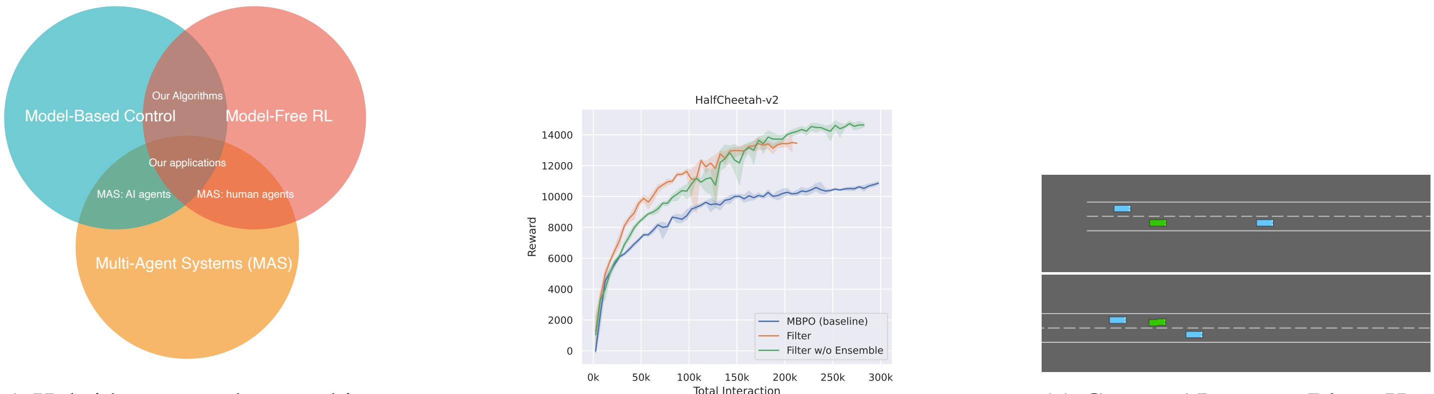
RESEARCH STATEMENT

In my doctoral research at University of Illinois Chicago, I have applied *model-based Reinforcement Learning* (RL), *switched control theory*, *stochastic game theory*, and *composite optimization* to develop algorithms that leverage model information to accelerate model-free training (Fig. 1a), achieving greater sample efficiency. The algorithms designed require fewer interactions with the environment to achieve the same level of optimality (Fig. 1b). Additionally, we extended the algorithm to hybrid multi-agent systems involving human participants, where humans are treated as black-box agents (model-free) while other agents follow a model-based approach, enabling effective collaboration between the AI and human agents without the need to control or model human behavior (Fig. 1c).



(a) Hybrid systems that combine model-based control and model-free reinforcement learning

(b) Filter and Filter w/o Ensemble are our algorithms

(c) Green: AI agents; Blue: Human agents. The AI agent successfully changed the lane.

Figure 1: Summary of Research

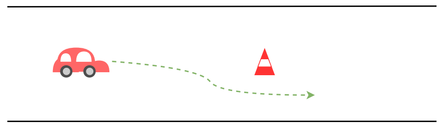
My current research extends these hybrid principles to the frontier of Large Language Models (LLMs) in decision-making systems. In a groundbreaking collaboration with Stony Brook University, we created the first system to integrate LLMs into real-world, closed-loop autonomous driving (Dong et al., 2024). Our architecture resolves the fundamental latency challenges that previously prevented LLMs from being used in real-time systems by strategically positioning them as high-level instructors while maintaining rapid local control mechanisms (Fig. 2).

Looking forward, I am expanding these methodologies to empower a broader range of end-users in deploying AI systems. This work draws on my entrepreneurial experience as a co-founder of LIII NETWORK, where we developed GNU $\TeX_{\text{MACS}}$/Mogan—a WYSIWYG $\TeX$-like editor that accelerates academic writing—and Goldfish Scheme, an interpreter that bridges the gap between Scheme and Python ecosystems. Particularly notable is our implementation of literate programming (Knuth, 1984) in Goldfish Scheme, which exemplifies my commitment to making complex systems more comprehensible and maintainable.

