Appendix

In this supplementary material, we first present the templates for QAs and prompts of pre-training and fine-tuning data (Appendix A). Then, we provide detailed metric information for our closed-loop evaluation (Appendix B), as well as the specific 14 scenario types defined in nuPlan challenges [3] (Appendix C). In the end, we report additional quantitative results conducted to further substantiate the efficacy of our approach (Appendix D).

A Data Templates

A.1 Pre-training Data

Planning-QA We assessed the capabilities of Llama2-13B [4] via a zero-shot methodology and discovered that its extensive pre-training data provides a solid foundation in traffic rule comprehension. However, its limited mathematical prowess poses challenges in grasping and deducing the connections between instructions and numerical expressions. To address this, we introduced a language-based *Planning-QA* dataset aimed at transitioning LLM from a general model to a specialized model adept in autonomous driving planning. This enhancement focuses on refining its capability in instruction interpretation and reasoning.

Concretely, we delineated the level of autonomous driving planning into three granularities: 1) high-level instructions: formulated through velocity commands including *stop*, *accelerate*, *decelerate*, *maintain speed*, and routing commands including *turn left*, *turn right*, *go straight*, 2) control: assessing the values of velocity and acceleration, 3) and waypoint: encompassing a series of points. Six question types were devised to articulate the transitional relationships across the *high-level instructions - control - waypoint* spectrum, and each QA-pair is adapted based on log data from nuPlan [1]. Fig. S1a illustrates the universal system prompt template applicable to all questions, whereas Fig. S1b-S1g display specific examples for each question type alongside their respective answer, substituting <**Question**> and <**Answer**> within the system prompt.

- Input:
Role: You are now an autonomous driving driver, you need to answer questions based on the given high-level instructions, control or waypoints information.
Coordinate: Your current coordinate is at (0,0), the positive direction of the x-axis is to the front, and the positive direction of the y-axis is to the left. The waypoints are represented as a series of equally time-spaced points, with the time interval between two adjacent points being 0.5 s.
Question: ***<Question>***
Final Answer:

- Target:
<Answer>

(a) system prompt template for Planning-QA dataset

<Question>
Given the waypoints ***[[1.543 -0.123] [4.337 -0.251] [7.463 -0.469] [10.919 -0.8]]*** *and the high-level instructions* ***turn right for 9.402 m, during the way, we decelerate****, are they consistent? If not, what should the corresponding high-level instructions be? And what exactly is different about the high-level instructions?*

<Answer>
- Positive answer:
The high-level instructions corresponding to the given waypoints is ***turn right for 9.402 m, during the way, we decelerate****, which is consistent with the given high-level instructions.*

- Negative answer:
The high-level instructions corresponding to the given waypoints is ***go straight for 9.402 m, during the way, we accelerate****, which is not consistent with the given high-level instructions* ***turn right for 9.402 m, during the way, we decelerate****. The given high-level instructions requires a* ***turn right*** *and a* ***decelerate****, but the waypoints correspond to* ***go straight*** *and* ***accelerate****.*

(b) *waypoints - high-level instructions*

<Question>
The driving waypoints ***[[1.493 -0.574] [4.073 -0.577] [6.653 -0.58] [9.233 -0.583]]****, initial speed* ***5.16 m/s****, accelerating time* ***1.5 s*** *and acceleration* ***1.05 m/s**2*** *are known. Known that acceleration is a fixed value, does the given acceleration matches with the waypoints? If not, what should the real acceleration value be?*

<Answer>
- Positive answer:
Through the given waypoints and initial velocity, the acceleration corresponding to the waypoints is ***1.05 m/s**2****, which is consistent with the given acceleration.*

- Negative answer:
Through the given waypoints and initial velocity, the acceleration corresponding to the waypoints is ***0.1 m/s**2****, which is* ***smaller (resp. greater)*** *than the given acceleration* ***1.05 m/s**2****.*

