Project Milestone: Decision Transformer Warm-Start for Obstacle-Aware Minimum-Time Racing Trajectory Optimization

Victor Sebastian Martinez Perez [1]

Abstract

This project asks whether an offline sequence model, a Decision Transformer (DT), can amortize nonlinear constrained trajectory optimization for autonomous racing by providing high-quality warm starts. Obstacle-aware minimum-time lap optimization is formulated as a spatial direct-collocation nonlinear program (NLP) solved by IPOPT over an 8-state single-track vehicle model with nonlinear tire forces, parameterized to the Stanford Dynamic Design Lab *VW Golf GTI* test vehicle. In an initial batch evaluation on a 260 m oval track with randomized circular obstacles, only 20% of scenarios are accepted on the first IPOPT attempt; an acceptance-gated retry policy restores 100% acceptance at higher computational cost. The milestone deliverable is a reliable expert solver and evaluation harness, plus a dataset/feature contract and DT input/output specification. The final goal is to reduce IPOPT iterations and wall-clock time (and eliminate most retries) while preserving feasibility and near-baseline lap time.

1. Introduction

Minimum-time racing line optimization is a canonical "drive-at-the-limits" problem: the planner must coordinate steering and longitudinal force near the tire friction boundary while respecting hard geometric constraints (track edges and safety margins). In this regime, model choice and initialization matter—small errors in friction modeling, load transfer, or actuator limits can translate into qualitatively different local minima and very different feasibility outcomes (Kapania et al., 2015; Subosits & Gerdes, 2021; Aggarwal & Gerdes, 2025).

This project focuses on *closed-track* lap optimization with static obstacles. The vehicle is modeled as a nonlinear single-track (bicycle) system with load-transfer states and nonlinear tire forces, parameterized to the full-size *2018/2019 VW Golf GTI* used in prior Stanford Dynamic Design Lab experiments (Subosits & Gerdes, 2021; Aggarwal & Gerdes, 2025). Track geometry (centerline, width, curvature, and banking/grade when present) is loaded from precomputed map files; experiments in this milestone use the "Medium Oval" track (approximately 260 m lap length) with randomly placed circular obstacles and an inflated vehicle footprint.

Repeatedly solving this nonlinear constrained optimal control problem is expensive. Even with modern interior-point methods, performance is highly sensitive to warm starts: poor initializations can increase iterations, stall, or converge quickly to a trajectory that satisfies discretized constraints but violates dense collision checks. The current expert pipeline therefore includes an *acceptance-gated retry* mechanism (e.g., increasing discretization or obstacle sampling) to reliably produce safe demonstrations, but this increases compute and complicates dataset generation.

Decision Transformers (Chen et al., 2021) treat offline RL as conditional sequence modeling, conditioning on return-to-go (RTG) to generate actions consistent with high-return expert demonstrations. In this setting, the DT is not required to be *the controller*: it only needs to propose a dynamically coherent, near-feasible initial trajectory for a downstream optimizer. This hybrid view mirrors transformer warm-starting for spacecraft trajectory optimization (Guffanti & D'Amico, 2024), but racing adds harder contact-like nonlinearities (tire saturation) and strict geometric safety constraints. The hypothesis is that a DT trained on accepted expert solutions can produce initializations that reduce IPOPT iterations and eliminate most retries while preserving final lap time and constraint satisfaction.

2. Approach

2.1. Vehicle model and minimum-time NLP

The dynamics model follows the single-track formulations used in recent autonomous racing studies (Subosits &
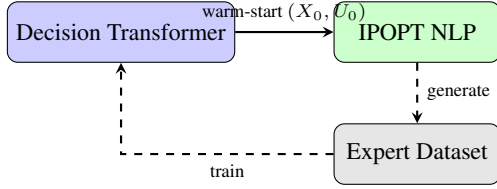
[1]Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, USA. Correspondence to: Victor Sebastian Martinez Perez <sebasmp@stanford.edu>.

Figure 1. Hybrid warm-starting pipeline. DT proposes a trajectory initialization; IPOPT enforces hard constraints and refines to a locally optimal solution.

Gerdes, 2021; Aggarwal & Gerdes, 2025): left/right tires are lumped per axle, planar rigid-body dynamics are retained, and nonlinear tire forces are represented with a Fiala-style brush model. The trajectory is parameterized by arc length s to align optimization constraints with track geometry. Let $k = 0, \ldots, N$ index spatial nodes with spacing Δs, and let decision variables be $\mathbf{X} = [\mathbf{x}_0, \ldots, \mathbf{x}_N]$ and $\mathbf{U} = [\mathbf{u}_0, \ldots, \mathbf{u}_N]$, with

$$\mathbf{x} = [u_x, u_y, r, \Delta F_{z,\text{long}}, \Delta F_{z,\text{lat}}, t, e, \Delta \psi]^\top, \quad \mathbf{u} = [\delta, F_x]^\top. \tag{1}$$