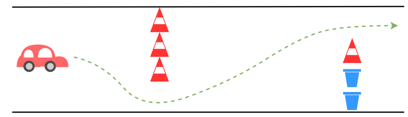
These varied projects share a common thread: making sophisticated AI-involved systems more practical and accessible while maintaining their theoretical rigor. As AI continues to evolve, I am dedicated to developing hybrid approaches that combine the best aspects of different methodologies, ultimately creating AI systems that can be effectively deployed in real-world applications ranging from autonomous vehicles to healthcare and scientific research.

1 COMBINING MODEL-BASED CONTROL AND MODEL-FREE LEARNING

The challenge of controlling nonlinear systems lies at the heart of many real-world applications. Traditional approaches face a critical trade-off: model-based methods are efficient but potentially inaccurate, while model-free methods are accurate but data-hungry. This dilemma is particularly acute in physical systems, where extensive data collection can be costly or impractical.



(a) Training environment for end-to-end autonomous driving



(b) Testing environment for LLMs-assisted end-to-end autonomous driving

Figure 2: Training and testing environments (the experiment is conducted in real-world, closed-loop; this image is for presentation purposes only). The LLMs enhance the generalization of autonomous driving systems.

Our solution bridges this gap through hybrid algorithms that combine the strengths of both approaches. We developed methods based on composite optimization (**Yansong Li** and Han, 2022) and Dyna-style learning (**Yansong Li** et al., 2024) that achieve optimal performance with significantly less data than pure model-free approaches. These algorithms strategically integrate approximation models with model-free methods to accelerate learning while maintaining accuracy.

Continuous Control on Legged Robots Our theoretical advances found practical application in the challenging domain of legged robotics. We developed a novel *Dyna-style learning with data filter* algorithm (**Yansong Li** et al., 2024) that addresses a fundamental challenge in model-based learning: the negative impact of biased data. By introducing an intelligent filtering mechanism, our algorithm selectively uses only the most reliable model-generated data to accelerate model-free training. We validated this approach in MuJoCo, where it achieved state-of-the-art performance while requiring significantly fewer environmental interactions (Fig. 1b).

Potential Impacts The success of our hybrid algorithms opens new possibilities for deploying AI in complex, real-world environments. Their reduced need for environmental interactions directly translates to safer and more cost-effective deployment—a crucial advantage when training physical robots. Moreover, the model-based components provide clear insights into AI decision-making, enabling easier debugging and modification. This interpretability is essential for applications where understanding and adjusting AI behavior is critical for safety and performance.

2 HYBRID SYSTEMS IN DECISION MAKING

My research deployed the hybrid principles beyond the combination of models in model-free RL. The exploration has led to solutions in human-AI collaboration and language model integration for autonomous systems.

Multi-Agent Systems Involving Humans Human behavior presents a unique challenge in multi-agent systems due to its complexity and unpredictability. Rather than attempting to create precise models of human behavior, we developed a novel approach that treats humans as black-box components within the system. This perspective allows us to observe and adapt to human actions without making assumptions about their decision-making process. Meanwhile, we maintain explicit models for AI agents, creating an effective hybrid architecture. This approach led to two significant advances. First, we developed an algorithm for Stackelberg games (**Yansong Li** and Han, 2023) that enables AI agents to collaborate effectively with human followers without modeling human behavior. We then extended this framework to handle the more complex domain of stochastic games (**Yansong Li** and Han, 2024), demonstrating the adaptability of our hybrid approach.

Large Language Models for Decision Making We applied a hybrid system to integrate LLMs into autonomous systems. Our key insight is that while LLMs excel at high-level reasoning, they are poorly suited for direct control tasks due to latency issues. This led us to develop a novel architecture for autonomous driving that leverages LLMs' strengths while avoiding their limitations. Our system (Fig. 2) represents a breakthrough in LLM application for autonomous driving. We created the first successful implementation of LLMs in closed-loop, real-world driving by decoupling high-level decision-making from direct control. The architecture combines multimodal LLMs for strategic decisions with rapid local control systems, effectively addressing the latency challenges that hindered previous attempts. Our experiments demonstrate that this hybrid approach significantly enhances the vehicle's ability to navigate novel environments (Dong et al., 2024).