(c) *waypoints - control*

<Question>
Driving following the lane, where the centerline ahead is ***[[1.543 -0.123] [1.789 -0.132] [2.035 -0.141] [2.281 -0.151] [2.526 -0.161] [2.772 -0.172] … [13.058 -1.053] [13.302 -1.084] [13.546 -1.116]]****. And the initial speed* ***11.721 m/s****, acceleration* ***1.318 m/s**2****, driving time* ***0.5 s*** *are known, what should the future waypoints be?*

<Answer>
Based on the given input, the waypoints in the future ***0.5 s*** *is* ***[[1.543 -0.123] [7.558 -0.476]]****.*

(d) *high-level instructions - waypoint*

<Question>
Known that the initial speed is ***5.305 m/s****. Can we stop in* ***3.0 s*** *if the acceleration is* ***-0.402 m/s**2****?*

<Answer>
- Positive answer:
Based on the given input, we could stop in ***3.0 s****.*

- Negative answer:
Based on the given input, we could not stop in ***3.0 s****.*

(e) *high-level instructions - control*

<Question>
Does the waypoints ***[[1.482 -0.853] [4.136 -0.853] [6.789 -0.848] [9.442 -0.85]]*** *consistent with the condition that driving follow the lane* ***[[1.482 -0.853] ... [11.218 -0.855 [11.429 -0.856] [11.641 -0.857] [11.853 -0.858]]****, where the initial speed is* ***5.307 m/s****, acceleration is* ***1.7 m/s**2****? If not, what should the real acceleration be? If it does, how long is the driving time?*

<Answer>
- Positive answer:
The waypoint is consistent with the given input. And the driving time corresponding to the waypoints ***1.5 s****.*

- Negative answer:
The waypoint is not consistent with the given input. The acceleration corresponding to the given input is ***0.1 m/s**2****, but the given acceleration is* ***1.7 m/s**2****.*

(f) *control - waypoint*

<Question>

*Given the initial speed **1.033 m/s** and acceleration **-0.007 m/s**2**, and driving along the centerline **[[1.413 -0.19] [1.668 -0.187] [1.922 -0.183] … [13.123 -0.012] [13.378 -0.013] [13.632 -0.015] [13.887 -0.017]]**, according to the current state, what should the corresponding high-level instructions be? Is it consistent with the given high-level instructions **turn left for 115.5 m, during the way, we stop**?*

<Answer>

- Positive answer:

*Based on the given input, the high-level instructions should be **turn left for 115.5 m, during the way, we stop**, which is consistent with the given high-level instructions.*

- Negative answer:

*Based on the given input, the high-level instructions should be **go straight for 115.5 m, during the way, we decelerate**, which is not consistent with the given high-level instructions **turn left for 115.5 m, during the way, we stop**. The given high-level instructions requires a **turn left** and a **stop**, but the waypoints correspond to **go straight** and **decelerate**.*

(g) *control - high-level instructions*

Fig. S1: Six question types were devised to articulate the transitional relationships across the *high-level instructions - control - waypoints* spectrum. All data are generated based on logs from nuPlan, with the red parts being subject to change based on actual data. Some data have both positive and negative answers, which are both displayed in the corresponding templates. The waypoints are abbreviated due to space limitations.

Reasoning1K Additionally, we have crafted 1,000 QA-pairs using GPT-4, tailored to tasks that require not just decision-making but also in-depth reasoning. Each sample includes a question that sets the context, provides scene descriptions, outlines considerations based on traffic rules, details the anchor trajectory from real-time planners, and specifies requirements for the answers, while the answer delivers trajectory predictions along with the decisions and the underlying thought process. By training with mixed of *Reasoning1K* and *Planning-QA*, we have substantially enriched the dataset's diversity and complexity. This approach aims to bolster LLM's ability to deduce answers through logical reasoning, seamlessly integrating traffic rules with detailed, scenario-based analysis. Fig. S2a and Fig. S2b showcase an example of the input and its corresponding target from the dataset, respectively.

