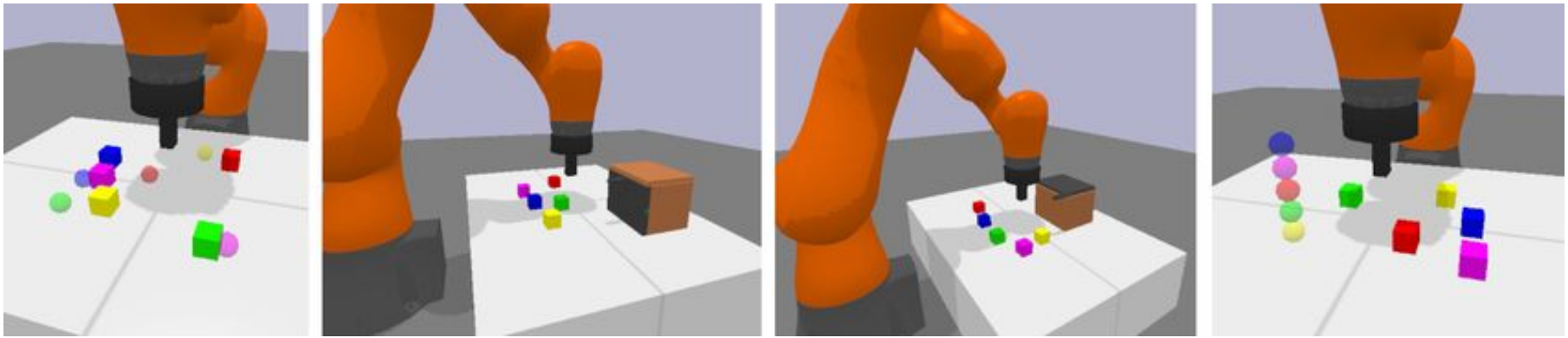
# CSC2626 Final Project

## Planning with Diffusion Models

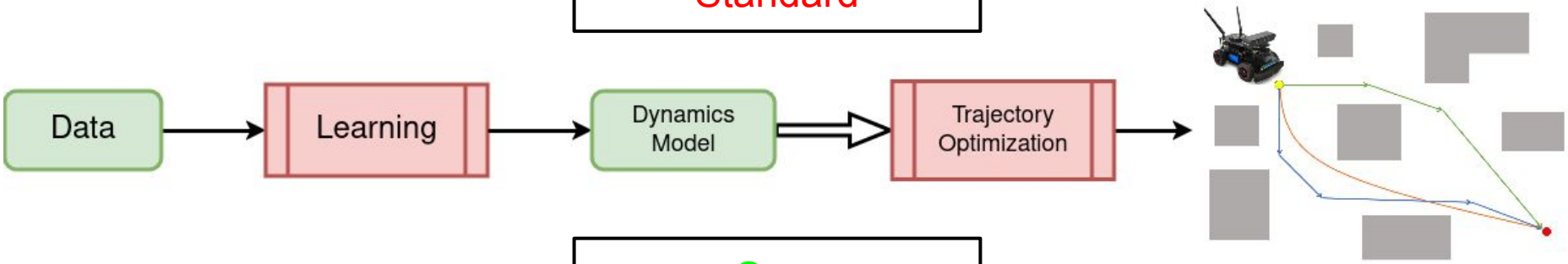


**Planning with Diffusion for Flexible Behavior Synthesis**

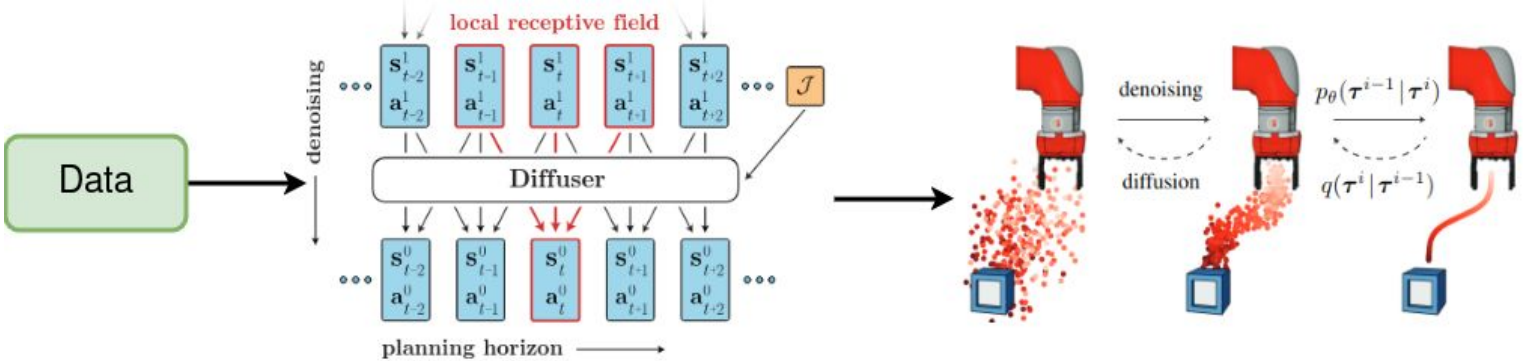
Michael Janner [*1] Yilun Du [*2] Joshua B. Tenenbaum [2] Sergey Levine [1]