3 LARGE LANGUAGE MODELS FOR LITERATE PROGRAMMING

Modernizing Scientific Document Creation: GNU TeXmacs and Mogan The challenge of efficient mathematical writing has long been a bottleneck in scientific communication. GNU TeXmacs pioneered a solution with its WYSIWYG TeX-like editor, enabling users to type equations at handwriting speed. Building on this foundation, we identified crucial areas for improvement in performance, framework modernization, and user experience.

These insights led us to develop Mogan Research, a fork of GNU TeXmacs. Our version maintains the core strength of rapid equation editing while implementing significant technical upgrades. Table 1 details these improvements, which span performance optimization, modernized framework adoption, and enhanced user interface design.

	GNU TeXmacs	**Mogan Research**
Performance	slow	fast
Qt Framework	mainly in Qt 4 (some in Qt 5)	mainly in Qt 6 (some in Qt 5)
Scheme Engine	GNU Guile 1.8.x (deprecated by GNU)	Goldfish Scheme (based on S7 Scheme)
Build Tool	GNU Autotools	xmake

Table 1: Comparison between GNU TeXmacsand Mogan Research

Literate Programming in the Era of LLMs: Goldfish Scheme as a Case-Study The evolution of Mogan required us to address a fundamental challenge in its architecture: the Scheme interpreter. While TeXmacs relied on the deprecated GNU Guile 1.8.x, our initial choice of S7 Scheme also proved limiting due to incomplete R7RS and Unicode support. This prompted us to develop *Goldfish Scheme*[1] (Shen et al., 2024).

What sets Goldfish Scheme apart is its innovative development approach using literate programming, a methodology introduced by Knuth (1984). Building on Knuth's original vision (Knuth, 1984), we see literate programming not merely as a documentation method but as a fundamental shift in how code is organized and understood. Rather than following traditional top-down or bottom-up approaches, code becomes a "web" of interconnected ideas[2] that naturally aligns with how LLMs process and understand information. This alignment between literate programming and LLM capabilities has guided our development of Goldfish Scheme. We designed the language with two key principles in mind:

- *Simplicity for Inference*: We leveraged Scheme's minimalist design to create code that LLMs can more easily analyze and understand, leading to more accurate and reliable responses.
- *Python Compatibility*: By incorporating Python-like features, we bridge the gap between Scheme's elegant simplicity and the extensive training data available to LLMs through Python's widespread use.

The emergence of Large Language Models presents a unique opportunity to revitalize literate programming for modern software development. Through our research, we raised a compelling hypothesis about the synergy between these two approaches:

> *Projects developed using literate programming are inherently more understandable for LLMs than projects containing even the most detailed comments.*

[1]Goldfish Scheme is named with a playful reference to the idea that goldfish have a 7-second (7s) memory, the inverse of S7. In the future, we hope LLMs can help create new functions for Goldfish Scheme within those same 7s. The literate programming document for Goldfish is available in `https://github.com/LiiiLabs/goldfish/releases/tag/v17.10.9`.

[2]In our implementation, code exists in interconnected chunks that tell a story about their purpose and relationships. We provide this functionality as a Mogan plugin, automatically exporting chunks into organized project files and folders.

Our experiments in Shen et al. (2024) provide strong evidence for this hypothesis. The narrative structure of literate programming, where code and documentation form a cohesive story, aligns naturally with LLMs' ability to process and understand natural language. This alignment suggests a promising future where literate programming could become the preferred methodology for creating AI-comprehensible software.

FUTURE RESEARCH DIRECTIONS

How can we deploy learning-based decision-making policies in real-world real-time? The deployment of learning-based decision-making systems in real-world applications presents three fundamental challenges that drive my future research agenda:

- *Data requirements*: Real-world training demands extensive environmental interaction, making traditional learning approaches impractical for many applications.
- *Real-time performance*: The computational complexity of learning-based policies often conflicts with the strict timing constraints of real-world decision-making.
- *Adaptability to environmental changes*: Current learning systems struggle to generalize beyond their training data, limiting their effectiveness in dynamic environments.

