Razer
Case Study

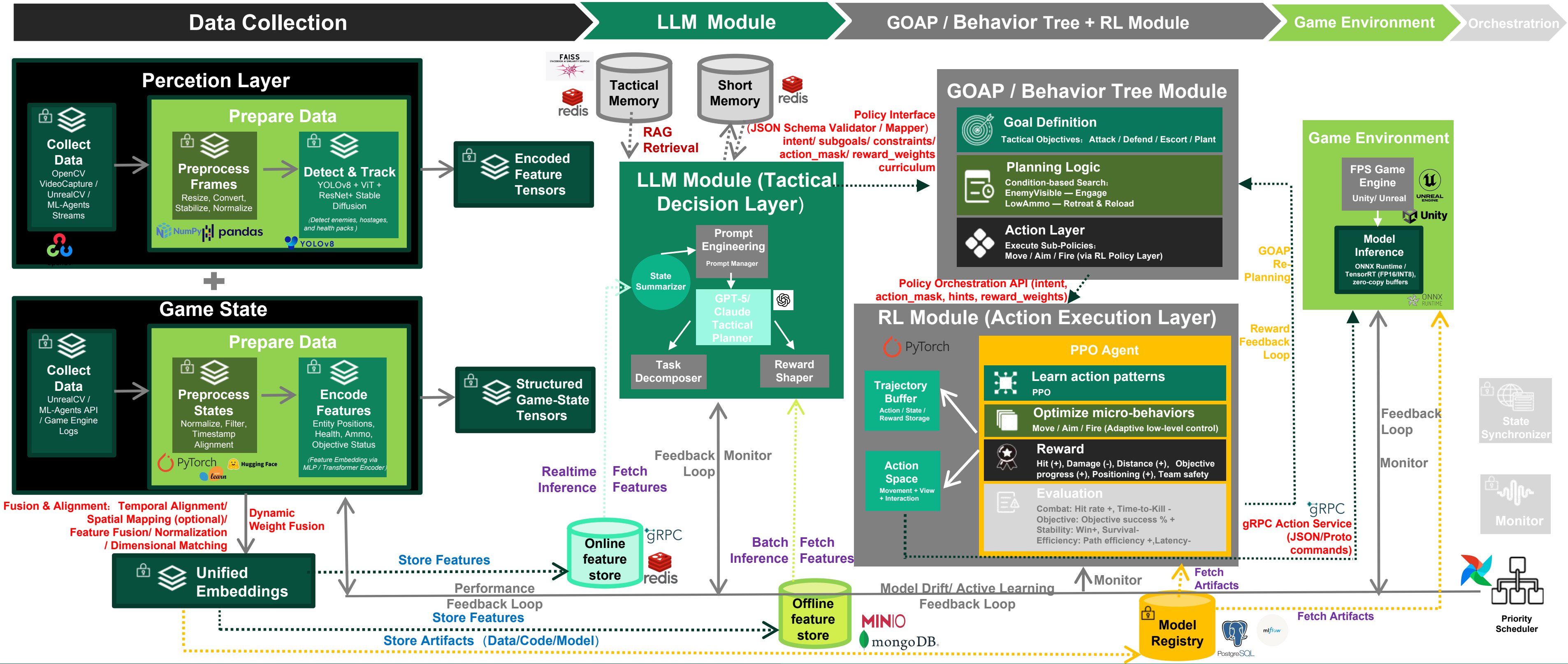
Hybrid LLM + RL Intelligent Agent Design for FPS Game Simulation

Hu Yu
Singapore Management University | MITB

Casestudy

Implementation of Gameplay agents using Reinforcement Learning(RL)

Machine Learning Lifecycle



Casestudy

Core Technical Contributions and System Optimizations

Key Innovations in Hybrid LLM + GOAP / Behavior Tree + RL Agent Design

- Combines LLM's global reasoning with RL's fine-grained control for human-like deliberation.Injects world knowledge from LLMs into the RL value function.
- Language Based Reward Shaping and Policy Alignment. Simulative Reasoning and Embodied Learning Loop.

RAG & Orchestration Design

- Retrieval Setup: Hybrid RAG using Cosine Similarity + BM25, followed by a reranker for relevance refinement. (Still decide if use RL based ranking model, +10-20ms latency + deployment cost)
- Retention Policy: short-term data stays in Redis (30–180 s), and long-term behaviors in vector or cold storage (weeks or longer) for analytics and personalization.

