Open-World Robot Planning and Learning

Panpan Cai
Full-time Faculty, Shanghai Innovation Institute
Associate Professor, School of Artificial Intelligence, Shanghai Jiao Tong University

My research focuses on *large-scale decision making* for robots that must operate in *complex, uncertain, and dynamic* real-world environments, so that they can interact seamlessly with humans and accomplish challenging long-horizon tasks. Decision making in such settings is fundamentally hard: robots have noisy sensors and limited models; the world is full of human-controlled agents with subtle and shifting intentions; and useful tasks span very long-horizon plans whose state-action space is astronomical. Decisive progress requires combining the rigor of *model-based planning* with the breadth of *learning from large-scale data*.

Two application domains drive and stress-test the algorithms my group develops: **(i) autonomous driving in dense urban traffic**, where the robot must reason about many heterogeneous agents in real time, and **(ii) home-service mobile manipulation**, where the robot must interpret open-ended human instructions, reason over partially observed kitchens and apartments, and execute long-horizon plans built from contact-rich primitives. Concretely, my research advances three intertwined directions: *modeling and reasoning under uncertainty*, *real-time planning at scale*, and *integrating planning with learning and foundation models*.

Modeling and reasoning under uncertainty

A useful model not only captures the complexity of real-world dynamics and human behaviours, but also represents intrinsic uncertainties in principled ways. For *autonomous driving*, I have developed traffic behavioural models that formalize interactions among heterogeneous agents as constrained optimization in the velocity space (PORCA [1], GAMMA [2]), and an open-source simulator, SUMMIT [3], that produces realistic driving simulations on real-world maps for end-to-end algorithm evaluation. For *home-service robots*, the dominant uncertainties shift from continuous motion to symbolic and semantic ambiguity — ambiguous human instructions, hidden or unknown object locations, and open-vocabulary object types. Tru-POMDP [4] tackles these challenges by combining a hierarchical *Tree of Hypotheses*, queried from large language models, with an *open-ended POMDP* formulation that supports rigorous Bayesian belief tracking and efficient belief-space planning. On complex object-rearrangement tasks across diverse kitchens, Tru-POMDP significantly outperforms LLM-based and LLM–tree-search hybrid planners, with stronger robustness to ambiguity and occlusion.

Real-time planning at scale

Sophisticated planning brings combinatorial complexity — the well-known “curse of dimensionality” and “curse of history”. I attack this from a *systems-and-algorithms co-design* angle. HyP-DESPOT [5, 6] introduces hybrid CPU/GPU parallel belief-tree search that speeds up large-scale POMDP planning by hundreds of times and is widely used as an open-source planner. Building on this foundation, my recent work pushes the *practicalization* of POMDP planning on both target domains.

Generic urban driving. Hi-Drive [7] formulates planning under behavioural and trajectory uncertainty as a hierarchical POMDP, treating driver models as high-level decision actions to manage exponential complexity, and refines trajectories via importance-sampling optimization over critical agents. On real-world urban benchmarks, Hi-Drive significantly outperforms state-of-the-art planning- and learning-based baselines across diverse driving situations. Vec-QMDP [8] further aligns POMDP search with modern CPUs' SIMD architecture via Data-Oriented Design and a hierarchical parallelism scheme (sub-trees across cores, vectorized expansion and collision checking within SIMD lanes), achieving 227x–1073x speedup over state-of-the-art serial planners and millisecond-level latency on large-scale benchmarks such as nuPlan. Together, Hi-Drive and VecQMDP move principled POMDP planning from constrained crowd-driving demos into the general urban driving stack, deployable on commodity CPUs without GPUs.

Motion planning for mobile-manipulation. Neural Randomized Planning (NRP) [9] addresses long-range whole-body motion planning in cluttered household environments by combining global sampling-based motion planning with a local neural sampler. NRP exploits the search structure of the global planner to stitch together learned local sampling distributions into an adaptive global one, scaling to high-dimensional configuration spaces while preserving classical planning guarantees. Despite being trained only in simulation, NRP transfers zero-shot to a real robot operating in novel households.

Integrating planning with learning and foundation models

Robots should improve from experience. Yet planning alone becomes intractable when problems are large-scale, while learning alone is brittle when domains shift. I have advocated a paradigm of *"think locally, learn globally"* — planning supplies online, problem-aware optimization, and learning amortizes global priors and heuristics.

Driving. LeTS-Drive [10] learns global priors offline from tree search and uses them as heuristics to guide online belief-tree search, enabling sophisticated decision making in large, highly interactive crowds. LEADER [11] extends this idea by learning attention over driving behaviours for planning under uncertainty (CoRL 2022 Best Paper Finalist).

Home-service mobile manipulation. The same paradigm has matured in the era of foundation models. UniDomain [12] pretrains a unified PDDL domain (3,137 operators, 2,875 predicates, 16,481 causal edges) from 12,393 manipulation videos; given a target task class, it retrieves and systematically fuses task-relevant atomics into compositional meta-domains. UniPlan [13] extends UniDomain to long-horizon mobile manipulation in large indoor environments by unifying scene topology, visuals, and robot capabilities into a holistic PDDL representation, grounded online by a vision-language model on a visual-topological scene map. AHAT [14] couples an LLM trained to map abstract human instructions plus textual scene graphs into PDDL subgoals with a new reinforcement-learning algorithm (TGPO) that injects external correction of intermediate reasoning into GRPO, yielding scalable long-horizon plans for human-style household tasks across very large environments.

