Shankara Anand
Varun Gupta
CS221 Project Proposal

Finding Cell Differentiation Pathways With Reinforcement Learning

One of biology's most complex and ubiquitous phenomena is early stem cell fate differentiation. Cell fate is mediated by an interplay of both environmental factors and intracellular components, particularly gene transcription factors.[1] Previously, scientists have approached cell differentiation through bulk assays, which provided insight on average cell population characteristics. Today, microfluidic high-throughput qRT-PCR has enabled the collection of single cell gene expression data.[2] Statistical analysis of such data provides insights into interactions among the factors and how the factors determine cellular fate. For example, unsupervised analysis of mouse embryonic stem cells suggests regulatory networks in early stage blood-cell development.[3,4] We will attempt to further this analysis by building a stochastic, cell differentiation pathway using the same single cell gene expression data and reinforcement learning.

Task Definition

A single input for our system is a single cell gene expression dataset. For example, we consider the mouse embryonic dataset collected and studied by Moignard et al.: 3,934 cells "with blood-forming potential captured at four time points between E7.0 and E8.5" and their levels of expression for 46 different genes. In addition to gene expression profiles, cells in the input datasets are each accompanied with label that denotes the cell's maturity stage or stem cell type. In the mouse embryonic dataset, there are five such classifications. Three classifications indicate the three ordinal stages of development: primitive streak (PS), neural plate (NP), and head fold (HF). The fourth and fifth classifications represent two different cellular fates and thus differentiation that HF cells undergo: 4SG (erythroid fate) and 4SFG (endothelial fate).

A single output for our system is the following: the optimal differentiation policy for the cells. Thus, given a cell with a certain initial distribution (initial gene expression profile), our model should be able to predict the cellular fate of this cell. In the context of the mouse embryonic dataset, our model should be able to predict the cellular fate of any PS, NP, or HF cell (whether that fate is endothelial or erythroid in nature).

Model Approach

The goal and scope of this project is the following: given a set of single cell expression data, build and tune a stochastic process that models each cell as a state, where state is composed of the continuous gene expression profile and cell label. For instance, the process built from the the mouse embryonic dataset would have 3,934 states, with each state having 47 components: expression levels of 46 genes and a classification label of PS, NP, HF, 4SG, or 4SFG. Tuning the process involves tuning rewards associated with transitioning between states; higher rewards can be associated with maturing to the next label. Terminating states are those cells that represent fates in the dataset. We can use reinforcement learning to find an optimal policy and differentiation model. For any given any starting state in the model (and thus an initial gene expression profile and classification), the model determines the transitions that take it from state to state and finally a terminating state, which indicates a cellular fate. For example, we could examine any particular NP cell and see how such a cell would mature according to the optimum policy.

We approach this problem assuming a level of randomness as suggested by the widely-held "equivalence hypothesis" of cell differentiation.[5] This hypothesis suggests that individual cells in early embryos are

homogenous, and that any noticable differences among these cells are noise which may eventually shift towards finite numbers of fixed, differentiated cell types.[5] In other words, some probability distribution characterizes the initial distribution of possible gene expression profiles of the earliest progenitor cells; the initial gene expression profile of any particular cell likely pushes it stochastically towards a particular cellular fate, thus representing a dynamic system where initial conditions matter. Therefore, our proposed methodology of using Markov decision processes is reasonable.

Evaluation

We will validate the predicted optimal policies and thus predicted cellular fates against published transcription factors and transitions that have been elucidated as crucial for determining cell fate. For example, in one study, Moignard et al. used transgenic assays to confirm the particular interaction of transcription factors they found through bioinformatic techniques.[3] In other study, Moignard et al. used their single cell gene expression analysis to confirm known factors that control endothelial fates.[4] We measure accuracy by how well an optimal policy (differentation model) involves known regulatory pathway transitions.