A.2 Fine-tuning Data

The pre-training phase aims to utilize language-based data to bolster the LLM's understanding of instructions, whereas the fine-tuning phase seeks to further refine this comprehension by aligning it with vectorized scenes. This entails a process characterized by multi-modal inputs and vectorized supervision. Inputs encompass two parts: linguistic prompts and vectorized scene data. The linguistic portion includes fixed system prompts and log-based routing instructions. Within this template, the <**map**> token is substituted with tokens of vectorized scene at the point of input. Instructions are generated from the forthcoming 8 seconds' logs through a hand-crafted rules, encompassing four instruction categories—*turn left*, *turn right*, *go straight*, and *stop* — paired with pertinent distance details. Notably, the *stop* is only employed when the actual path length for

- Input:
Role: You are now an autonomous driving driver, I will provide you with a predicted future trajectory. Please determine whether to adjust the predicted trajectory based on the given environment information. Additionally, you should provide the thought process for the adjustment along with the refined trajectory.

Context:
- Trajectories: Trajectories are represented by 16 points. Each point is represented by (x, y). The time interval between two points is 0.5s.

Environment:
- Speed: Car speed on X-axis is 6.3m/s, car speed on Y-axis is -0.0m/s.
- Location: Car location is (0.0, 0.0).
- Acceleration: Car acceleration on X-axis is 1.29m/s^2, car acceleration on Y-axis is 0.0m/s^2.
- Lane: You are in the 1st lane from the left. ***There are 1 lanes in total****, which means there are 0 lanes on your left, and 0 lane on your right.* ***The width of each lane is 3.75m.***
- Sampled 16 points of the center of the 1st lane from the left.: [(1.46, 0.0) ... (70.47, -21.34)]
- ***Traffic Light: traffic lights of incoming lane connector are: GREEN***
- ***Other Vehicles: There are 0 vehicles or pedestrians near you.***

Traffic Rules:
1. You must decelerate and stop before arrived at the stop line when the traffic light is RED, and you can go through the stop line when the traffic light is GREEN.
2. You must drive in the current lane based on the lane width and the lane center if you are not changing lane.
3. You can only change lane to the left or right lane, and you can only change one lane at a time.
4. ***You can accelerate if there is no car in front of you, but you can not exceed the speed limit of 50.0m/s.***
5. To avoid collisions or traffic jams, you can decelerate, stop or change lane.
6. By default, other vehicles will not change lanes on their own.

Inputs:
- Predicted Ego Future Trajectories in next 8.0s : [(1.46, 0.0), (4.73, 0.47) ... (42.56, 7.91), (43.6, 7.99)]

Outputs:
- Refined Ego Future Trajectories in next in next 8.0s.
- Decision: How to adjust the predicted trajectory, such as accelerate, decelerate, stop or change lane, etc.
- Thought process: Think step by step why the Refined trajectory is better than the predicted trajectory. You can only analysis based on the environment information. Besides, You must contain the follow things:
1. Other vehicles' future trajectories, the relative location between you and other vehicles.
2. Traffic rules.

(a) example of input for Reasoning1K

- Target:
Refined Ego Future Trajectories in next 8.0s : [(4.63, 0.29) ... (53.01, 10.38), (55.96, 10.95)]

Decision: The predicted trajectory needs to be adjusted to align more closely with the center of the lane. This can be achieved by maintaining the current acceleration and speed on X-axis, while slightly increasing the speed on the Y-axis.

Thought Process:
1. *Other Vehicles' Future Trajectories and Relative Location:* ***There are no other vehicles or pedestrians near us, so we don't need to consider their future trajectories or our relative location to them. This gives us more flexibility in adjusting our trajectory.***
2. *Traffic Rules****: The traffic light is green, so we are allowed to proceed through the intersection without stopping. We are currently in the only lane, so we cannot change lanes.*** *However, we need to* ***stay within the lane, which is 3.75m wide****. The predicted trajectory deviates from the center of the lane, which is not ideal.* ***There is no speed limit constraint as our current speed is well below the limit of 50.0m/s.*** *Since there are no other vehicles in front of us, we can maintain or even slightly increase our speed if necessary.*

Based on these considerations, it's clear that we need to adjust our trajectory to stay closer to the center of the lane. This can be achieved by maintaining our current speed and acceleration on X-axis, while slightly increasing our speed on the Y-axis. This will allow us to move more towards the center of the lane, ensuring a safer and more efficient drive.

