Individual Development Plan (June 2022)

Will Dumm

Career Goals

Long-Term and 10-year Goals

In the long term, my goal is to be productive in a flexible, self-directed but collaborative, location-agnostic position contributing as a programmer to an interesting project in science or applied math. This job will be:

- senior-level

- flexible, in the sense that I have some freedom to pursue projects that interest me

- creative, not purely an application of technical skills

- related to a goal or cause that is important to me

My current position offers most of these characteristics already, but I can learn to take advantage of them better. For example, I can use my freedom in my position to be more creative about how I approach my projects. I think this long-term goal is achievable within the next ten years.

Last Year's Achievements

- Learned lots of background about phylogenetics and adjacent topics

- Wrote a Python implementation of the history DAG

- Wrote Parsimony Plateau paper on parsimony in the history DAG

- Rewrote (parts of) `gctree` to utilize history DAG for expanded tree search, more efficient tree computations, and additional ranking criteria.

- Began planning and writing paper describing improvements to gctree.

- Familiarized myself with `Usher` and `matOptimize`, and helped get mutation annotated DAG (MAD) implementation started

- Used `Usher` to verify that parsimony allows considerable tree uncertainty, at least for SARS-CoV-2 data.

This Year's Goals

- (from last year) Becoming a proficient c++ programmer

- (help to) complete implementation of the MAD, integration with `matOptimize`, and parsimony tree search.

- Implementation of (more) useful applications of the history DAG, including

    - a more general tree search motivated by an arbitrary weight function or likelihood, informed by context-sensitive model.
    - implementation in some form of the ideas described in the 'conditional damara distribution' grant application.

- Complete gctree v2 paper, with posterior predictive check, and simulation validation evaluating performance of gctree with hDAG, isotype- and mutability- parsimony compared to branching process likelihood alone.

- Use hDAG to understand the shape of the parsimony plateau, and develop ways to understand the shape of the hDAG in general. This will hopefully be another paper, with William's help.

- Submit hDAG / Parsimony Plateau paper

- Continue managing interns, making sure that tehy are productive and supported, and getting the experience they want, and learning what that takes.

Development of Project-Specific Knowledge

Project Description

My projects for the foreseeable future involve applications of the history-DAG, as described in the goals for this year.

New Skills and Knowledge Gained

- Biology background, including basic immunology and phylogenetics

- Basic understanding of applications of MCMC and variational inference to Bayesian phylogenetics

- Detailed understanding of parsimony and Fitch, Sankoff algorithms

- Re-familiarization with Python, DAG-related algorithms, shell, and cluster use

- Basic understanding of many of the group's projects

- Familiarity with variational inference applied to phylogenetics

- Extremely detailed understanding of `gctree` inference

Skills and Knowledge Needed

The following skills will be needed to make project progress:

- Continued familiarization with c++

- Detailed understanding of the Usher MATUtils and MATOptimize codebase

- Continued progress in understanding the limitations of the history DAG in our applications

- Efficient, productive, and supportive intern management

Career Skills

Strengths

Communication Skills

In my last IDP version, I identified ways in which I could improve my communication skills. There have been a few opportunities since then to present my project progress, for example in group meeting presentations and paper writing. However, I haven't meaningfully implemented my main proposal for improving communication skills, which I still think will be helpful: I should improve my general readiness to explain project ideas without preparation, for example by verbalizing each day a new idea that I had that day.

Strengths and Other Opportunities for Improvement

I expect to improve my time management in this position:

- Keeping sight of immediate goals for each day or subproject, and staying on track

- Knowing better when to shift attention when progress on a task slows

- Having a more complete array of useful tasks in mind, and better recognizing directions with high potential for progress

- (New) Better prioritizing tasks, especially by efficiently delegating appropriate tasks to interns, or deferring them in favor of those with higher immediate value

I can address the first two items by pausing to check in with my goals and progress throughout the day. The third can be helped by continuing to write down any ideas or questions that come up while working on other tasks.

Other opportunities for improvement:

- I would like to better anticipate how I can help collaborators, like Ognian and interns, instead of waiting for them to identify what they need from me. Spending some time carefully thinking through their proposed tasks and making notes about key ideas and potential pitfalls will help with this.

- I would like to take more time to think broadly about projects, brainstorm new project ideas, and communicating them to someone instead of letting them languish in my notes.

- Finally, although code testing is already important to me, I'm not satisfied with the number of issues caused by my code this year. I should force myself to spend more time testing even fairly informal scripts.

Opportunities for Contacts and Collaboration

I'm looking forward to continuing to work with interns. I also hope to learn a lot from Ognian as he continues to implement the MAD, and from the Usher team as we integrate `larch` and `matOptimize`.