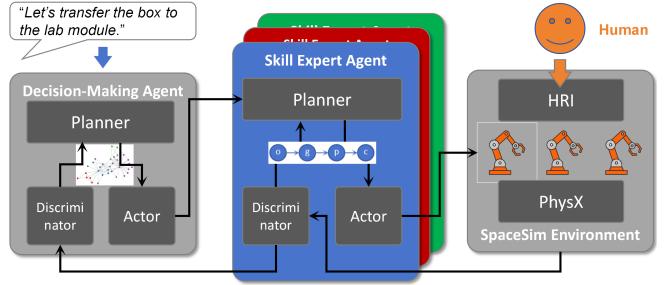


Figure 3: System working flow.

**Collaborative Planning.** Given a long-horizon task described in natural language, the DMA Planner first decomposes it into multiple subtasks and allocates each subtask to either a SEA (preferentially) or a human executor based on estimated affordance. To express the collaborative relationships among these subtasks, the DMA constructs a directed *collaboration graph* (CoG, Figure 4) within its working memory. The capability for collaborative task planning originates from the foundation model [OpenAI, 2023], which directly outputs the collaboration graph in JSON format guided by prompt engineering [Mo and Xin, 2023]. CoGs can represent both serial and parallel collaborations. Serial collaboration refers to a sequential order between task nodes, where a task node must wait until the preceding one is completed. Conversely, parallel nodes allow multiple subtasks to start simultaneously, with the next subtask initiating only after all preceding subtasks have concluded.

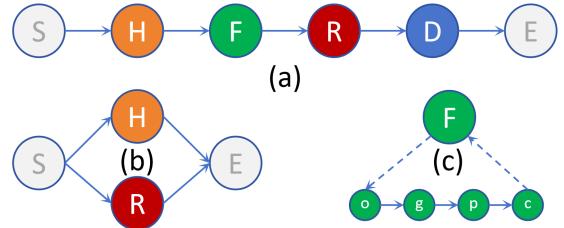


Figure 4: Serial (a) and parallel (b) subtask collaboration graphs. (c) A subtask is decomposed into a basic skill chain.

Subtasks are described in text instructions. For one subtask instruction, a Planner in the designated SEA further decomposes it into a *skill chain* [Mishra *et al.*, 2023] consisting of basic skills (*e.g.*, “(g)rasp”, “(o)pen”, “(p)ick”, etc.). A basic skill is an RL policy learned via Proximal Policy Optimization (PPO) [Schulman *et al.*, 2017], which resides in the long-term memory (skill library) of the SEA.

**Task Execution.** Actors transform instructions into actions. DMA Actor maintains communication between SEAs and ensures that the security boundaries of each SEA are not violated during execution. For each skill instruction in a skill chain, the Actor in a SEA invokes the corresponding RL policy to output the skill action vector directly to the robot. We utilize reward engineering to facilitate skill smoothing. The observation and training objectives of RL include not only the position of the termination point but also the instantaneous

state at the termination of the skill, such as the velocity, acceleration, and 6-DoF pose of the manipulated object. This ensures a smooth connection between the end of one skill and the beginning of the next.

**Result Evaluation.** Given the task instruction, SEA Discriminator generates a structured state description of current observations with the help of a vision-language model [Liu *et al.*, 2023] and identifies whether the current state matches the goal of the task. The Discriminator of DMA collects messages from each SEA Discriminator in a multi-process manner and provides the DMA Planner with judgments on whether the tasks of each collaborator are accomplished.

With this closed workflow, planning, execution, and discrimination form an agent-driven HMRC loop, navigating collaborators to traverse the collaboration graph sequentially and ultimately completing the task.

## 4 Long-horizon HMRC Manipulation

In our **demo video**<sup>1</sup>, we demonstrate certain interesting phenomena of weightlessness and basic skills of robots in the simulation environment. Specifically, we demonstrate two representative long-horizon HMRC cases.