Decision Flow:

- The orchestration layer replaces LangGraph's workflow logic to reduce latency.
- It first checks whether the local knowledge base contains relevant information.
- If matched → use RAG results directly;
- if not → trigger external Gen-AI retrieval (LLM-based reasoning).
- Goal: Achieve lower latency, lightweight coordination, and more accurate tactical outputs.
- and enable quick recovery after interruptions.

Feature Fusion & Alignment

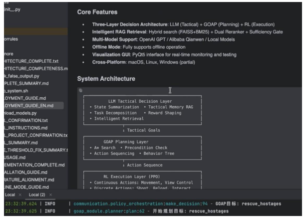
- Introduced Dynamic Weight Fusion (DWF) for adaptive balancing between visual and state embeddings.
- A lightweight gating network analyzes modality confidence in real time and adjusts fusion weights dynamically. Enhances robustness in noisy or incomplete environments and improves downstream decision stability.

Low Latency Design

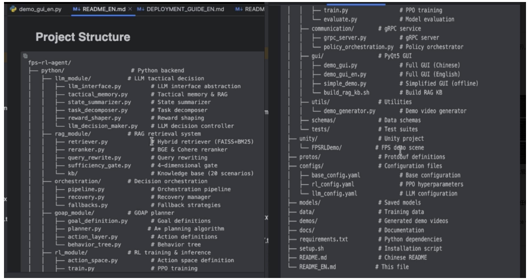
- OpenCV / FFmpeg C++ capture (zero-copy, GPU buffer)
- Use self orchestration instead of langchain/langGragh
- CUDA-based resize / normalize (no Python overhead)
- Shared Memory Bridge (C++ & Python) < 1 ms latency
- Unity/Unreal native plugin for action injection & reward callback
- TensorRT / ONNX Runtime (FP16/INT8) inference engine

Casestudy

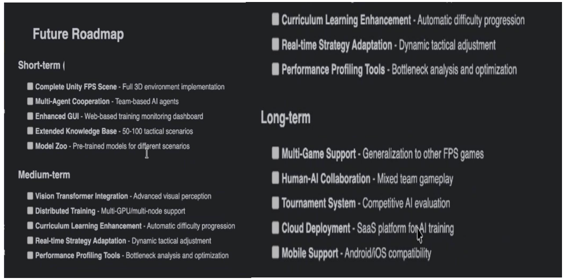
FPS RL Agent System Demo



System Architecture & Core Features

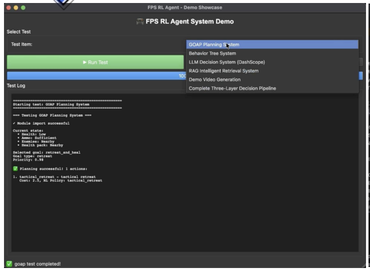


Project Structure & Modules

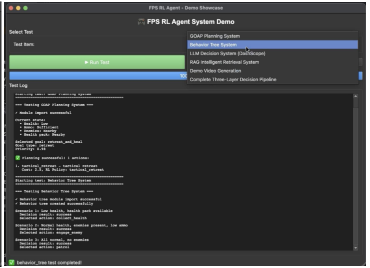


Future Road Map

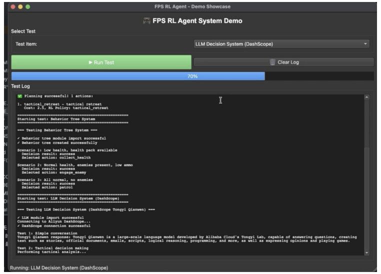
Integration Completed - All subsystems were tested independently and successfully integrated into a three-layer decision pipeline



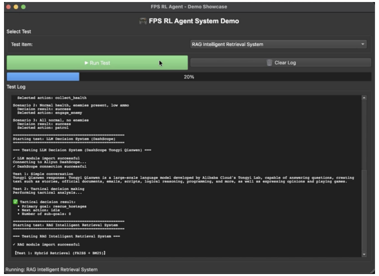
GOAP Planning System Testing



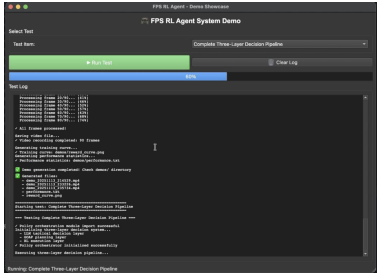
RL/ Behavior Tree System Testing



LLM Decision System Testing



RAG Retrival System Testing



Fully Integration

Next: RL policy training + Vision Transformer integration

Casestudy

Research Motivation for Next-Gen Enhancement

1.Generalist Game Agents & Multi-Task Adaptation
Recent works propose generalist architectures capable of learning across diverse game environments without retraining.Techniques such as shared world models, meta-RL, and task-conditioned policies enable agents to generalize tactical knowledge beyond a single FPS task.