# Motivation and Definition of the Problem

Standard

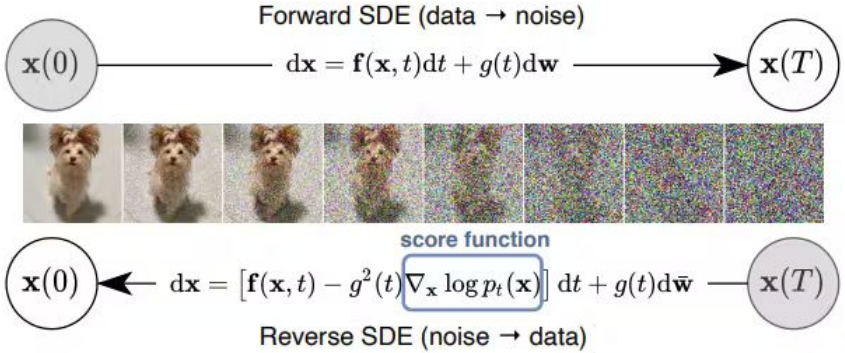



Ours

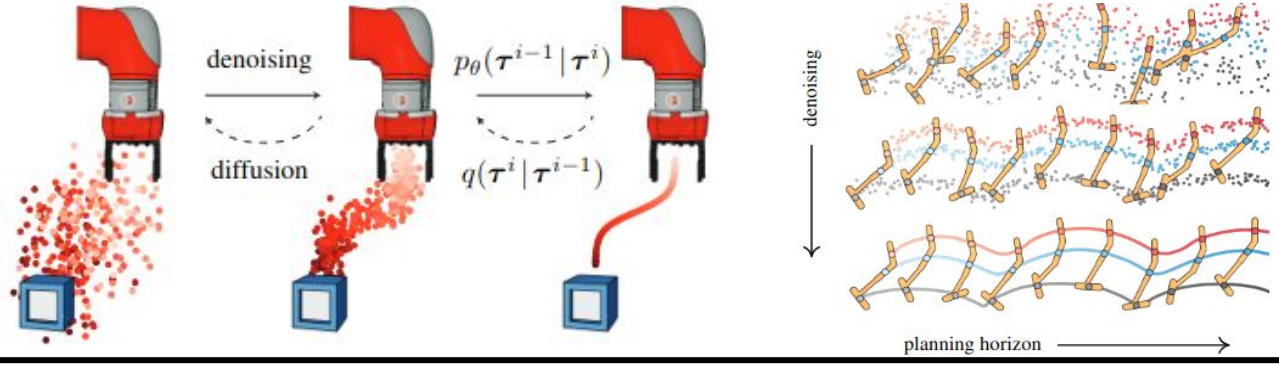


# Putting Prior Work into Context - Diffusion Models

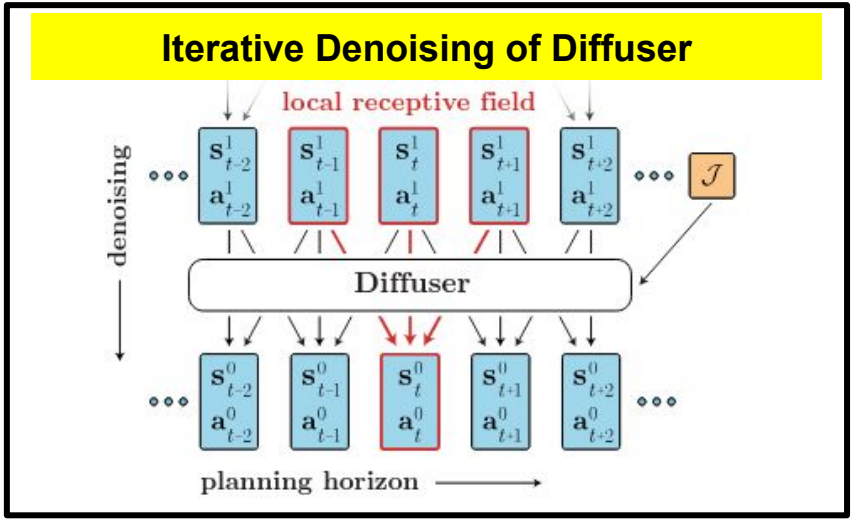
## Standard Diffusion Model

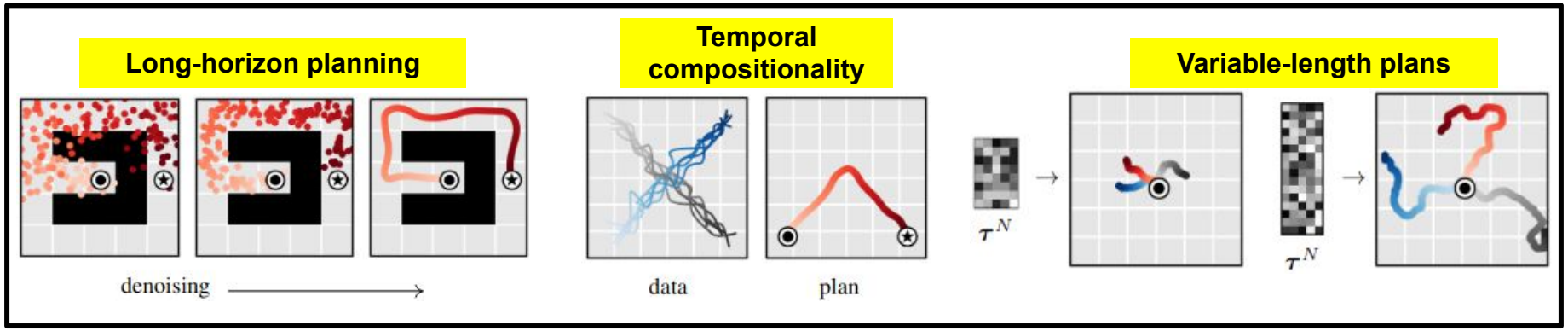



## Diffusion Process in Diffuser Paper



# Putting Prior Work into Context - Diffuser for trajectory planning



# Scoping - Goal Conditioned RL

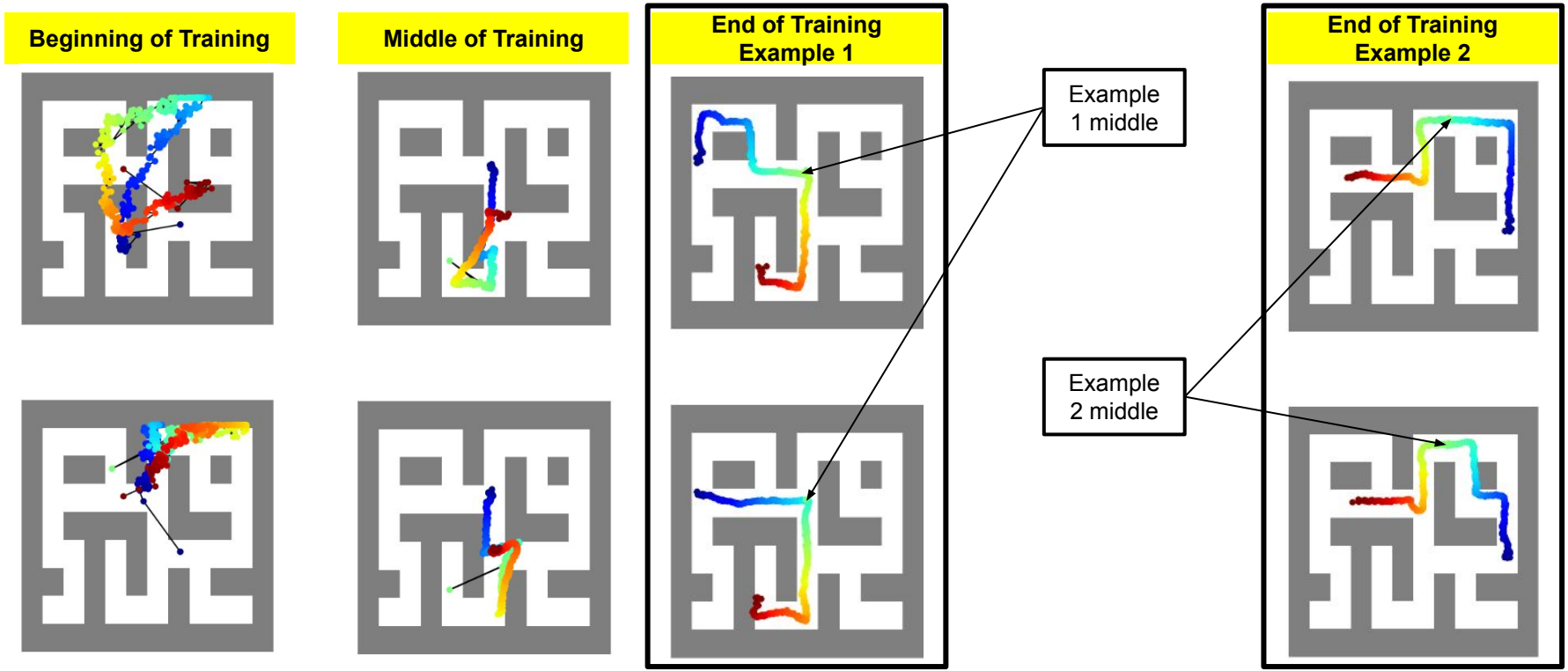
**Goal Conditioned RL is similar to Image Inpainting**

**Current Method of Conditioning simply replaces predicted with correct state**

**Algorithm 1** Guided Diffusion Planning

**Require** Diffuser $\mu_\theta$, guide $\mathcal{J}$, scale $\alpha$, covariances $\Sigma^i$
**while** not done **do**
  Observe state s; initialize plan $\boldsymbol{\tau}^N \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  **for** $i = N, \ldots, 1$ **do**
    // parameters of reverse transition
    $\mu \leftarrow \mu_\theta(\boldsymbol{\tau}^i)$
    // guide using gradients of return
    $\boldsymbol{\tau}^{i-1} \sim \mathcal{N}(\mu + \alpha\Sigma\nabla\mathcal{J}(\mu), \Sigma^i)$
    // constrain first state of plan
    $\boldsymbol{\tau}^{i-1}_{s_0} \leftarrow \boldsymbol{s}$
  Execute first action of plan $\boldsymbol{\tau}^0_{\mathbf{a}_0}$