### 4.1 Case 1: Floating Objects Rearrangement

In this demo (Figure 5), humans and robots work together to collect objects floating in the space station. To simplify, Planners use the object state information provided by the simulator, such as spatial position and size. DMA assigns subtasks to humans or robots based on the distance between the floating objects and the executors, thereby ensuring that there is no mutual interference between them.

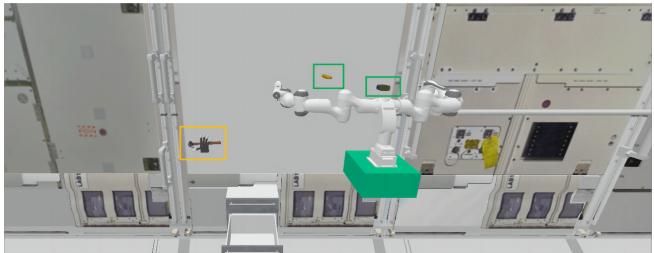


Figure 5: Floating objects rearrangement.

### 4.2 Case 2: Relay Object Transport

In long-horizon planning, Agents not only seek to minimize human workload as much as possible but also comprehensively consider the skill availability of different executors and environmental constraints. The prompts provide affordance information about the capabilities and limitations of different robots. For example, the rail-type robot can stably grasp objects but can only move along rails within a single module. The free-flying robot can pass through cabin connections but struggles with stably dragging objects over long distances. The dexterous robot can perform delicate operations but is limited to objects fixed on the tabletop. In this case (Figure 1),

humans collaborate with multiple robots under the control of SpaceAgents-1 to complete the task of transporting cargo between multiple modules, effectively reducing human labor.

## 4.3 Results

In Table 1, we compare the success rates of the same dexterous robot completing identical tasks under microgravity or gravity conditions [Gu *et al.*, 2023]. We observe that it is more difficult for robots to complete tasks in microgravity. It is almost impossible for robots to successfully utilize existing skill strategies directly in microgravity, which reflects the importance of learning in a simulation environment. The success rates of grasping skills for Free-flying robots and Rail-type robots reach 0.84.

Robots	P	T	CF	OC	CC	OB	PB
Dexterous (0g)	0.88	0.93	0.72	0.79	0.96	0.93	0.78
Dexterous (1g)	0.91	0.98	-	0.92	0.95	0.93	0.82

Table 1: Success rates of basic skills (**PickCube**, **ThrowCube**, **CatchFloatingCube**, **OpenCabinetDoor**, **CloseCabinetDoor**, **OpenBox**, **PickfromBox**) in microgravity (0g) or earth gravity (1g). “-” represents the robot does not have this skill.

We also compare the performance of SpaceAgents-1 with that of human experts (Table 2). With teleoperation, individuals act as the controllers of the robot instead of the agent. Across two collaborative tasks, SpaceAgents-1 demonstrates human-comparable performance in long-horizon task decomposition, attributable to the capabilities of the foundation model. However, regarding skill execution, humans encounter challenges in stably controlling robots in microgravity without prior training, contrasting starkly with terrestrial conditions. Nonetheless, a considerable number of failures are observed in the experiment, highlighting the irreplaceability of neither robots nor humans.

Tasks	Experts	SpaceAgents-1
Floating Objects Rearrangement	0.91	0.87
Relay Object Transport	0.46	0.39

Table 2: Success rates on long-horizon tasks.

## 5 Conclusion

This paper presents SpaceAgents-1, an HMRC system in microgravity. We design a hierarchical multi-agent collaboration architecture, which is driven by the foundation model to complete various tasks such as collaborative planning, embodied execution and effect evaluation. The space physics simulation environment we built allows the robots to learn operational skills and collaborative processes in microgravity at a very low cost. We will open-source this system in future work, hoping that this effort can inspire more researchers.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 61906195).

