Tracing the Influence of Predecessors on Trajectory Prediction
Supplementary Material

Mengmeng Liu, Hao Cheng, Michael Ying Yang
University of Twente, The Netherlands
mengmeng.liu1998@gmail.com; {h.cheng-2, michael.yang}@utwente.nl

In this supplementary material, we provide further information about the computational performance of the proposed Predecessor and Successor (PnS) method in Section A and experimental results in Section B.

A. Computational performance

Table 1 demonstrates the computational performance of PnS with the Predecessor Tracing module in terms of model size and inference speed. It can be seen that LAformer+PnS demonstrates a good inference speed, *e.g.* $22\,\mathrm{ms}$ for 12 agents and $95\,\mathrm{ms}$ for 32 agents, which is faster than the data sampling rate of nuScenes ($10\,\mathrm{Hz}$). Also, in the same batch-size setting, LAformer+PnS is much more lightweight in terms of model size and also has a faster inference speed than LAformer+HD. This is not surprising because compared with the Predecessor Tracing module, extracting HD map information requires more powerful hidden layers and takes more time to align lane segments with motion dynamics to guide the prediction.

Model	#Params	Inference speed	
		Batch size	Time (ms)
LAformer [2]+PnS	377K	32	95
LAformer [2]+PnS	377K	12	22
LAformer [2]+HD	482K	12	82
LAformer [2]+HD	482K	32	215

Table 1. Model size and inference speed.

B. Further analysis of experimental results

Predecessor Tracing. To further analyze the efficacy of the Predecessor Tracing module, we evaluate the performance difference by adding distance and angle thresholds in the process of identifying potential predecessors, as shown in Table 2.

Specifically, we only search for potential predecessors within $20\,\mathrm{m}$ in L_2 distance and $[-90, 90]$ degrees in the field-of-view of the successor, while ignoring other agents outside this area. This setting ensures that the identified predecessor agents are spatially close to the successor and driving in a feasible direction, thus avoiding false positive predecessor identification. However, only a marginal performance difference measured in mFDE_{10} is found by adding these thresholds. As we mentioned in the main paper, when there is no agent within this area satisfying these thresholds, the prediction rolls back to conditioning on the past trajectories with an empty predecessor. Nevertheless, distant agents or agents with large angles may also carry benefits information. For instance, forward driving agents can indicate the driving direction of the lane even these agents are far away from the successor, and an agent may change to the opposite lane after making a U-turn. These thresholds may result in a reduced recall of the true predecessors. Based on this observation we assume that the lack of observations and reduced recall due to the thresholds may contribute to the performance difference.

Threshold	mADE_{10}	mFDE_{10}
-	1.21	2.32
$\surd$	1.21	2.34

Table 2. Thresholding the distance and heading angles for identifying potential predecessors. According to this configuration, the candidate predecessors are identified within $20\,\mathrm{m}$ L_2 distance and $[-90, 90]$ degrees in the field-of-view of the successor. Other agents outside this area are ignored.

Challenging cases. Figure 1 presents examples of challenging cases where LAformer+PnS faces difficulties with scene constraints. Due to the limited observation of predecessors, LAformer+PnS cannot always accurately align its multimodal predictions in accordance with lane connections and driving directions. For instance, in the first-row scenario, the model predicts only straight-forward driving, despite through traffic and left-turn traffic sharing a significant segment of the same lane. This is because few vehicles were observed for the left turn. In the other three scenarios, even though at least one predicted trajectory overlaps well