My research has already made significant strides in addressing these challenges through hybrid approaches. Our algorithms combining model-based control with model-free learning have successfully reduced data requirements, while our LLM-assisted driving framework has demonstrated improvements in real-time performance and environmental adaptability.

Looking ahead, I plan to expand this research to address two critical remaining challenges: the *exploration vs. exploitation balance* and *safety and stability* in real-world deployments. Through collaborations with experts in robotics, control theory, and game theory, I aim to develop more robust AI systems that can safely and effectively adapt to real-world applications. This work will build upon my existing research while opening new avenues for investigation in autonomous systems, human-AI interaction, and safe AI deployment.

Can LLMs help for decision-making in real-world? Integrating Large Language Models into real-world decision-making systems requires careful consideration of their fundamental limitations. Our research indicates that LLMs are better suited for supporting decision-making processes than making direct decisions. This conclusion stems from three critical constraints:

- *High Latency:* LLM inference creates response delays incompatible with real-time decision requirements.
- *Limited Context Awareness:* LLMs fail to understand dynamic situations, leading to disastrous decisions.
- *Reliability Concerns:* The sensitivity of LLM outputs to minor input variations introduces unpredictability that makes them unsuitable for critical decision-making roles.

These insights have led us to pursue hybrid system architectures where LLMs serve as strategic advisors rather than direct decision-makers. This approach parallels my earlier work on combining model-based and model-free methods, suggesting promising directions for integration. We can create more robust and reliable decision-making frameworks by leveraging LLMs' strengths in high-level reasoning while compensating for their limitations through specialized control systems. My future research will explore these hybrid architectures through collaborations that span machine learning, control theory, and systems design.

Renaissance of Literate Programming in the Era of LLMs The use of literate programming extends beyond traditional programming paradigms into the realm of prompt engineering, offering new ways to structure and communicate code. While challenges remain in areas such as collaboration workflows, version control systems, and standardization of writing practices, we are actively working with our colleagues at LIII NETWORK to address these issues. Our goal is to create a modern literate programming framework that fully harnesses the capabilities of LLMs while maintaining the practical needs of software development.

REFERENCES

Dong, Z., Zhu, Y., **Yansong Li**, Mahon, K., and Sun, Y. (2024). Generalizing end-to-end autonomous driving in real-world environments using zero-shot LLMs. In *8th Annual Conference on Robot Learning*.

Knuth, D. E. (1984). Literate programming. *Comput. J.*, 27(2):97–111.

Shen, D., Liu, N., **Yansong Li**, Wang, D., He, L., and Zhou, Y. (2024). *Goldfish Scheme: A Scheme Interpreter with Python-Like Standard Library*. LIII NETWORK.

Yansong Li, Dong, Z., Luo, E., Wu, Y., Wu, S., and Han, S. (2024). When to trust your data: Enhancing dyna-style model-based reinforcement learning with data filter.

Yansong Li and Han, S. (2022). Accelerating model-free policy optimization using model-based gradient: A composite optimization perspective. In Firoozi, R., Mehr, N., Yel, E., Antonova, R., Bohg, J., Schwager, M., and Kochenderfer, M. J., editors, *Learning for Dynamics and Control Conference, L4DC 2022, 23-24 June 2022, Stanford University, Stanford, CA, USA*, volume 168 of *Proceedings of Machine Learning Research*, pages 304–315. PMLR.

Yansong Li and Han, S. (2023). Solving strongly convex and smooth stackelberg games without modeling the follower. In *American Control Conference, ACC 2023, San Diego, CA, USA, May 31 - June 2, 2023*, pages 2332–2337. IEEE.

Yansong Li and Han, S. (2024). Efficient collaboration with unknown agents: Ignoring similar agents without checking similarity. In Dastani, M., Sichman, J. S., Alechina, N., and Dignum, V., editors, *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2024, Auckland, New Zealand, May 6-10, 2024*, pages 2363–2365. International Foundation for Autonomous Agents and Multiagent Systems / ACM.