Single-Cell Environment and Proximal Trajectory Inference using Collaborative Agent Reinforcement

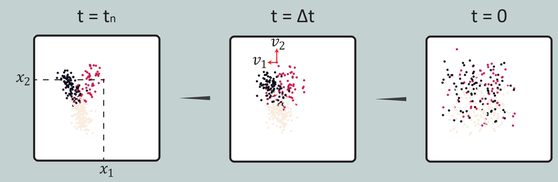
Noah Cohen Kalafut[1,2] and Daifeng Wang[1,2,3,*]

[1]Department of Computer Sciences, Wisconsin, US
[2]Waisman Center, University of Wisconsin-Madison, Wisconsin, US
[3]Department of Biostatistics and Medical Informatics, Wisconsin, US

INTRODUCTION

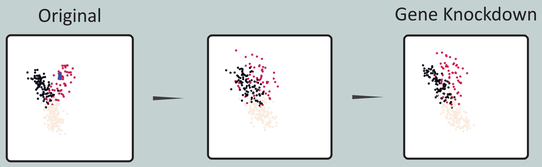
Recent techniques enable functional characterization of single-cells, which allows the study of cellular and molecular mechanisms in complex biological processes including cell development. Many methods have been developed to utilize these single-cell datasets to reveal cell developmental trajectories such as dimensionality reduction and pseudotime. However, these methods generally produce static snapshots of the data, challenging a deeper understanding of the mechanistic dynamics underlying cell development. To address this, we have developed scEPTIC-RL (single-cell Environment and Proximal Trajectory Inference using Collaborative Reinforcement Learning), a multi-agent reinforcement learning model to recapitulate the dynamic progression of cells during development. scEPTIC-RL takes single-cell data, either single or multimodality, and trains a collaborative reinforcement learning model that governs cell-cell dynamic interactions driving development. Particularly, it models single cells as individual agents which coordinate progression on a latent space through interacting with neighboring cells. The trained model can further prioritize cellular features and in-silico predict the dependencies of cell development from feature perturbations (e.g., gene knockdown). We apply scEPTIC-RL to both simulation and real-word single-cell multiomics datasets including brain development and cancers, revealing potential novel mechanistic insights on gene expression and regulation in those complex developmental processes.

RESULTS

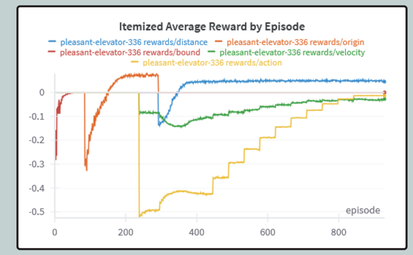


While provided no phenotype information, scEPTICAL was able to separate cell types in a 2-dimensional latent space. Each cell acts individually, allowing for more in-depth downstream analyses.

From the integrated latent space, cells were permuted to the position of progenitor cell types. Model iterations were mapped to timeframes using annotations from the developmental dataset



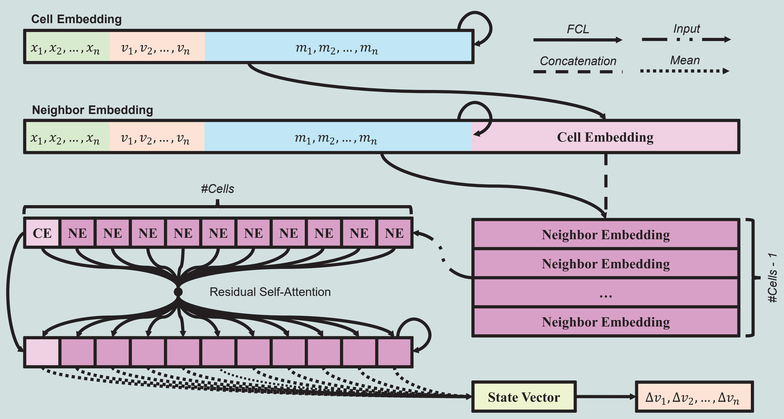
The flexibility of the model allows for modal representations to be permuted (e.g. gene knockdown) and the resultant changes of the latent space may be analyzed for significance.



Further, several accuracy statistics (i.e. XXX, XXX) show scEPTICAL outperforms XXX for integration, XXX for trajectory prediction, and correctly identifies more important genes than XXX from raw gene expression data...

METHODS

$$L_{\text{CLIP}}(\theta) = \text{MIN}\left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}\hat{A}_t, \text{CLIP}\left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}, 1-\epsilon, 1+\epsilon\right)\hat{A}_t\right)$$



In this environment, each single-cell has a position and velocity. The objective of the environment is to emulate the inter-cell distances demonstrated in each modality within the defined dimensions. The central objective here is the distance reward, which increases as the inter-cell distances more closely approximate those of the real modalities.

$$R_D^{t+1} = D^{t+1} - D^t, \quad D^t = \sum_{i=1}^{m}(\text{dist}(M_i) - \text{dist}(X^t))^2$$

In addition, several penalties are added for stability. In order, these are the boundary, velocity, and action penalties. The boundary penalty adds a fixed penalty if the cell touches a simulation boundary. The velocity penalty increases with velocity, incentivizing slower movement. The action penalty increases with the magnitude of actions from the policy, rewarding smoother cell movement.

$$[R_B^t]_i = \begin{cases} -1 & |x_j| = 1 \text{ for some } j \\ 0 & \text{else} \end{cases}, \quad R_V^t = -\sum_{i=1}^{n_d}(v_i^t)^2, \quad R_A^t = -\sum_{i=1}^{n_d}(a_i^t)^2$$

These four rewards and penalties make a stable environment for single-cell integration with any number of modalities. Additional hyperparameters, such as reward and penalty scaling, may be utilized to fine-tune the resulting policies.

CONCLUSION

- ★ scEPTIC-RL is able to effectively separate cells by phenotype, such as cell type
- ★ scEPTIC-RL reliably reconstructs cell developmental and disease progression trajectories
- ★ scEPTIC-RL consistently prioritizes genes and cell features important to development and phenotype

Overall, scEPTIC-RL provides reliable trajectory reconstructions for cell progression and disease development in an easily distributable and highly generalizable manner.

[1] Jie Liu, Yuanhao Huang, Ritambhara Singh, Jean-Philippe Vert, and William Stafford Noble. "Jointly Embedding Multiple Single-Cell Omics Measurements". en. In: Algorithms Bioinform 143 (Sept. 2019).
[2] Alexandro E. Trevino, Fabian M˙uller, Jimena Andersen, Laksshman Sundaram, Arwa Kathiria, Anna Shcherbina, Kyle Farh, Howard Y. Chang, Anca M. Pas, ca, Anshul Kundaje, Sergiu P. Pas, ca, and William J. Greenleaf. "Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution". In: Cell 184.19 (Sept. 2021), 50535069.e23. issn: 0092-8674. doi: 10 . 1016 / j . cell . 2021 . 07 . 039. url: https://doi.org/10.1016/j.cell.2021.07.039.
[3] Zhi-Jie Cao and Ge Gao. "Multi-omics single-cell data integration and regulatory inference with graph-linked embedding". In: Nature Biotechnology 40.10 (Oct. 2022), pp. 1458–1466. issn: 1546-1696. doi: 10.1038/s41587022-01284-4. url: https://doi.org/10.1038/s41587-022-01284-4.