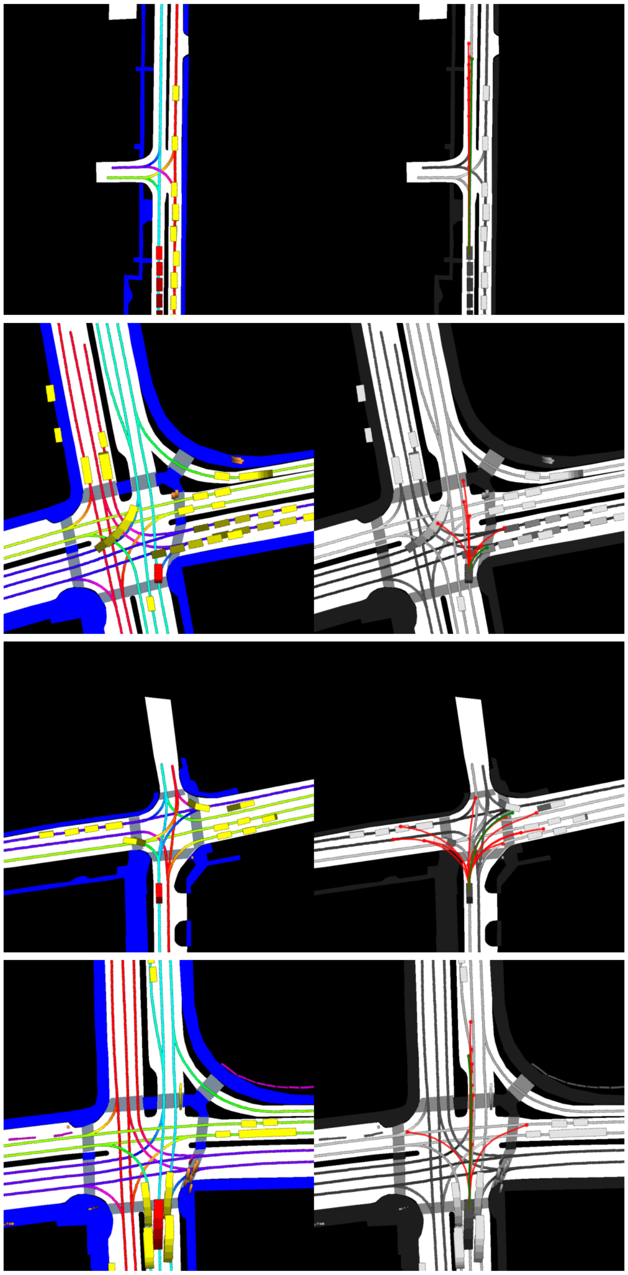

Figure 1. Challenging cases where LAformer+PnS has difficulties in following the lane connections. Following [1], we adopt the same visualization scheme to present the traffic situations on the left column and the corresponding predictions on the right column. The predictions are in red and the corresponding ground truth trajectories are in green.

with the corresponding ground truth trajectory, some of the predicted trajectories are not feasible in terms of driving directions and lane connections. This indicates that in some cases, it is not sufficient to estimate scene constraints based only on the predecessors' trajectories.

Overall, as illustrated in Table 3 in the main paper, the inclusion of PnS with the Predecessor Tracing module significantly improves prediction accuracy when compared to the map-less setting. We envision that our Predecessor Tracing module can inspire the development of more sophisticated techniques for identifying predecessors to address the increased challenges in trajectory prediction tasks that lack explicit scene information. We defer the exploration of more sophisticated solutions, such as knowledge distillation [3], for addressing scenarios involving neither HD maps nor predecessors, to future work.

References

[1] Nachiket Deo, Eric Wolff, and Oscar Beijbom. Multimodal trajectory prediction conditioned on lane-graph traversals. In *Conference on Robot Learning*, pages 203–212. PMLR, 2022. 2

[2] Mengmeng Liu, Hao Cheng, Lin Chen, Hellward Broszio, Jiangtao Li, Runjiang Zhao, Monika Sester, and Michael Ying Yang. Laformer: Trajectory prediction for autonomous driving with lane-aware scene constraints. *arXiv preprint arXiv:2302.13933*, 2023. 1

[3] Yiqi Zhong, Zhenyang Ni, Siheng Chen, and Ulrich Neumann. Aware of the history: Trajectory forecasting with the local behavior data. In *European Conference on Computer Vision*, pages 393–409. Springer, 2022. 2