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 nonlinear constrained optimization problem, solved by IPOPT over a single-track vehicle model with nonlinear tire forces, parameterized to the Stanford Dynamic Design Lab *2018 VW Golf GTI* test vehicle. The project approach is a hybrid pipeline in which DT predicts a trajectory initialization from return-conditioned expert data, and IPOPT then enforces hard dynamics and safety constraints to refine that initialization into a feasible local optimum. 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 drive-at-the-limits control problem in which a planner must coordinate steering and longitudinal force near the tire friction boundary while satisfying strict geometric constraints such as track edges and safety margins. In this regime, model fidelity and initialization quality significantly affect performance: small errors in friction modeling, load transfer, or actuator limits can produce qualitatively different local minima and very different feasibility outcomes in nonlinear solvers. Prior work highlights how sensitive racing optimization is to modeling choices and solver behavior (Kapania et al., 2015;

Subosits & Gerdes, 2021; Aggarwal & Gerdes, 2025).

This project addresses closed-track lap optimization in the presence of static obstacles. The vehicle is represented by a nonlinear single-track (bicycle) model with load-transfer states and empirical nonlinear tire forces, parameterized to the 2018/2019 VW Golf GTI used in earlier Stanford Dynamic Design Lab experiments (Subosits & Gerdes, 2021; Aggarwal & Gerdes, 2025). Solving minimum-time racing via direct collocation yields a nonlinear constrained optimal control problem that is computationally expensive and sensitive to initial guesses.

Sequence models such as the Decision Transformer (DT) reframe offline reinforcement learning as conditional sequence modeling, where trajectories are generated by conditioning on return-to-go and past states/actions without explicit policy optimization (Chen et al., 2021) — an approach that leverages transformer architectures originally developed for language modeling. In this setting, the DT is not intended to replace the controller; instead, it is tasked with proposing dynamically coherent, near-feasible initial trajectories for a downstream optimizer. This hybrid perspective resembles transformer-based warm-starting in other trajectory optimization domains (Guffanti & D'Amico, 2024) but must contend with the harder contact-like nonlinearities of tire saturation and stringent geometric safety constraints present in racing.

The core hypothesis is that training a DT on expert solutions accepted by a high-fidelity optimizer will yield initializations that reduce IPOPT iterations and runtime, eliminate most retry attempts, and preserve final lap time and constraint satisfaction.

2. Approach

2.1. Vehicle model and minimum-time NLP

The vehicle dynamics model uses a nonlinear single-track (bicycle) formulation following recent autonomous racing studies (Subosits & Gerdes, 2021; Aggarwal & Gerdes, 2025). Tires on each axle are lumped into single equivalent wheels, planar rigid-body motion is retained, and lateral/longitudinal forces are computed with a Fiala-style non-

[1]Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, USA. Correspondence to: Victor Sebastian Martinez Perez <sebasmp@stanford.edu>.

linear tire model. Load transfer between left and right sides is included as explicit states to capture normal-force redistribution under combined braking, cornering, and acceleration.

To align problem structure with track geometry, the trajectory is parameterized in terms of arc length s. The spatial domain is discretized at $N+1$ nodes indexed $k = 0, \ldots, N$ with uniform spacing Δs. Decision variables consist of state vectors $\mathbf{X} = [\mathbf{x}_0, \ldots, \mathbf{x}_N]$ and control vectors $\mathbf{U} = [\mathbf{u}_0, \ldots, \mathbf{u}_N]$:

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

Here, u_x and u_y are body-frame longitudinal and lateral speeds, r is yaw rate, δ is steering angle, and F_x is longitudinal force. Load-transfer states $\Delta F_{z,\text{long}}$ and $\Delta F_{z,\text{lat}}$ represent normal force deviations due to load shift. Path deviation e and heading error $\Delta\psi$ define position and orientation relative to the track centerline.

The optimization objective minimizes final time t_N, regularized with a control-smoothness term:

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

State evolution is enforced with trapezoidal collocation in spatial form:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\Delta s}{2} \left(f_s(\mathbf{x}_k, \mathbf{u}_k) + f_s(\mathbf{x}_{k+1}, \mathbf{u}_{k+1}) \right), \quad (3)$$

where $f_s(\cdot)$ denotes dynamics per unit arc length. This maintains sparsity and aligns with standard racing optimization practice.

