The Allard Lab Manual

Jun

Apr 08, 2022

CONTENTS:

This is a continuing work in progress. The source file is on github at our allardlab repo. You are welcome to fork and pull-request improvements. There is also a compiled latex version.

CHAPTER

ONE

MISSION STATEMENT

1.1 Peer reviewing

Virtue: When we peer-review, we are constructive and we try to make their paper better.

For me (Jun), peer-reviewing is the scientist's version of a musician practicing scales. This is the action that makes me improve the most. It hones the skill of analyzing the technical approach, recognizing a gap, judging and improving rigor, and evaluating impact.

After a Nature editorial came out on the role of trainees in peer review, we adopt the "Bovyn rule": We always have a synchronous meeting to discuss the manuscript (not just asynchronous discussion).

1.2 The mission of our research group

The mission of our research group is to make *rigorous, relevant and elegant* contributions to the scientific knowledge of humanity. We use computational, mathematical and biophysical approaches to figure out *how living cells use force, space and time* in their *problem-solving strategies*. We work so that the basic science discoveries we make become part of the *worldwide, multi-generational tapestry of scientific knowledge* that benefits all people.

It may sound rote, but every part of this 3-sentence paragraph has operational meaning.

- rigorous
- relevant: addresses a gaps in knowledge
- force, space and time: physics, mechanics, stat mech (entropy, diffusion)
- cellular problem solving strategies
- worldwide, multi-generational tapestry of scientific knowledge

1.3 Value-virtue pairs

To achieve our mission, we adopt certain principles, or "virtues". Every virtue comes from a value.

- **Value: We try to make other people's science better.**
  - Virtue: Peer Reviewing See this *Peer reviewing* section.
  - Virtue: We *memorialize* our own stuff, both for internal future use and possible external future use. See Dev Ops section

- **Value: We value time, the most scarce and precious resource.**
  - Virtue: Being on-time.
  - Jun's Rule If Jun is late, Jun agrees to put $1/min/attendee into an Allard Lab discretionary fund.
- **Value: We make other people want to work with us.**
  - Virtue: Being a good listener, being excessively respectful of others. There are essential policies that have to do with respecting our colleagues, led by the wider scientific community, the UCI community, and our other affiliated communities. Besides meeting those standards, we try to go beyond them.
- Value: We think our projects are cool, in addition to useful. - *[Jun Question for Team] Virtue: Should we re-start Twitter? More fun visuals.*

CHAPTER

TWO

ELEMENTS OF A SCIENTIFIC CONTRIBUTION

2.1 Jun's paper writing tips and trick

2.1.1 How I write papers

0. Find some role-model papers, just to get inspiration for how to arrange figures, expose statements, organize subsections. Ideally the role-model paper is from a similar journal to the one we are targeting.
1. Outline the story: What are the Results headings and all the subfigures that tell this story? Sketch abstract.
2. Subfigures in Matlab .eps files
3. Assemble figures, write captions. Nice meaty captions: A good paper can be followed by reading captions alone (without reading Main Text).

Now relax. Lots of work ahead, but if you did 1-3 correctly, the rest is like a ball rolling down a hill.

4. Write Results text, Methods/Model text, Supplement. A good paper can be followed by reading Results text alone (without looking at figs).
5. Big literature review, spend a few days, re-read ~15 papers, search for anything we missed.
6. Write Discussion text. Connect to the field. Rank order paragraphs from most important to least important. Be upfront, but not apologetic, about limitations of work.
7. Write Intro text. The main purpose of Intro is to describe the gap in knowledge. Writing such a short overview of the field requires you to have an opinion of the field – mark of scientific maturity.
8. Assemble, polish. Clean up references. Give to lots of people for feedback. Get ideas for suggested referees.
9. Draft cover letter to the handling editor

- Omer's 4 key questions to answer in a cover letter: (1) How will this work make others think differently and move the field forward? (2) How does our work relate to current literature? (3) Who is the most relevant audience for the work? (4) What has the work accomplished and what has it not achieved?

Now relax. Measured in "wall clock" time, you might be halfway. The next step requires a stomach: *receiving peer review*.

2.1.2 Jun's Revision process

My workflow for resubmissions is as follows. It's kind of cumbersome but I find it works.

