

Extending Discrete Morse Theory to Neuron Reconstruction

Ben Holmgren, Ryan Hansen

November 2, 2020

Background

In all, our project can be divided into a series of tasks, as follows:

1. (Main project - analyze algo) To begin, we need to actually analyze and fully understand the TMD algorithm as proposed by Hess et al. in terms of inputs, outputs, and time/space complexity. Specifications of the inputs, outputs and the time complexity of the TMD algorithm are included in Hess' work, but there is no rigorous analysis of the space complexity of the algorithm. We've provided a rigorous statement of all of these items below.
2. (Propose new algo, show does same stuff) The focal point of this project is to demonstrate how discrete morse theory can not only be used as an equivalent route to classifying neurons, but also to argue for its superiority. Step two is basically just the first part - proving we can substitute a filtration and barcode analysis (TMD) with generating a discrete morse function and subsequent analysis. Since Kathryn Hess has already shown classification of neurons can be attributed to branching, and branching occurs at critical points, discrete morse shows promise for classification of neurons!
3. (Space Complexity better, time complexity same) The next step is obviously proof of morse theory superiority. While we think there may be several, the primary advantage for morse theory is space complexity. According to a random google search, the human brain contains nearly 86 million neurons, meaning with regard to simulating real brain function, the ability to represent singular neurons with better space complexity would be a major contribution. Additionally, we need to show that using Morse theory retains the original time complexity of the TMD algorithm to make sure we have only benefits.
4. We also hope to show additional features that using a discrete Morse function allows. We are suspicious of two other advantages morse theory may provide the community at large. First, we want to show that using Morse theory allows for the pinpointing of specific regions within a neuron, which may relate to neuron functionality. Kathryn Hess does do something similar, but we think we are adding something unique. Second, we think searching the morse function data structure is much faster than searching the lengthy chain of branches utilized in the TMD algorithm, which we are suspicious could also be meaningful..... who knows what kind of buried treasure lies within.....

Progress

Thus far, we've successfully accomplished tasks 1-3. Our progress for each is loosely as follows:

1. The algorithm’s input, output, and runtime are clearly expressed on page 7 of Hess et al. For an analysis of the runtime, we claim that the space complexity of the algorithm is linear with respect to the number of nodes on the neuron. We prove this by considering the number of nodes in an active branch of the neuron. At worst, this is equivalent to the total number of nodes. As a lower bound, there could possibly be only one node in an active branch. But, no matter what, there must be a maximal active branch with a considerable fraction of the nodes. Meaning that the number of active nodes has a linear relationship to the total number of nodes, and the space complexity of the TMD algorithm is order of n .
2. The filtration specified in the TMD algorithm essentially uses a filtration to find the persistence of branches throughout the neuron. To generate an equivalent Morse function, we first need to define underlying vertex data present in the image of a neuron. Letting N denote the set of nodes in the neuron image, we define an underlying function $f : N \rightarrow \mathbb{R}$ where f defines the path-based distance from a source node to any other given node throughout the neuron. Then, we’re able to assign a modified discrete Morse function to the complex. The function is modified in that critical cells are defined by any vertex which is larger than degree 2. (Because the complex is guaranteed to be an acyclic graph, we know that only critical vertices are possible). Then we gain matchings from a Morse function revealing of the flow into any critical cell in the complex. In doing so, we gain branchwise distance in the number of collapses needed to reach critical cells, and in doing so gain a metric of persistence in each individual branch. Thus, we gain all of the features given by the filtration in the TMD algorithm in using a modified Morse function.
3. From the proposed setup, we gain that we can generate a discrete Morse function in linear time with the underlying vertex data. Notably, this discrete Morse function is guaranteed to be “high performing” in that the critical cells are guaranteed to be directly reflective of the areas where the neuron branches, since we guarantee that the underlying vertex data used by the algorithm proposed by Holmgren et al. is increasing along the path from the source to any given node, and that the function f is injective. And, since the underlying vertex data is also guaranteed to be sorted radially along the neuron, we are able to improve the functionality of Holmgren et al. Indeed, the runtime in Holmgren et al. is guaranteed to remain linear with respect to the simplices of the neuron (i.e. roughly double the number of vertices for a graph), so the TMD algorithm has equivalent runtime. We note the logic behind morse theory feasibility is contingent on critical point being equivalent to branching. The proof of our chosen morse function to be ‘high performing’ morse function is a trivial \iff proof we will go into more detail in our formal writeup. However, the algorithm by Holmgren et al. when adapted to a lexicographically sorted Hasse diagram of the underlying neuron is able to operate with constant space complexity. Improving the space complexity of the TMD algorithm. As we mentioned earlier, this my friends is a game changer.

Remaining Work

As for the remainder of the project, we still need to work on specifying how to collapse simplices in an assigned Morse function to highlight structural differences within different kinds of neurons. We have a rough idea of how to go about this, but need to elaborate fully on the theory behind this process. And, we need to more formally write up our developments thus far, especially our framing of the problem and our provided improvements with regards to its space complexity. The abstract math we were kind of thinking about is we know the morse functions can classify neurons,

meaning they are some kind of invariant on differing neuronal spaces. Our hope, then, is to find some kind of distance between morse functions to further hone in on what specifically differentiates functionality. Kathryn Hess does some neat heat maps, but morse functions could give us much more specific information on key differences. The general consensus of the neuroscience community is nobody knows what parts of the neuron are super meaningful, so we want to explore what morse functions can tell us. Last, we wanted to look into whether or not it is useful to search in the morse function data structure, because as we mentioned earlier, it is much easier than sifting through the memoized string of branches used in the TMD algo.

Works cited

A Topological Representation of Branching Neuronal Morphologies, Hess et al.

If You Must Choose Among Your Children, Pick the Right One, Fasy, Holmgren et al.