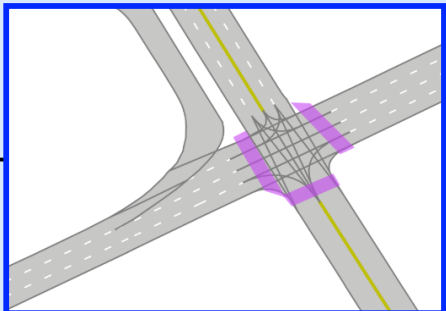
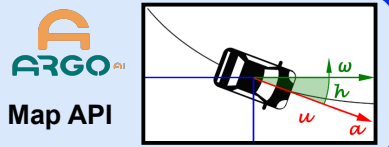


# Input

HD Map



Physics-based Heuristics



Agents Past Trajectories



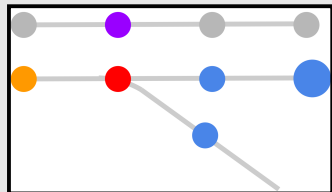
$T_0$   
 $T_{-1}$   
 $T_{-2}$   
 $T_{-3}$   
...  
 $T_{-n}$

$\Delta v +$   
Metadata

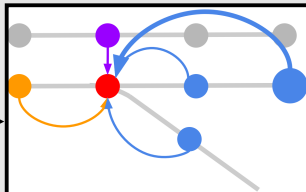
Heuristic Lane Proposals

Map Encoder

Build Graph



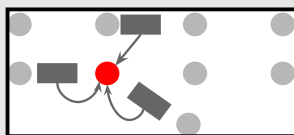
LaneGCN



Latent graph

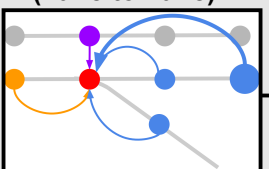
Fusion Cycle

Cross-Attention  
(Agent to Lane)



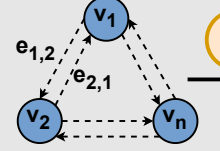
1

LaneGCN  
(Lane to Lane)



2

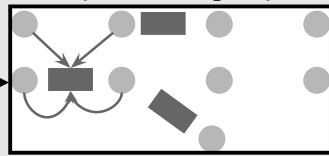
Crystal-GCN  
(Agent to Agent)



4

3

Cross-Attention  
(Lane to Agent)



Social Encoder

Positional Encoding

Linear  
Encoder (FC)

+

$\sim$

+

$\times L_h$

MHSA

+

Add & Norm

+

Feed Forward

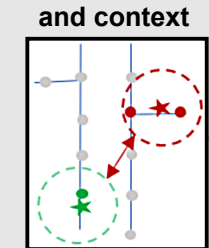
+

Add & Norm

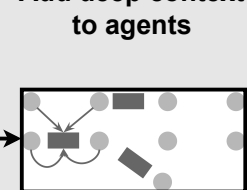
Latent actors

Goal Areas estimation

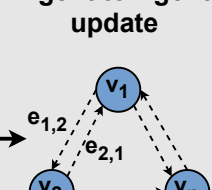
Estimate goals  
and context



Add deep context  
to agents



Agent to Agent  
update



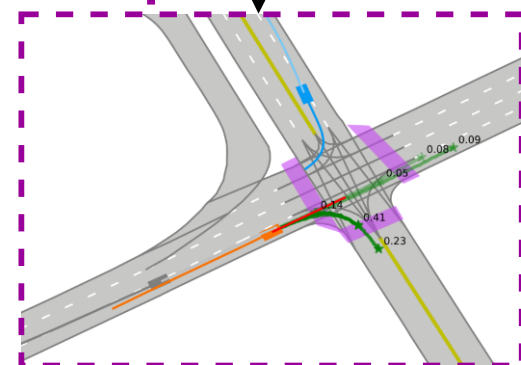
Latent actors with  
deep area context

Multimodal Decoder  
(Regression & Confidences)

Preliminary  
multimodal predictions

Motion  
Refinement

Output



Refined Multimodal trajectories