- I give every Reviewer comment a code, like "Rev1Minor3".
- I create THREE google docs in a folder here.
  - A Big Notes doc, with the Reviewer comments, todo lists / Action Items, paragraph drafts for the response letter, paragraph drafts for the new manuscript, and other notes.
  - A "Dashboard" doc, with very short summary of each comment, so we can check them off as we go – like a progress dashboard.
  - Another doc for the actual careful response letter. This is mostly blank for now, but then can be made quickly following cut-and-paste from Big Notes, and then edited carefully, especially for tone, which is easy to get wrong.

2.1.3 Writing tips

- *cut cut cut cut cut*
- One paper should have one main message. Repeat main message 5 times: In Title, Abstract, Intro, Results, Discussion. (In rare cases, 2 main messages.) All other messages go in Results and Discussion
- *cut cut cut cut cut*
- Writing a paper is a continuum of fact-reporting ("we did this", "we observed this") to interpretation ("this suggests that…", "this might have implications for…"). Caption is most factual, followed by Results text, followed by figure caption first sentence and Result headings. Discussion is most interpretation. For some historical reason, the figure caption's first sentence is the interpretation, while the subfigure sentences are factual.
- Savage and Yeh "Novelist Cormac McCarthy's tips on how to write a great science paper, Nature 2019 I disagree with like half of these.

2.2 Being an effective scientific communicator

We spend a lot of time on communication, optimizing presentations, poster, paper figures. This is for a few reasons.

1. Clear thinking means clear communicating, and clear communicating is clear thinking. There is almost never a distinction in effort.
2. Science gains from network effects, but only if each lab/scientist devotes effort to communicating and coordinating. Like a high-performance computer devoting some of its cpu time to parallelization.

2.3 The elements

There is a standard layout – the time-tested most powerful layout – for a scientific contribution that works particularly well.

- Big question/background
- Specific question/hypothesis – what you will deliver in this work
- "Here we"
- Method

- Results
- Impact

Here are some examples of paper abstracts and layouts:

- What paper elements? `Slides with examples of paper layouts`.

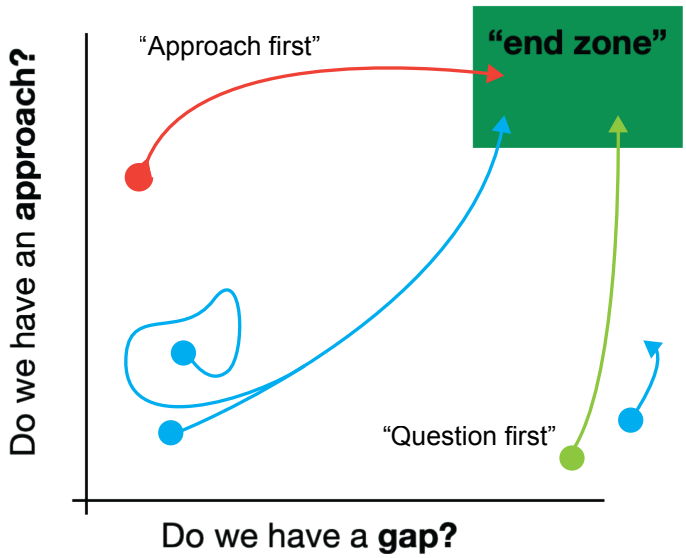
2.4 Approach and gap are of equal importance

"When I sit down with colleagues over a beer at a meeting, we don't go over the facts, we don't talk about what's known; we talk about what we'd like to figure out, about what needs to be done… This crucial element in science was being left out for the students." - Stuart Firestein (Columbia Neuroscience)

Doing science involves two activities of equal importance:

(1) Identifying the gap in current knowledge/needs

(2) Modeling/experiment approach and carrying out that approach

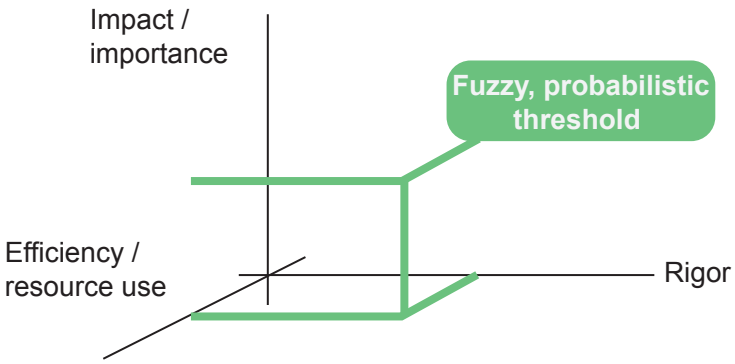


An analogy is "product-market fit".