Baseline and Oracle

The baseline implementation uses the mouse embryonic dataset. It calculates the centroids of the cells that have already differentiated (4SG and 4SFG), and then assigns the remaining progenitor cells (PS, NP, and HF) to the closest centroid. We use Euclidean distances based on differences in gene expression profiles (thus, 46 dimensions). We report two errors: the average distance between each progenitor and the centroid to which it was assigned only in terms of the genes that are known to determine that particular fate, and the average distance between each progenitor and the centroid to which it was not assigned in the opposite terms. For example, for a NP cell assigned to an endothelial (4SFG) fate, we calculated the first error as the distance between the cell and the endothelial (4SFG) centroid in terms of Cdh5, Erg, HoxB4, Sox7, and Sox17 factors, and the second error as the distance between the same cell and the erythroid centroid (4SG) centroid in terms of Gata1, Gfi1b, HbbbH1, Ikaros, Myb, and Nfe2 factors. The baseline finds that the first error is significantly larger than the second, indicating that classification based on distance provides little insight into the cellular transitions for cells that are squarely in-between either centroid.

Ideally, biological assays could mark each progenitor stem cell, measure its gene expression profile (and maturity) every time step, and test its final cellular fate. With measurements of many such stem cells, a frequentist oracle could calculate the optimum policy as the expectation over these randomly-realized differentiation pathways. We approximate this situation in our model by assuming that each measured cell in our datasets represents a possible state.

Conclusion

The gap between the oracle and baseline implementation is the identification of what exactly is meaningful in pushing cells from state to state. Additionally, the baseline assumes nothing about the life cycle or time point in which our cells are sampled for gene expression data. The genetic makeup of cells may be wildly different due to different needs in their development cycles. On the other hand, the oracle is unrealistic to reach by artificial intelligence standards, due to the complex range of factors that may play a role in cell differentiation and the stochasticity of biology. The use of reinforcement learning given rewards that can be tuned may provide additional benefit over the baseline implementation, especially since we currently are not proposing any metric for transition probabilities between our various states.

Sources

[1] Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663–676 (2006)
[2] Citri, A., Pang, Z. P., Sudhof, T. C., Wernig, M. & Malenka, R. C. Comprehensive qPCR profiling of gene expression in single neuronal cells. Nat. Protoc. 7, 118–127 (2012)
[3] Moignard, Victoria, et al. "Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis." Nature cell biology 15.4 (2013): 363-372.
[4] Moignard, Victoria, et al. "Decoding the regulatory network of early blood development from single-cell gene expression measurements." Nature biotechnology 33.3 (2015): 269-276.
[5] Biase FH, Cao X, Zhong S. Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing. Genome Research. 2014;24(11):1787-1796.

```
## typesort sortA sortB
##
##           3175   759
```

Compute centroids of fate types:

```
fourg <- Nor[celltypes == "4G",]
fourgf <- Nor[celltypes == "4GF",]

fourg_centroid <- colMeans(fourg)
fourgf_centroid <- colMeans(fourgf)
```

Assign each cell to closest centroid:

```
assign <- function(profile) {
  dist_fourg <- sqrt(sum((profile - fourg_centroid)^2))
  dist_fourgf <- sqrt(sum((profile - fourgf_centroid)^2))
  return(if(dist_fourg < dist_fourgf) "4G" else "4GF")
}
assignments = apply(Nor, 1, assign)
```

```
endothelial <- c("Cdh5", "Erg", "HoxB4", "Sox7", "Sox17")
erythroid <- c("Gata1", "Gfi1b", "HbbbH1", "Ikaros", "Myb", "Nfe2")
fourgf_rel <- t(data.frame(fourgf_centroid))
fourg_rel <- t(data.frame(fourg_centroid))

erythroid_distance <- function(profile) {
  return(sqrt(sum((profile - fourgf_rel[,erythroid])^2)))
}
endothelial_distance <- function(profile) {
  return(sqrt(sum((profile - fourg_rel[,endothelial])^2)))
}

erythroid_distances <- apply(Nor[,erythroid], 1, erythroid_distance)
endothelial_distances <- apply(Nor[,endothelial], 1, endothelial_distance)
assignment_error <- (sum(erythroid_distances[assignments == "4G"]) + sum(endothelial_
distances[assignments=="4GF"]))/nrow(Nor)
non_assignment_error <- (sum(erythroid_distances[assignments == "4GF"]) + sum(endothe
lial_distances[assignments == "4G"]))/nrow(Nor)
```

The mean assignment error is 21.1119312, and the mean non-assignment error is 8.2270436. Ideally, the assignment error should be much smaller than the non-assignment error, which would indicate that a cell is much closer to its assigned centroid in terms of the transcription factors relevant to that fate than to the centroid it was not assigned to in terms of transcription factors relevant to the fate it did not achieve.