2. Multi-Modal Fusion Beyond Static Integration
Dynamic Weight Fusion (DWF) + Cross-Modal Semantic Alignment

3. Temporal-Spatial Reasoning in Continuous Environments
Integrating temporal attention or trajectory transformers could improve memory of long-term enemy movement or strategy shifts. Adds temporal awareness for smoother, context-consistent agent behavior.

4. Simulation-to-Reasoning Feedback (LLM as Reflective Memory)
Methods such as unsupervised skill discovery (DIAYN) or HRL (Hierarchical RL) let the system autonomously define meaningful intermediate goals.
Could extend GOAP behavior trees with learned skill primitives rather than hand-coded nodes.

5. LLM System Resilience & Performance Optimisation
Circuit breaker: Automatically fall back to the rule based system when the LLM service times out or becomes unavailable.
Metrics: Add latency & success-rate probes.
Caching: Reuse LLM decisions for frequent game states.

References

1. Tencent AI Lab. Agents Play Thousands of 3D Video Games: Towards Generalist Game Agents, 2024.
2. Wu et al. Think in Games: Learning to Reason in Games via Reinforcement Learning with Large Language Models, 2024.
3. DeepMind. XLand: Generalist Reinforcement Learning Agents, 2021–2023.
4. OpenAI. Procgen Benchmark and PPO Baselines, 2020.
5. Eysenbach, B., et al. Diversity is All You Need (DIAYN): Unsupervised Skill Discovery in RL, NeurIPS, 2018.
6. Vezhnevets, A., et al. FeUdal Networks for Hierarchical Reinforcement Learning, ICML, 2017.
7. Finn, C., et al. Model-Agnostic Meta-Learning for Fast Adaptation (MAML), ICML, 2017.
8. Li, Y., et al. LLM-Guided Reward Shaping and Curriculum Learning, 2023.
9. Chen, Z., et al. Reflective Memory and Self-Critique Mechanisms in LLM-RL Agents, 2024.
10. Huang, J. & Gao, Y. Language Models as Planners: Integrating LLMs with RL for Decision Making, 2023.
11. OpenAI. AutoGPT, Voyager, and Gato Series: Autonomous Generalist Agents, 2022–2024.
12. Orkin, J. Goal-Oriented Action Planning (GOAP) for Game AI, GDC, 2003.
13. Isla, D. Handling Complexity in the Halo 2 AI, GDC, 2005.
14. Zhang, L. Hybrid GOAP-RL Architectures in Modern Games, IEEE CoG, 2023.
15. Liu, K., et al. Dynamic Weight Fusion Networks for Multimodal Adaptation, CVPR, 2023.
16. Radford, A., et al. CLIP: Connecting Vision and Language, NeurIPS, 2021.
17. Wang, H., et al. BEVT: BERT Pretraining of Video Transformers, ICCV, 2023.
18. Shazeer, N., et al. Mixture-of-Experts Architectures for Efficient Multimodal Fusion, Google AI, 2023.
19. Mehta, R. Lightweight Multimodal Architectures for Edge Deployment, 2023.
20. NVIDIA. ONNX Runtime and TensorRT Acceleration Guide, Technical Report, 2024.
21. NVIDIA. INT8 / FP8 / INT4 Quantization for Edge AI, White Paper, 2024.
22. DeepMind. Asynchronous PPO and Distributed RL Pipelines, 2022.
23. Microsoft Research. CUDA-Based Preprocessing and Shared Memory Optimization, 2023.
24. Radford, A., et al. CLIP: Learning Transferable Visual Representations from Natural Language Supervision, NeurIPS, 2021.
25. DeepMind. Gato: A Generalist Agent, Nature, 2022.
26. Parisotto, E., et al. Transformer-based PPO and Memory-Augmented Reinforcement Learning, 2023.
27. Schaeffer, J., et al. Game AI Taxonomy: From Rules to Learning Agents, ACM Comp. Games Survey, 2020.