It's ok to have one before you have the other (exploratory projects, "approach first" projects, "question first" projects,…) and it almost always happens that the approach and gap change throughout a project.

2.5 The importance rigor resources trade-off

A useful coordinate system to "project" (verb) success as a project progresses:



(1) Importance/impact of the result: This importance criteria is often the easiest to retreat on, because there is an attitude in the scientific community that scientists are bad at predicting future importance anyway.

(2) Rigor/completeness. How much did we explore all the "nooks and crannies", eg parameter sweeps? How much did we reject alternative possibilities? How much did we test the tools?

(3) Efficiency/use of resources/speed: Are we going to get this paper done in time, so people can move on to even bigger and brighter things!

There is a fuzzy, probabilistic minimal threshold for each of these.

Above the threshold, there is a trade-off between them. Teams benefit from discussing this trade-off.

Don't underestimate the value of speed! There are many exciting things you can get on board, but "The bus leaves at 9:00", meaning there is a specific point in time you need to hit a milestone. Don't fear these things!

2.6 Figures!

- What makes a great figure? `Slides about figures`.

Upshot: You need the ability to make a figure that is publication grade, informative, and digestible. *Not a pixel out of place.*

The "Masterful Inaction" principle: it's ok to make a quick figure, as long as you're doing it because you have the ability to make a publication-quality figure and have deliberately made the decision not to.

2.7 Paper writing

The *paper writing process*

CHAPTER

THREE

MEMORIALIZATION AND THE DEVOPS STRATEGY

3.1 DevOps for Systems Biologists

- What is "DevOps"? `DevOps keynote slides`.

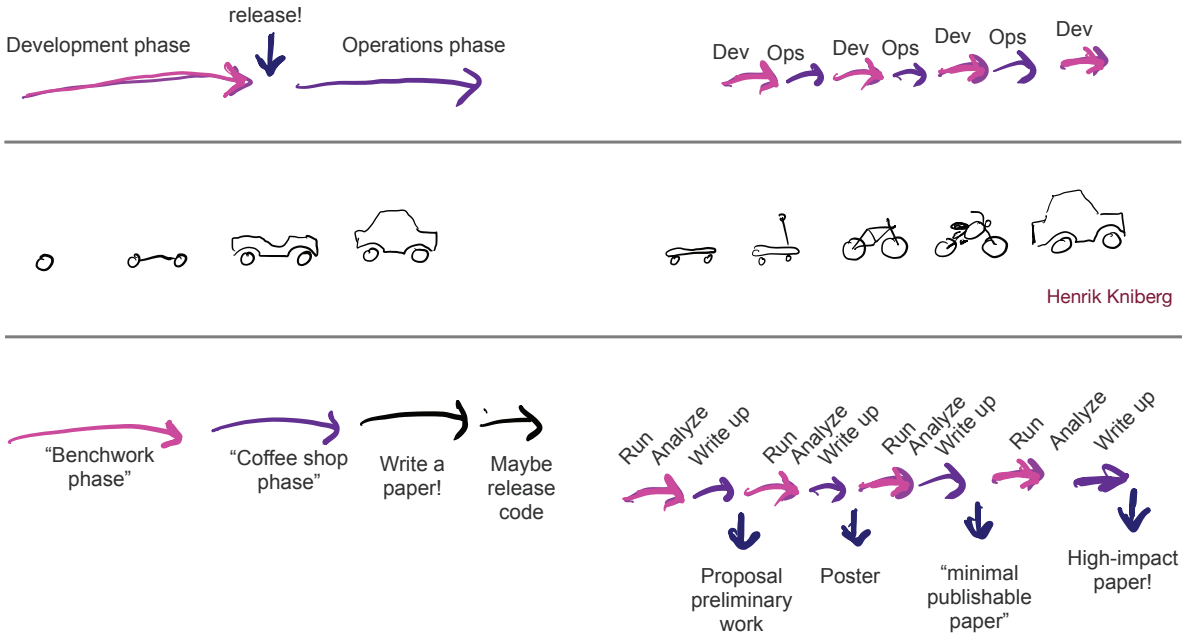
3.2 Upshot

In order to…

- …make our contributions make other people's science better (within our group, and other labs)
- …keep our own peace of mind!
- …foster healthy, fun and productive collaborations
- …maximize our own efficiency