Hard constraints include:

$$\dot{s}_k \geq \varepsilon_s, \qquad\qquad 1 - \kappa(s_k)e_k \geq \varepsilon_\kappa, \qquad (4)$$

ensuring forward progress and avoiding singular Frenet coordinates, respectively. Additional constraints enforce track bounds and actuator limits on δ and F_x. For closed-lap solutions, periodic boundary conditions are imposed on all variables except time: $\mathbf{x}_0 = \mathbf{x}_N$ and $\mathbf{u}_0 = \mathbf{u}_N$, with $t_0 = 0$ and t_N unconstrained. The sparse nonlinear program (NLP) is implemented in CasADi and solved with IPOPT.

2.2. Obstacle handling and acceptance-gated retries

Static circular obstacles are introduced through spatial clearance constraints evaluated at each discretization node:

$$\|p(s_k, e_k) - p_{\text{obs},i}\|_2^2 \geq (r_i + m_i + R_{\text{veh}} + c)^2, \quad (5)$$

where $p(s, e)$ maps Frenet coordinates to world coordinates, $p_{\text{obs},i}$ and r_i define obstacle i, m_i is an obstacle margin, R_{veh}

is an inflated vehicle footprint radius, and c is additional clearance. Node-level enforcement is tractable, and safety is verified with a dense post-solve collision check along each segment.

Because IPOPT can converge to a discretized solution that violates dense collision checks—especially under coarse discretization—an *acceptance-gated retry policy* is applied. If IPOPT reports convergence but dense clearance fails beyond tolerance, the same scenario is re-solved with increased node count N and/or finer obstacle sampling. This yields a reliable set of accepted expert demonstrations while surfacing failure modes that warm-starting should reduce.

2.3. Dataset contract (expert trajectories)

Each accepted solve is recorded as an episode with:

(i) a reproducibility header (map identifier/hash, discretization parameters 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) from Frenet mapping, local track properties $(\kappa(s_k), \text{half-width})$, and return-to-go defined as

$$\text{RTG}_k = -\sum_{j=k}^{N-1} \Delta t_j. \qquad (6)$$

Including load-transfer states and full solver outputs ensures future extensibility even if the Decision Transformer observes only a subset of states.

2.4. Decision Transformer warm-start (planned)

The Decision Transformer is trained as a causal decoder predicting controls conditioned on past information and return-to-go. At each step k, the input vector is

$$\mathbf{z}_k = [\text{RTG}_k, \ \mathbf{o}_k, \ \mathbf{u}_{k-1}], \qquad (7)$$

where $\mathbf{o}_k$ includes observable state components $(u_x, u_y, r, e, \Delta\psi)$, local track features $(\kappa, \text{half-width})$, and a fixed-size obstacle tensor representing the M nearest obstacles within a look-ahead window, encoded in Frenet-relative coordinates.

During inference, the DT generates a full trajectory $(\mathbf{X}_0, \mathbf{U}_0)$ that is passed directly to IPOPT as an initialization. The expectation is that these warm starts will reduce IPOPT iteration count and runtime while maintaining feasibility and solution quality.

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 it
Baseline attempt	0.20	17.7	3
Selected (after retries)	1.00	21.6	3

Figure 1. Representative IPOPT solution on the oval track with obstacles (current pipeline output).

3. Milestone progress and preliminary results

Track geometry for these results is loaded from precomputed map files (centerline, width, and curvature, with banking/grade when available); the reported experiments use the *Medium Oval* map (approximately 260 m lap length) with randomized circular obstacles and an inflated vehicle footprint.

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 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 placements 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.
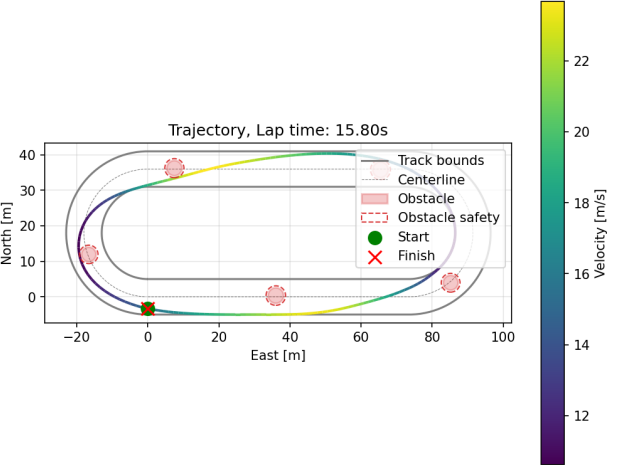
4. Remaining work

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.