# Experiment: Train Model to also condition on middle of trajectory



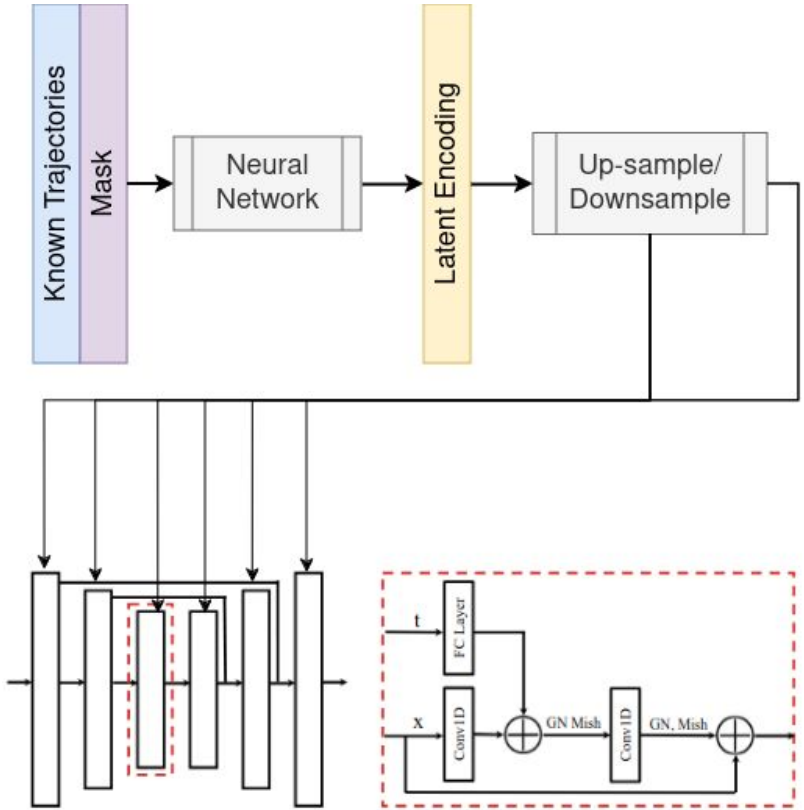
# Method: Changing conditioning mechanism

**Before**

**Algorithm 1** Guided Diffusion Planning

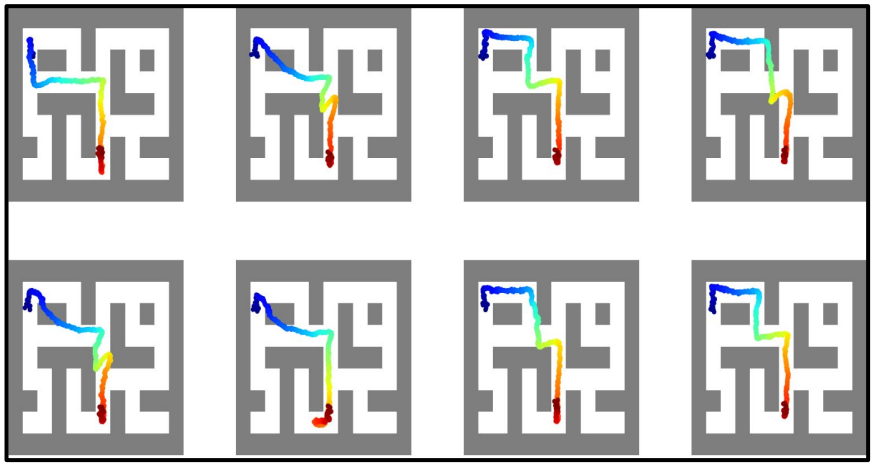
**Require** Diffuser $\mu_\theta$, guide $\mathcal{J}$, scale $\alpha$, covariances $\Sigma^i$
**while** not done **do**
  Observe state $\mathbf{s}$; initialize plan $\boldsymbol{\tau}^N \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  **for** $i = N, \ldots, 1$ **do**
    // parameters of reverse transition
    $\mu \leftarrow \mu_\theta(\boldsymbol{\tau}^i)$
    // guide using gradients of return
    $\boldsymbol{\tau}^{i-1} \sim \mathcal{N}(\mu + \alpha\Sigma\nabla\mathcal{J}(\mu), \Sigma^i)$
    // constrain first state of plan
    $\boldsymbol{\tau}^{i-1}_{\mathbf{s}_0} \leftarrow \boldsymbol{s}$
  Execute first action of plan $\boldsymbol{\tau}^0_{\mathbf{a}_0}$