(b) example of corresponding target for Reasoning1K

Fig. S2: An example from *Reasoning1K*, where S2a displays the input, and S2b shows the corresponding target. The reasoning-related parts are highlighted in color, with the same color in both input and target representing the cause and effect in the reasoning process respectively.

Role: You are now an autonomous driving driver, and I will provide you with the environment information including Ego Car Information, Agents Information and Map Information.

Environment: ***<map>***

Routing instructions: ***Go straight in 9.01 m. Turn right in 119.93 m.***

Please predict the future waypoints based on the given environmental information and routing instructions.

Final Answer:

Fig. S3: Prompt template for fine-tuning stage.

the next 8 seconds falls below 0.5 meters. An illustration of the input template is provided in Fig. S3.

B Closed-loop Metrics

In this section, we elucidate the metrics of closed-loop evaluation deployed within the nuPlan [1] framework. The calculation for the composite metric *Score* is detailed in Formula 1.

Drivable Area Compliance (Drivable) Ego should remain within the designated drivable area. If a frame is identified where the maximum distance from any corner of the ego's bounding box to the nearest drivable area exceeds a predetermined threshold, the drivable area compliance score is assigned a value of 0; otherwise, it is assigned a value of 1.

Driving Direction Compliance (Direct.) The evaluation of the ego's compliance with the correct driving direction is encapsulated within this metric. It stipulates that a distance traversed in the opposite direction of less than 2 meters within a 1-second interval is deemed compliant, warranting a score of 1. A traversal exceeding 6 meters within the same timeframe indicates a significant deviation from compliance, thus yielding a score of 0. Intermediate distances are assigned a proportional score of 0.5.

Ego is Comfortable (Comf.) This metric quantifies the comfort of the trajectory based on a series of kinematic and dynamic thresholds including longitudinal acceleration, lateral acceleration, yaw acceleration, yaw velocity, longitudinal jerk, and overall jerk magnitude. Compliance across all specified parameters results in a trajectory being classified as "comfortable", meriting a score of 1; otherwise, the score is 0.

Ego Progress Along Expert Route (Prog.) This metric represents the ratio of the ego‘s progress per frame to that of an expertly defined route. The closer the ego's progress is to the expert route's progress, the higher the score, with the maximum score being 1.

No Ego At-fault Collisions (Coll.) This metric is integral to the safety evaluation, quantifying instances where the ego is directly responsible for collisions. An absence of at-fault collisions throughout the evaluation period scores a 1. A singular collision with a static object, such as a traffic cone, is deemed a minor infraction, scoring a 0.5, whereas all other scenarios result in a score of 0, accentuating the paramount importance of collision avoidance.

Speed Limit Compliance (Lim.) The ego's adherence to speed regulations is scrutinized by calculating the average instance of speed limit violations per frame. This metric deducts points based on the proximity of the violation to the maximum speed threshold, promoting strict compliance with speed limits.
Time to Collision within Bound (TTC) This metric evaluates the hypothetical time to collision (TTC) with any external agent if the ego were to continue on its trajectory with unaltered speed and heading. A TTC exceeding 0.95 seconds is considered safe, thus earning a score of 1. Lower TTC values signify elevated risk, consequently scoring a 0, highlighting the significance of maintaining a safe buffer from other road users.
Making Progress (MP) An ancillary metric to the Ego Progress Along Expert Route, it assigns a score of 1 if the progress exceeds a threshold of 0.2, otherwise scoring 0. Though not directly incorporated into the main text due to its derivative nature, it contributes to the final score.
Score The final scenario score is calculated by combining the above metrics in a predetermined manner, providing a comprehensive assessment of the autonomous vehicle's performance within the evaluated scenarios:

$$\begin{aligned} Score = (Prog * 5 + TTC * 5 + Lim * 4 + Comf * 2) \div 16 \\ * Coll * Drivable * MP * Direct \end{aligned} \tag{1}$$