<sup>1</sup>Youtube Video (<https://youtu.be/GLSjUUtp32k>)

## References

- [Ackerman *et al.*, 2022] Evan Ackerman, Rahul Rao, and Julianne Pepitone. Robotics: Meet the lunar gateway’s robot caretakers: With people seldom on board, the space station will rely on autonomy. *IEEE Spectrum*, 59:7–13, 2022.
- [Bualat *et al.*, 2018] Maria G Bualat, Trey Smith, Ernest E Smith, Terrence Fong, and DW Wheeler. Astrobee: A new tool for iss operations. In *2018 SpaceOps Conference*, page 2517, 2018.
- [Durante *et al.*, 2024] Zane Durante, Qiuyuan Huang, Naoki Wake, Ran Gong, Jae Sung Park, Bidipta Sarkar, Rohan Taori, Yusuke Noda, Demetri Terzopoulos, Yejin Choi, et al. Agent ai: Surveying the horizons of multimodal interaction. *arXiv preprint arXiv:2401.03568*, 2024.
- [Glauser *et al.*, 2019] Oliver Glauser, Shihao Wu, Daniele Panozzo, Otmar Hilliges, and Olga Sorkine-Hornung. Interactive hand pose estimation using a stretch-sensing soft glove. *ACM Transactions on Graphics (TOG)*, 38:1 – 15, 2019.
- [Gu *et al.*, 2023] Jiayuan Gu, Fanbo Xiang, Xuanlin Li, Zhan Ling, Xiqiang Liu, Tongzhou Mu, Yihe Tang, Stone Tao, Xinyue Wei, Yunchao Yao, Xiaodi Yuan, Pengwei Xie, Zhiao Huang, Rui Chen, and Hao Su. Maniskill2: A unified benchmark for generalizable manipulation skills. In *International Conference on Learning Representations (ICLR)*, 2023.
- [Li *et al.*, 2022] Xueai Li, Dapeng Yang, and Hong Liu. China’s space robotics for on-orbit servicing: the state of the art. *National Science Review*, 10, 2022.
- [Lippi *et al.*, 2022] Martina Lippi, Paolo Di Lillo, and Alessandro Marino. A task allocation framework for human multi-robot collaborative settings. *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7614–7620, 2022.
- [Liu *et al.*, 2023] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning. In *NeurIPS*, 2023.
- [Mishra *et al.*, 2023] Utkarsh Aashu Mishra, Shangjie Xue, Yongxin Chen, and Danfei Xu. Generative skill chaining: Long-horizon skill planning with diffusion models. In *7th Annual Conference on Robot Learning (CoRL)*, 2023.
- [Mo and Xin, 2023] Shentong Mo and Miao Xin. Tree of uncertain thoughts reasoning for large language models. *ArXiv*, abs/2309.07694, 2023.
- [OpenAI, 2023] OpenAI. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.
- [Schulman *et al.*, 2017] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *ArXiv*, abs/1707.06347, 2017.
- [Xi *et al.*, 2023] Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, Rui Zheng, Xiaoran Fan, Xiao Wang, Limao Xiong, Qin Liu, Yuhao Zhou, Weiran Wang, Changhao Jiang, Yicheng Zou, Xiangyang Liu, Zhangyue Yin, Shihan Dou, Rongxiang Weng, Wensen Cheng, Qi Zhang, Wenjuan Qin, Yongyan Zheng, Xipeng Qiu, Xuanjing Huan, and Tao Gui. The rise and potential of large language model based agents: A survey. *ArXiv*, abs/2309.07864, 2023.
- [Xiang *et al.*, 2020] Fanbo Xiang, Yuzhe Qin, Kaichun Mo, Yikuan Xia, Hao Zhu, Fangchen Liu, Minghua Liu, Hanxiao Jiang, Yifu Yuan, He Wang, Li Yi, Angel X. Chang, Leonidas J. Guibas, and Hao Su. Sapien: A simulated part-based interactive environment. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11094–11104, 2020.
- [Zhou *et al.*, 2020] Yuxiao Zhou, Marc Habermann, Weipeng Xu, Ikhsanul Habibie, Christian Theobalt, and Feng Xu. Monocular real-time hand shape and motion capture using multi-modal data. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5345–5354, 2020.