This state choice extends the standard $(u_x, u_y, r, t, e, \Delta\psi, \Delta F_z)$ representation by retaining both longitudinal and lateral load-transfer states used by the current simulator/optimizer implementation. Kinematic path relations follow the same structure as prior work:

$$\dot{s} = \frac{u_x \cos \Delta\psi - u_y \sin \Delta\psi}{1 - \kappa(s)e}, \qquad \dot{e} = u_x \sin \Delta\psi + u_y \cos \Delta\psi, \tag{2}$$

$$\dot{\Delta\psi} = r - \kappa(s)\dot{s}.$$

The NLP objective is minimum lap time with a lightweight control-smoothness regularizer:

$$\min_{\mathbf{X},\mathbf{U}} \ J = t_N + \lambda_u \sum_{k=0}^{N-1} \|\mathbf{u}_{k+1} - \mathbf{u}_k\|_2^2. \tag{3}$$

Dynamic continuity is enforced by trapezoidal collocation in spatial form, consistent with racing trajectory-optimization practice:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\Delta s}{2} \left(f_s(\mathbf{x}_k, \mathbf{u}_k) + f_s(\mathbf{x}_{k+1}, \mathbf{u}_{k+1}) \right), \tag{4}$$

where $f_s(\cdot)$ denotes the spatial-form dynamics $\dot{\mathbf{x}}/\dot{s}$, and the along-track rate $\dot{s}$ appears implicitly through the kinematics. The NLP includes hard track bounds, actuator limits, and two numerically critical constraints:

$$\dot{s}_k \geq \varepsilon_s \quad \text{(forward progress)}, \tag{5}$$

$$1 - \kappa(s_k)e_k \geq \varepsilon_\kappa \quad \text{(Frenet non-singularity)}. \tag{6}$$

For full-lap solutions, periodic closure is imposed as $\mathbf{x}_0 = \mathbf{x}_N$ and $\mathbf{u}_0 = \mathbf{u}_N$ for all states/controls except time (with $t_0 = 0$ and t_N free). The resulting sparse NLP is solved with IPOPT through CasADi, matching the solver stack commonly used in the cited vehicle-planning literature.

2.2. Obstacle handling and acceptance-gated retries

Static circular obstacles are enforced via node-based clearance constraints in the world frame:

$$\|p(s_k, e_k) - p_{\text{obs},i}\|_2^2 \geq (r_i + m_i + R_{\text{veh}} + c)^2, \tag{7}$$

where $p(s, e)$ is the Frenet-to-ENU map, r_i is obstacle radius, m_i is a safety margin, R_{veh} is an inflated vehicle footprint radius, and c is an optional additional clearance. To avoid over-complicating the NLP early, obstacle constraints are kept simple in the optimizer, while safety is enforced with a dense post-solve collision check (sub-sampled along each segment).

An *acceptance-gated retry policy* is then applied: if IPOPT converges but violates dense clearance by more than a small tolerance, the same scenario is retried with higher discretization (N) and/or denser obstacle sampling. This produces reliable accepted expert demonstrations while also surfacing failure modes that DT warm-starting should reduce (e.g., converging quickly to a trajectory that "cuts" an obstacle).

2.3. Dataset contract (expert trajectories)

Each accepted solve is stored as an episode with: (i) a reproducibility header (map identifier/hash, N, Δs, obstacle list, solver settings, seed), (ii) per-step arrays $(s_k, \mathbf{x}_k, \mathbf{u}_k, \Delta t_k)$, and (iii) derived features: global pose (E_k, N_k, ψ_k), track curvature $\kappa(s_k)$, and RTG defined as $\text{RTG}_k = -\sum_{j=k}^{N-1} \Delta t_j$. Storing the full internal state (including load-transfer states) keeps the data future-proof even if DT only observes a subset.

2.4. Decision Transformer warm-start (planned)

DT will be trained as a causal decoder that predicts controls conditioned on RTG and a compact observation:

$$\mathbf{z}_k = [\text{RTG}_k, \ \mathbf{o}_k, \ \mathbf{u}_{k-1}], \tag{8}$$

where $\mathbf{o}_k$ includes $(u_x, u_y, r, e, \Delta\psi)$, local track features $(\kappa, \text{half-width})$, and a fixed-size obstacle tensor (e.g., the M nearest obstacles within a lookahead window, represented in Frenet-relative coordinates). At inference, the DT-generated (X_0, U_0) is passed directly to IPOPT as an initialization.

3. Milestone progress and preliminary results

Implemented. The current codebase includes: a unified vehicle dynamics model; a spatial direct-collocation IPOPT optimizer; consistent Frenet-to-ENU mapping between optimization and visualization; a compiled solver path; and batch evaluation scripts that sample random obstacle scenarios and apply acceptance-gated retries.

Early quantitative results. On the Medium Oval map with 3–6 randomly sampled obstacles per scenario, a

Table 1. Batch evaluation on Medium Oval (20 scenarios, mean 5 obstacles). "Baseline" refers to the first solve attempt; "Selected" refers to the accepted attempt after retries (higher N and/or denser obstacle sampling). Times/iterations reflect the recorded solver logs for this batch.