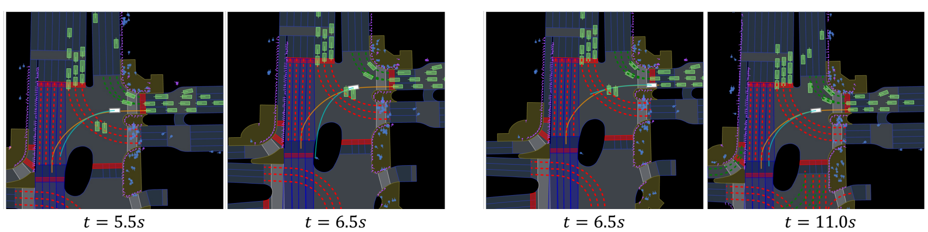
C Scenario Types

In this section, we delineate 14 scenario types employed in both the training and evaluation which is consistent with the nuPlan Challenge 2023 [3]. These scenarios encompass a broad spectrum of driving conditions to evaluate the adaptability and planning capabilities of autonomous vehicles. The scenarios are categorized as follows:
Behind Long Vehicle (type 0): The ego vehicle is positioned behind a long vehicle, maintaining a longitudinal distance of 3 to 10 meters within the same lane, where the lateral distance is less than 0.5 meters.
Changing Lane (type 1): Initiates a maneuver to transition towards an adjacent lane at the beginning of the scenario.
Following Lane with Lead (type 2): Involves the ego vehicle following a leading vehicle that is moving in the same lane with a velocity exceeding 3.5 m/s and a longitudinal distance of less than 7.5 meters.
High Lateral Acceleration (type 3): Characterized by the ego vehicle experiencing high acceleration ($1.5 < \text{acceleration} < 3\,\text{m/s}^2$) across the lateral axis, accompanied by a high yaw rate, without executing a turn.
High Magnitude Speed (type 4): The ego vehicle achieves a high velocity magnitude exceeding 9 m/s with low acceleration.
Low Magnitude Speed (type 5): Ego vehicle moves at low velocity magnitude ($0.3 < \text{velocity} < 1.2\,\text{m/s}$) with low acceleration, without coming to a stop.
Near Multiple Vehicles (type 6): The ego vehicle is proximate to multiple (more than six) moving vehicles within a distance of less than 8 meters while maintaining a velocity of more than 6 m/s.

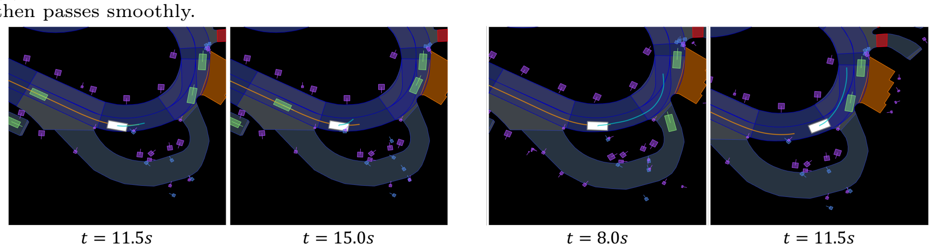
Starting Left Turn (type 7): Marks the commencement of a left turn by the ego vehicle across an intersection area, without being halted.
Starting Right Turn (type 8): Marks the commencement of a right turn by the ego vehicle across an intersection area, without being halted.
Starting Straight Traffic Light Intersection Traversal (type 9): The ego vehicle begins to traverse straight through an intersection controlled by traffic lights, without being stopped.
Stationary in Traffic (type 10): The ego vehicle remains stationary amidst traffic, surrounded by multiple (more than six) vehicles within an 8-meter radius.
Stopping with Lead (type 11): The scenario begins with the ego vehicle initiating deceleration (acceleration magnitude less than $-0.6\,\mathrm{m/s}^2$, velocity magnitude less than $0.3\,\mathrm{m/s}$) due to the presence of a leading vehicle ahead within a distance of less than 6 meters.
Traversing Pickup/Dropoff (type 12): The ego vehicle navigates through a pickup or drop-off area without stopping.
Waiting for Pedestrian to Cross (type 13): The ego vehicle waits for a pedestrian, who is within an 8-meter distance and less than 1.5 seconds away from the intersection, to cross a crosswalk. This occurs while the ego vehicle is not stopped and the pedestrian is not located in a pickup or drop-off area.