**After**

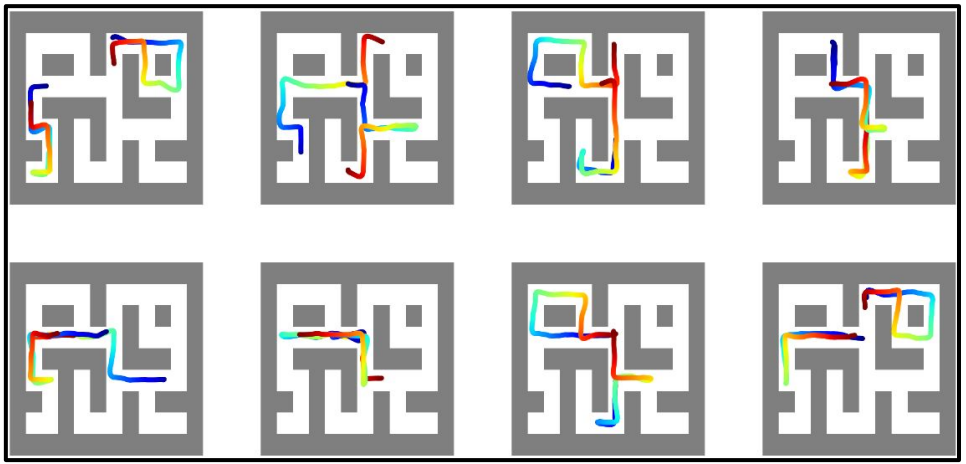


# Experiments

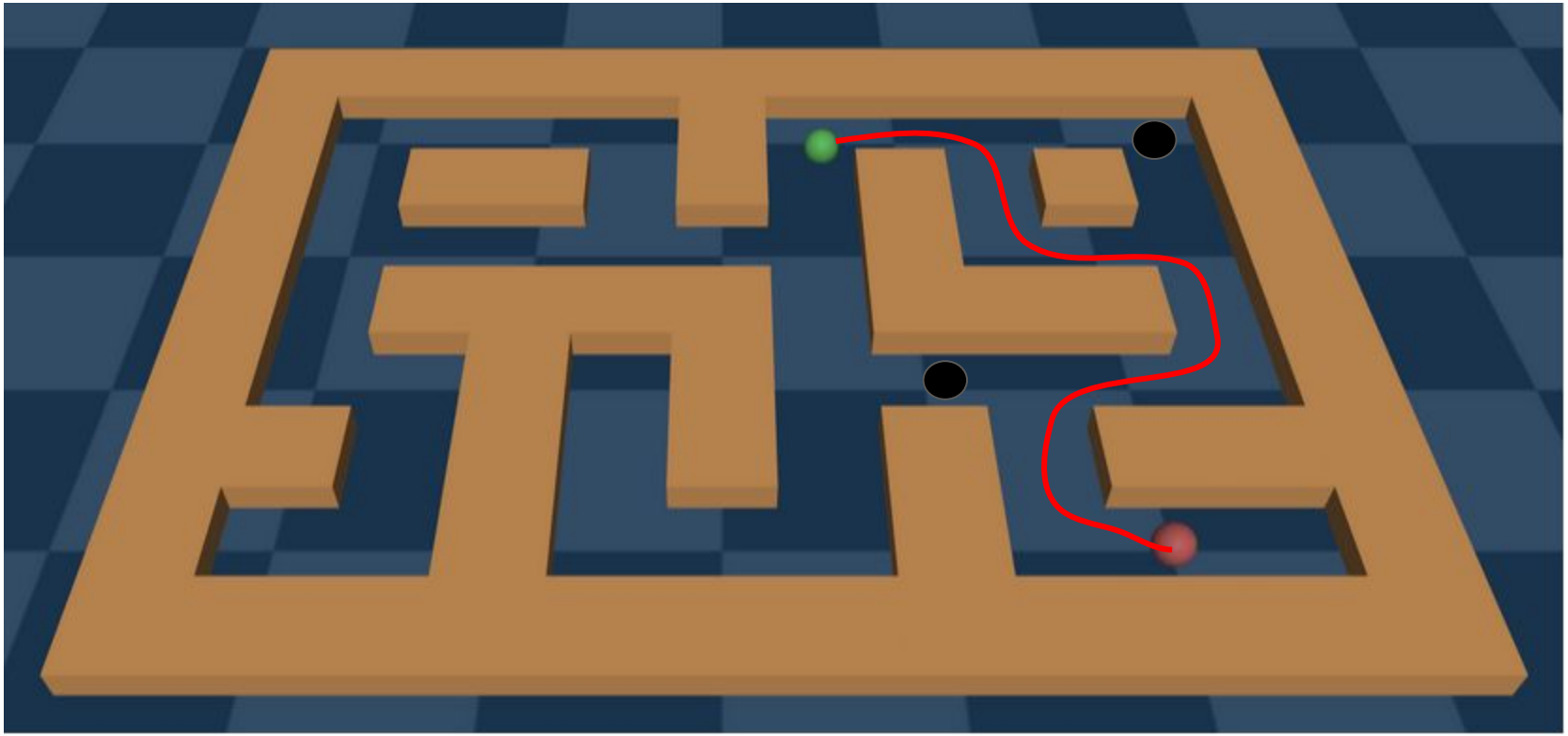
**Single Agent with new Conditioning Strategy**