Foundation models for physical intelligence. At the policy level, my group studies how foundation models can be made physically capable. ForceVLA [15] treats external force sensing as a first-class modality in Vision-Language-Action models via a force-aware Mixture-of-Experts, improving contact-rich tasks (e.g. plug insertion) by 23.2% over strong π_0-based baselines. MINT [16] argues that imitation learning should "Mimic Intent, Not just Trajectories": it disentangles behaviour intent from execution details via multi-scale frequency-space tokenization, and uses next-scale autoregression to enable one-shot skill transfer. I-Perceive [17] is a vision-language foundation model for active perception in mobile manipulation; it predicts task-relevant camera views from open-ended language instructions by fusing a VLM backbone with a geometric foundation model, generalizing zero-shot to novel scenes and tasks.

Summary and outlook

I enjoy making principled approaches *work* in the complex real world. The methodology my group has built — belief-space planning, behaviour and intent modelling, massive parallelization, and tight coupling with learning — was forged on autonomous driving and now also powers home-service mobile manipulation. The next phase of my research is to push the *practicalization* of large-scale decision making one step further: open-world planners that handle truly unbounded objects, instructions, and scene types; foundation models that reason and plan in concert with classical solvers; and scalable systems that bring the rigor of model-based planning to settings where humans live and work alongside robots.

References

[1] Y. Luo, P. Cai, A. Bera, D. Hsu, W.S. Lee, D. Manocha. PORCA: Modeling and planning for autonomous driving among many pedestrians. *IEEE RA-L*, 2018.

[2] Y. Luo*, P. Cai* (corresponding author), Y. Lee, D. Hsu. GAMMA: A General Agent Motion Model for Autonomous Driving. *IEEE RA-L*, 2022.

[3] P. Cai*, Y. Lee*, Y. Luo, D. Hsu. SUMMIT: A Simulator for Urban Driving in Massive Mixed Traffic. *IEEE ICRA*, 2020.

[4] W. Tang, X. He, Y. Huang, Y. Xiao, C. Lu, P. Cai (corresponding author). Tru-POMDP: Task Planning Under Uncertainty via Tree of Hypotheses and Open-Ended POMDPs. *NeurIPS*, 2025.

[5] P. Cai, Y. Luo, D. Hsu, W.S. Lee. HyP-DESPOT: A Hybrid Parallel Algorithm for Online Planning under Uncertainty. *RSS*, 2018.

[6] P. Cai, Y. Luo, D. Hsu, W.S. Lee. HyP-DESPOT: A Hybrid Parallel Algorithm for Online Planning under Uncertainty. *IJRR*, 2021.

[7] X. Jin, C. Zeng, S. Zhu, C. Liu, P. Cai (corresponding author). Hi-Drive: Hierarchical POMDP Planning for Safe Autonomous Driving in Diverse Urban Environments. *IEEE RA-L*, 2025.

[8] X. Jin, Y. Dong, B. Sun, H. Xu, Z. Hao, X. Lang, P. Cai (corresponding author). Vec-QMDP: Vectorized POMDP Planning on CPUs for Real-Time Autonomous Driving. *RSS*, 2026.

[9] Y. Lu, Y. Ma, D. Hsu, P. Cai (corresponding author). Neural Randomized Planning for Whole Body Robot Motion. *arXiv preprint*, 2024.

[10] P. Cai, Y. Luo, A. Saxena, D. Hsu, W.S. Lee. LeTS-Drive: Driving in a Crowd by Learning from Tree Search. *RSS*, 2019.

[11] M.H. Danesh*, P. Cai* (corresponding author), D. Hsu. LEADER: Learning Attention over Driving Behaviors for Planning under Uncertainty. *CoRL (Best Paper Finalist)*, 2022.

[12] H. Ye, Y. Xiao, C. Lu, P. Cai (corresponding author). UniDomain: Pretraining a Unified PDDL Domain from Real-World Demonstrations for Generalizable Robot Task Planning. *NeurIPS*, 2025.

[13] H. Ye, Y. Xiao, C. Lu, P. Cai (corresponding author). UniPlan: Vision-Language Task Planning for Mobile Manipulation with Unified PDDL Formulation. *arXiv preprint*, 2026.

[14] Z. Liu, Y. Li, R. Huang, C. Lu, P. Cai (corresponding author). Any House Any Task: Scalable Long-Horizon Planning for Abstract Human Tasks. *arXiv preprint*, 2026.

[15] J. Yu*, H. Liu*, Q. Yu*, J. Ren, C. Hao, H. Ding, G. Huang, G. Huang, Y. Song, P. Cai, W. Zhang, C. Lu. ForceVLA: Enhancing VLA Models with a Force-aware MoE for Contact-rich Manipulation. *NeurIPS*, 2025.

[16] R. Huang, C. Zeng, W. Tang, J. Cai, C. Lu, P. Cai (corresponding author). Mimic Intent, Not Just Trajectories. *RSS*, 2026.

[17] Y. Huang, Z. Wang, W. Tang, C. Lu, P. Cai (corresponding author). I-Perceive: A Foundation Model for Vision-Language Active Perception. *arXiv preprint*, 2026.