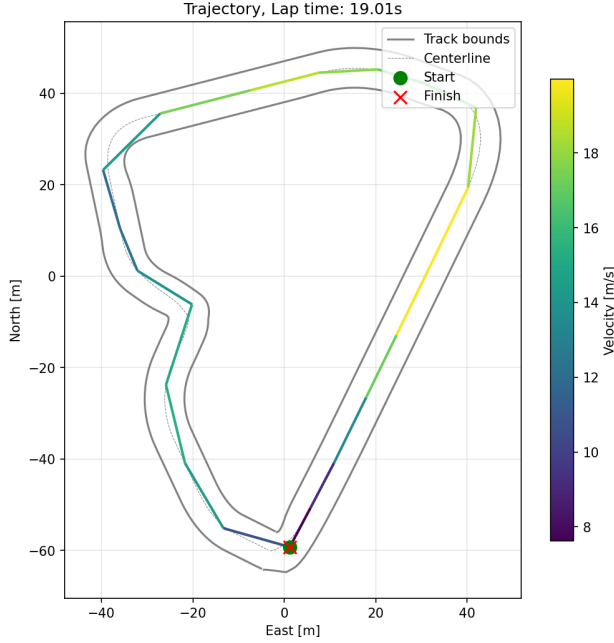| Method	Accept rate	Time [s] (mean)	IPOPT iters (mean)	Cost [s] (mean)
Baseline attempt	0.20	17.7	37.0	24.57
Selected (after retries)	1.00	21.6	32.6	22.80

Trajectory, Lap time: 19.01s

Figure 2. Representative IPOPT solution on the oval track with obstacles (current pipeline output).

batch of 20 scenarios achieved 100% IPOPT convergence and 100% acceptance under the practical clearance gate $\min(\text{dense clearance}) \geq -10^{-3}$ m. However, a single baseline attempt (with fixed N and obstacle subsampling) is accepted only 0.20 of the time; retries are frequently required to satisfy dense clearance.

Table 1 summarizes the batch statistics (selected attempt per scenario after retries), and contrasts them with the baseline attempt.

The selected attempts used the following retry outcomes: baseline accepted in 4/20 scenarios, higher-N accepted in 7/20, and higher obstacle subsampling accepted in 9/20. These statistics motivate DT warm-starting as a way to (i) reduce the need for retries and (ii) reduce the interior-point iterations required to reach a safe local optimum.

Observed limitations. Some adversarial obstacle place-

ments still yield converged but unsafe trajectories (significant negative dense clearance) even after retries, indicating that (a) the current obstacle constraint discretization may be too coarse in extreme scenes, and (b) the initializer and/or constraint enforcement needs strengthening before large-scale dataset generation.

4. Remaining work

Expert/data pipeline (near-term). (1) Finalize the Tier-1 solver settings and obstacle inflation so that acceptance-gated retries succeed on a broader set of obstacle scenes; (2) implement the canonical dataset schema and generate a sufficiently large set of accepted trajectories (obstacle-free base laps, then randomized obstacle layouts via a curriculum on obstacle count and size).

DT training and integration. (3) Train the Decision Transformer on RTG-conditioned sequences with track and obstacle features; (4) integrate DT outputs as IPOPT warm-start initializations (X_0, U_0).

Evaluation (final milestone). (5) Compare DT-warm-start vs. baseline initialization on: (i) IPOPT iterations, (ii) wall-clock solve time, (iii) acceptance rate without retries, and (iv) final lap time / constraint margins. Additional ablations will vary context length and obstacle feature sets to understand which information drives warm-start quality.

References

Aggarwal, R. and Gerdes, J. C. Friction-robust autonomous racing using trajectory optimization over multiple models. *IEEE Open Journal of Control Systems*, 2025. doi: 10.1109/OJCSYS.2025.3635449.

Chen, L., Lu, K., Rajeswaran, A., Lee, K., Grover, A., Laskin, M., Abbeel, P., Srinivas, A., and Mordatch, I. Decision transformer: Reinforcement learning via sequence modeling. In *Advances in Neural Information Processing Systems*, volume 34, pp. 15084–15097, 2021.

Guffanti, T. and D'Amico, S. Transformers for trajectory optimization with application to spacecraft rendezvous. In *IEEE Aerospace Conference*, pp. 1–15. IEEE, 2024.

Kapania, N. R., Subosits, J., and Gerdes, J. C. A sequential two-step algorithm for fast generation of vehicle racing trajectories. In *ASME Dynamic Systems and Control Conference (DSCC)*, 2015.

Subosits, J. K. and Gerdes, J. C. Impacts of model fidelity on trajectory optimization for autonomous vehicles in extreme maneuvers. *IEEE Transactions on Intelligent Vehicles*, 6(3):546–557, 2021. doi: 10.1109/TIV.2021.3051325.