D Visualization

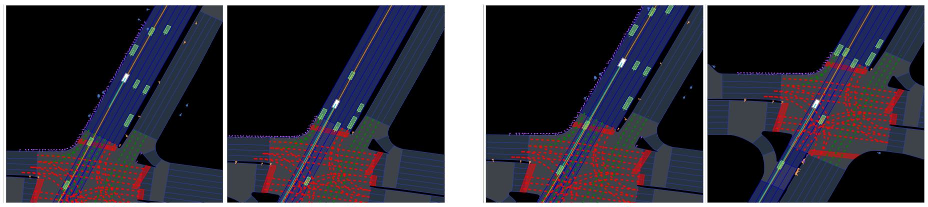
Fig. S4 additionally showcases a series of our visualization results and demonstrates its superior performance to GameFormer [2] across multiple metrics.



(a) GameFormer collides with other vehicles at 6.5s; AsyncDriver first yield to the vehicle going straight, then passes smoothly.

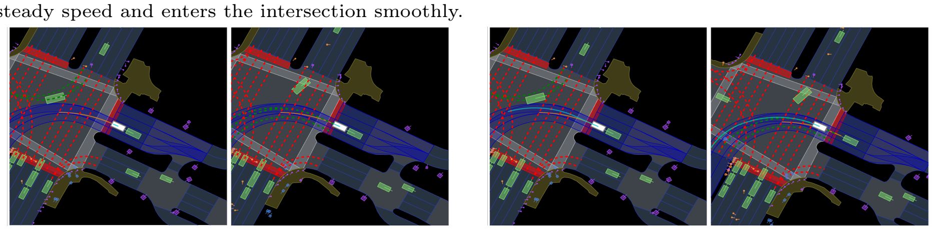


(b) GameFormer, upon reaching 11.5s, finds it difficult to evade pedestrians ahead, resulting in a collision at 15.0s; AsyncDriver, anticipating the situation in advance, passes smoothly before the pedestrians approach.



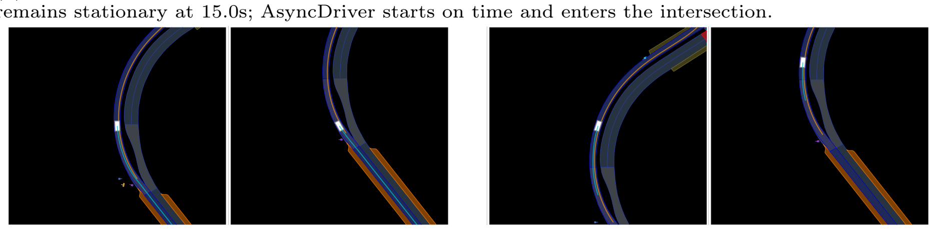
$t = 5.0s$ $t = 12.5s$ $t = 5.0s$ $t = 12.5s$

(c) GameFormer makes a misjudgment about the traffic light status at 5.0s, noticeably decelerating, leading to its failure to enter the intersection at 12.5s due to slow progress; AsyncDriver maintains a steady speed and enters the intersection smoothly.



