Multimodal Single Cell Data

$M^{(1)}$

$M^{(k)}$

$M^{(k)}, E^{(k)}$ Modal data and encoder

Inter-Cell Distances

i $D_{ij}^{(1)}$ j $D_{ij}^{(k)}$

$D^{(k)}, D_{ij}$ Distance for modality and latent space

Environmental Reward

$\vec{v}_i$ $\vec{x}_i$ D_{ij} $\vec{v}_j$ $\vec{x}_j$

Boundary Penalty

$$[R_B^t]_i = \begin{cases} -1 & \max(\vec{x}_i) = 1 \\ 0 & \text{else} \end{cases},$$

Velocity Penalty

$$[R_V^t]_i = -||\vec{v}_i^t||^2,$$

Action Penalty

$$[R_A^t]_i = -||\Delta\vec{v}_i^t||^2,$$

Distance Reward

$$[R_D^{t+1}]_i = \delta_i^{t+1} - \delta_i^t, \quad \delta_i^t = \sum_{k=1}^{n_m} \sum_{j=1 \neq i}^{n_c} (D_{ij}^{(k)} - D_{ij}^t)^2,$$

Total Reward

$$R^t = R_D^t + R_B^t + R_V^t + R_A^t$$

$\vec{x}_i, \vec{v}_i$ Cell position and velocity

$\vec{X}^{t_0}, \vec{V}^{t_0}$ R^{t_0}

$t = t_0$ $t = t_1$ $t = t_2$

Random Initialization

$\vec{X}^{t_0}, \vec{V}^{t_0}$ $\Delta\vec{V}^{t_0}$

Per-Cell Action Loop

$\vec{v}_i$ $\vec{x}_i$ $\vec{v}_j$ $\vec{x}_j$

$$\vec{a}_i = E_a\left(\left[\vec{x}_i, \vec{v}_i, E^{(1)}(M_i^{(1)}), \ldots, E^{(k)}(M_i^{(k)})\right]\right)$$

$$\vec{b}_{ij} = E_b\left(\left[\vec{x}_j, \vec{v}_j, E^{(1)}(M_j^{(1)}), \ldots, E^{(k)}(M_j^{(k)}), \vec{a}_i\right]\right)$$

Residual Self-Attention

mean $\vec{s}_i$ E_s $\Delta\vec{v}_i$

$\vec{a}_i, \vec{b}_{ij}$ Cell and neighbor embeddings

$\vec{s}_i, \Delta\vec{v}_i$ State and action vectors

E_a, E_b, E_s Cell, neighbor, and state encoders

π

Development and Disease Trajectory Reconstruction

Perturbation Analysis and Feature Prioritization

π Trained policy