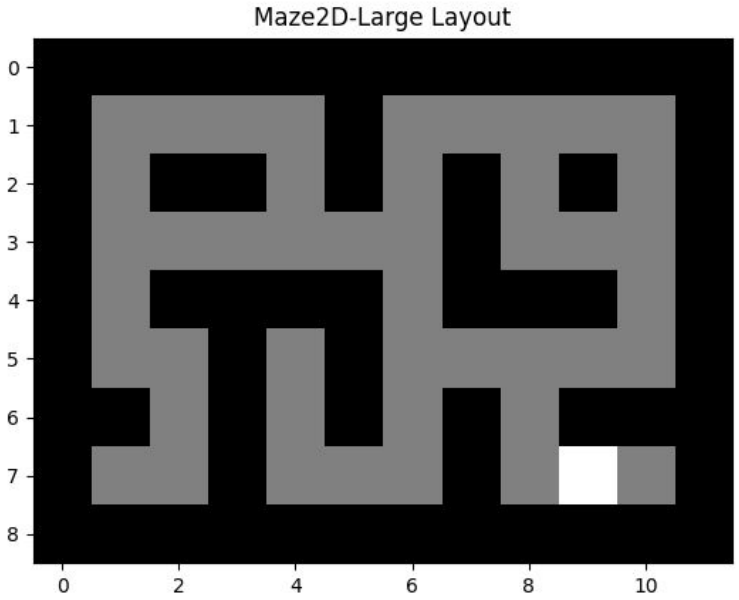
**Multi Agent with new Conditioning Strategy**



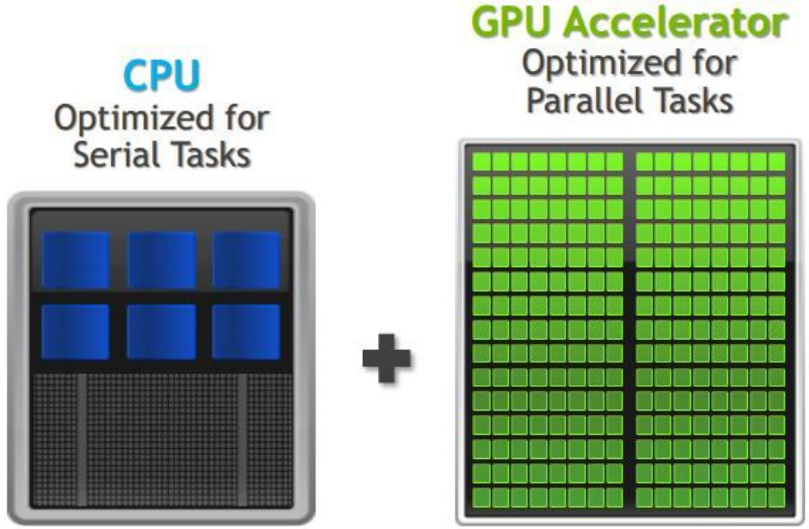
# In Progress - Obstacle Avoidance!



# Limitations




Static maze - will not generalize




GPU Compute Limited (NVIDIA T1200 Laptop GPU)