

W2W: Language-Model-Based Trajectory Prediction with Reinforcement Learning

Supplementary Material

6. Supplementary Details of Interaction-Aware Prompt Construction

This appendix gives a few implementation details of the interaction-aware prompt construction in Sec. 3.3. We only retain the experimental details required to reproduce the prompt construction process, including the preprocessing steps, threshold settings, a pseudocode snippet for interaction behavior judgment, and one prompt example. The same procedure is used in both training and test stages. All interaction cues are extracted solely from the observed 8-frame history trajectories, and future trajectory points are not used during the labeling process.

6.1. Observed-history-only preprocessing setting

For each sample, the preprocessing script first parses the observation text to obtain the target pedestrian trajectory and all neighboring pedestrian trajectories. It then computes distance and heading cues within the observed window only ($T_{obs} = 8$), and uses these quantities to assign one interaction label to each target–neighbor pair. The 12-frame future trajectory is not involved in this labeling process; it is only used in later stages: as the supervision target in the SFT stage, and for reward computation in the RL stage after the generated trajectory has been parsed.

6.2. Threshold settings in preprocessing

The thresholds used for interaction labeling include the distance thresholds d_c , d_{near} , d_{far} , $d_{f_{min}}$, and $d_{f_{max}}$, as well as the angular thresholds τ_{Align} and τ_{turn} . The distance thresholds are first defined in meters in the real scene and then mapped to the pixel coordinate system according to the meter-to-pixel conversion ratio of each scene. We set these distance thresholds to $d_c = 0.8\text{m}$, $d_{near} = 0.75\text{m}$, $d_{far} = 1.5\text{m}$, $d_{f_{min}} = 1.0\text{m}$, and $d_{f_{max}} = 2.0\text{m}$. For the angular thresholds, τ_{Align} is set to $\pi/8$ in the companion rule and to $\pi/6$ in the following rule, while τ_{turn} is set to $\pi/4$ in the obstacle rule. The relative-bearing thresholds use fixed angular boundaries: $\phi_{final} < \pi/4$ for companion, $\phi_{final} > \pi/3$ for obstacle, and $|\pi - \phi_{final}| < \pi/6$ for following.

6.3. Pseudocode and Example

An example prompt after preprocessing is shown below. The target history is always written first, followed by the retained interaction descriptions:

question: What trajectory does pedestrian 0 follow for the next 12 frames?

Algorithm 1 Interaction labeling

Require: target history s^{tar} , neighbors \mathcal{N} , thresholds

Ensure: retained interaction sentences for prompt construction

```
1: for each  $s^{nb} \in \mathcal{N}$  do
2:   compute distance and heading cues from the observed 8
   points
3:   if the obstacle rule is satisfied then
4:     label  $\leftarrow$  obstacle
5:   else if the companion rule is satisfied then
6:     label  $\leftarrow$  companion
7:   else if the following rule is satisfied then
8:     label  $\leftarrow$  following
9:   else
10:    label  $\leftarrow$  unrelated
11:  end if
12:  if label  $\in$  {obstacle, companion, following} then
13:    append one interaction sentence to the prompt context
14:  end if
15: end for
```

context: Pedestrian 0 moved along the trajectory [(106, 105), (111, 106), (116, 106), (121, 106), (125, 105), (128, 103), (132, 101), (136, 98)] for 8 frames. Pedestrian 1 is an obstacle to pedestrian 0, moved along the trajectory [(176, 102), (170, 103), (163, 105), (157, 106), (151, 106), (146, 106), (140, 106), (135, 105)] for 8 frames.

answer:

This example only uses the 8 observed history points when producing the interaction sentence. The 12-frame future trajectory is not referenced during interaction labeling and is used only as the supervision target for SFT (and as ground truth for reward computation after parsing during RL).