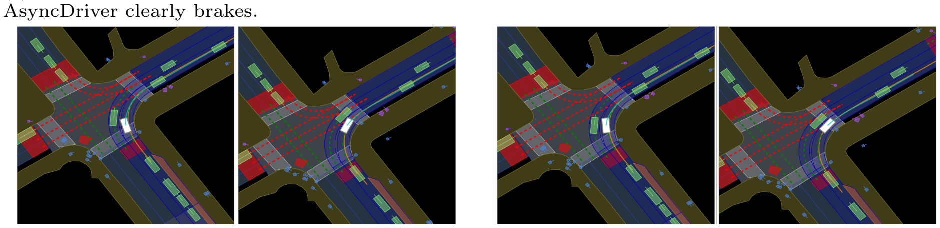
$t = 11.0s$ $t = 15.0s$ $t = 11.0s$ $t = 15.0s$

(d) GameFormer, after the traffic light turns from red to green at 11.0s, fails to start promptly and remains stationary at 15.0s; AsyncDriver starts on time and enters the intersection.



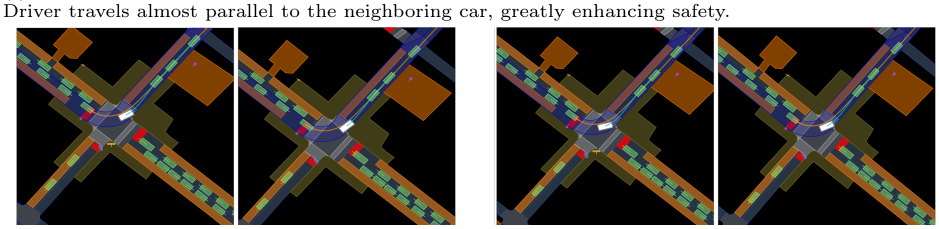
$t = 12.5s$ $t = 15.0s$ $t = 12.5s$ $t = 15.0s$

(e) At 12.5 seconds, when a pedestrian appears ahead, GameFormer does not evade or decelerate; AsyncDriver clearly brakes.



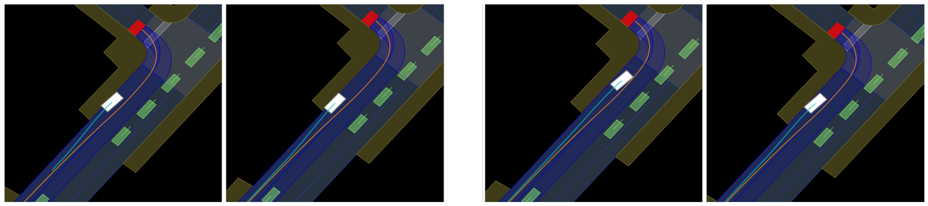
$t = 3.5s$ $t = 5.0s$ $t = 3.5s$ $t = 5.5s$

(f) GameFormer gets too close to the adjacent car while turning, increasing the risk of collision; AsyncDriver travels almost parallel to the neighboring car, greatly enhancing safety.



$t = 4.5s$ $t = 5.0s$ $t = 4.5s$ $t = 5.0s$

(g) GameFormer turns out of the drivable area; AsyncDriver does not.



$t = 10.0s$ $t = 11.0s$ $t = 10.0s$ $t = 11.0s$

(h) GameFormer exceeds the drivable area; AsyncDriver does not.

Fig. S4: The visualization results of GameFormer and AsyncDriver in various scenarios, with corresponding analytical explanations provided in subcaptions. The left two columns display the results for GameFormer, while the right two columns show the results for AsyncDriver.

References

1. Caesar, H., Kabzan, J., Tan, K.S., Fong, W.K., Wolff, E., Lang, A., Fletcher, L., Beijbom, O., Omari, S.: nuplan: A closed-loop ml-based planning benchmark for autonomous vehicles. arXiv preprint arXiv:2106.11810 (2021)
2. Huang, Z., Liu, H., Lv, C.: Gameformer: Game-theoretic modeling and learning of transformer-based interactive prediction and planning for autonomous driving. arXiv preprint arXiv:2303.05760 (2023)
3. Motional: nuplan challange. `https://github.com/motional/nuplan-devkit` (2023)
4. Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., et al.: Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288 (2023)