Project Proposal: Project KOSMOS

(Kerbal Operational & Strategic Mission Orchestration System)

Advancing Autonomous Spacecraft Operations with a Verifiable, Multi-Agent AI Architecture

Student: Carson Johnson, Yakub Mungai, Ethan Cleary
Date: August 19, 2025

1. Executive Summary

Recent breakthroughs in AI agent frameworks and LLM-based planning have made sophisticated multi-agent systems achievable within undergraduate project timelines. Project KOSMOS leverages these advances to build a multi-agent spacecraft control system that automatically generates formal mission plans from natural language commands. The key innovation is the introduction of a specialized **Planning Engineer** agent, which converts high-level mission goals into PDDL (Planning Domain Definition Language), enabling optimal and verifiable spacecraft operations.

2. Project Objectives

The primary objective of this project is to develop a multi-agent AI system capable of autonomous spacecraft mission execution from high-level natural language commands. Our central goal is to demonstrate a greater improvement over the baseline performance of the single-agent LLMSat system on complex, multi-stage scenarios. Success will be measured by achieving a greater mission completion rate in these scenarios, compared to the baseline, alongside a functional and scalable integration of industry-standard frameworks supporting three or more specialized agents.

3. Technical Architecture

The architecture of Project KOSMOS is designed to be both cutting-edge and pragmatic, aligning with current industry best practices. For the core multi-agent framework, we have selected **CrewAI** for its rapid, role-based development, which benchmark studies have shown to offer significantly faster execution and setup times. Should the need for more granular, production-grade state management arise, a migration path to **LangGraph** is planned as a contingency.

Our system will feature a "crew" of specialist agents, including a **Mission Commander** for high-level orchestration, a **Navigation Specialist** to execute maneuvers, a **Systems Engineer** for telemetry monitoring, and the novel **Planning Engineer**. This Planning Engineer is responsible for translating mission goals into formal PDDL problems, which are then solved by a classical planner to ensure optimal and verifiable execution sequences.

To manage memory, we will adopt a pragmatic Memory-as-a-Service approach, utilizing a managed vector database like **Pinecone** or a local alternative such as **Chroma**. This system will store the mission history, learned PDDL patterns, and successful maneuver sequences, which the agents can query to inform their planning.

4. Implementation Strategy & Timeline

The project is structured as a focused 12-week plan, designed to deliver a production-ready system with comprehensive evaluation results.

The first month will be dedicated to establishing the foundation. In **weeks 1-2**, we will set up the KSP environment with the kRPC interface and install the CrewAI framework, validating the connection by replicating the successful LLMSat Scenario A. In **weeks 3-4**, we will build a multi-agent prototype with basic agent roles and communication patterns, culminating in a deliverable of a working system executing simple orbital maneuvers.

The second month focuses on agent intelligence. In **weeks 5-6**, we will develop the specialist agents, focusing on the Planning Engineer's PDDL generation and the Navigation Specialist's classical planner integration. In **weeks 7-8**, we will integrate the full LLM→PDDL→Planner→Execution pipeline and the memory system, delivering a crew of specialized agents with enhanced reasoning capabilities.

The final month is for integration and evaluation. **Weeks 9-10** will involve end-to-end mission workflow implementation and rigorous testing on complex scenarios. **Weeks 11-12** will be dedicated to a comparative analysis against the LLMSat baseline and the preparation of the final research paper and project documentation.

5. Risk Mitigation

We have identified and planned for several potential risks. The primary challenge is the complexity of multi-agent coordination, which will be mitigated by leveraging CrewAI's proven role-based patterns. The reliability of LLM-generated PDDL will be managed by starting with simple domains and incorporating robust validation and debugging loops in the Planning Engineer's workflow. As contingencies, a simplified single-agent architecture or template-based planning can be adopted if necessary.

6. Expected Outcomes

This project will produce several key deliverables. The primary technical deliverable is a functional, multi-agent system capable of autonomous spacecraft control, accompanied by a comprehensive performance evaluation demonstrating its superiority over single-agent baselines, and control of spacecraft using Natural Language.

Through this project, the team will gain significant experience in production AI frameworks, multi-agent system design, and space simulation integration, alongside valuable skills in experimental design, performance evaluation, and software engineering best practices.

7. Conclusion & Project Viability

Project KOSMOS represents a pragmatic convergence of cutting-edge research objectives with industry-proven implementation strategies. By leveraging the dramatic improvements in AI agent frameworks and the mature KSP integration ecosystem, this project achieves ambitious technical goals within realistic undergraduate constraints. The expected impact is a working demonstration that multi-agent AI systems can achieve superior performance in complex spacecraft operations, providing both immediate practical value and foundational research for next-generation space mission autonomy.

6. References

[1] Carrasco, Alejandro, Victor Rodriguez-Fernandez, Richard Linares. "Large Language Models as Autonomous Spacecraft Operators in Kerbal Space Program." arXiv:2505.19896, 2025.

[2] Carrasco, Alejandro, Victor Rodriguez-Fernandez, Richard Linares. "Fine-tuning LLMs for Autonomous Spacecraft Control: A Case Study Using Kerbal Space Program." arXiv:2408.08676, 2024.

[3] Rodriguez-Fernandez, Victor, Alejandro Carrasco, Jason Cheng, Eli Scharf, Peng Mun Siew, Richard Linares. "Language Models are Spacecraft Operators." arXiv:2404.00413, 2024.

[4] Maranto, David. "LLMSat: A Large Language Model-Based Goal-Oriented Agent for Autonomous Space Exploration." arXiv:2405.01392, 2024.

[5] *Kerbal Space Program*, www.kerbalspaceprogram.com/. Accessed 19 Aug. 2025.

[6] Mahdavi, Sadegh, Raquel Aoki, Keyi Tang, Yanshuai Cao. "Leveraging Environment Interaction for Automated PDDL Translation and Planning with Large Language Models." arXiv:2407.12979, 2024.

[7] Wang, Guanzhi, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, Anima Anandkumar. "Voyager: An Open-Ended Embodied Agent with Large Language Models." arXiv:2305.16291, 2023.