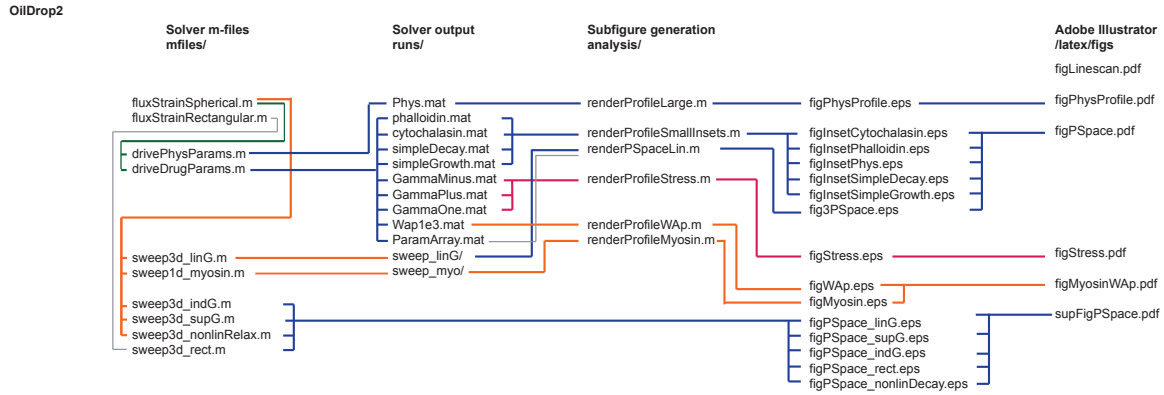
version-control can be used smoothly for entire projects (analysis and write-ups, not just code).

This allows you to use a small-batch, beginning-to-end, continuous-improvement "DevOps" approach (also known as "first, build a bike"). This approach has pros and cons.



3.3 Internal resources for DevOps performance

- A project workflow file with functions, scripts, input and output. The program VSCode with draw.io plugin is excellent for this. "Diagrams are a great way to communicate ideas visually and can be used to extend or sometimes even replace textual documentations of software projects.", Draw.io VS Code Integration.
  - Project `workflow` from Lewis et al 2014
  - Project `workflow` from Clemens et al 2021



- Start the latex doc early! Prototype latex doc that Jun uses personally in this git repo. (This is good for Project Seeds, and papers. Probably not good for thesis.)

3.4 When to go full-on DevOps, and when not to?

These tools, like the project workflow schematic, are sometimes useful but sometimes overkill. It is important to have the ability to use these tools, but then it is ok to decide not to use them.

- Wilson et al, "Good enough practices for scientific computing" PLoS Comp 2017

Here are some rough guideline

- **Even for the simplest code:**
  - Thoughtful variable names with consistent style (`camelCase`, `snake_style`, etc)
  - Comment at the top that says what the script tries to achieve
- **When you have a presentable figure**
  - git and version control
  - Design a single source of truth, within a scripts. i.e., design so you don't need to type in the same information twice, unless there is a test that shows an error if they mismatch.
- **Past the Norris limit of around 1500 lines**
  - The Norris limis is rough amount of code an untrained programmer can write before the code becomes so tangled that the author cannot debug or modify it without herculean effort.
  - Dedicated /doc directory
  - 10-20% of time "refactoring" (improving the code even if it's working fine as is). An expression from software engineering is "technical debt", the amount of disorganized clutter that slows future progress. How much time should you spend paying down technical debt (re-organizing your notes and directory structure, taking notes) versus producing results? This is called "refactoring".
- **Once the first draft project outline has crystalized (you know the question/gap and approach, and know the figures to gene**
  - Design a single source of truth, within the workflow.
  - `Quickstart.md` to do the simplest complete figure generation (generate data and plot it, e.g., in Matlab)
  - Project workflow diagram using draw.io or similar

CHAPTER

FOUR

WHAT IS A PHD?

- Fleming "Top tips for avoiding last-minute disasters and filing your thesis on time" Nature 2018
- Taylor "Twenty things I wish I'd known when I started my PhD, Nature 2018. I agree with all of these.

Three published or publishable papers, at least two of which are first-author, makes a PhD. Advancement is when you have a draft of the first paper, and the proposal and proof-of-concept for two more.

There are many equivalences (a high-impact paper probably counts for two; 3 middle-author papers probably count as a first-author paper; a widely-forked software..). Then, the onus is on the Candidate to argue this equivalence.

Different PhD programs have different standards. Different co-advisors have different standards.

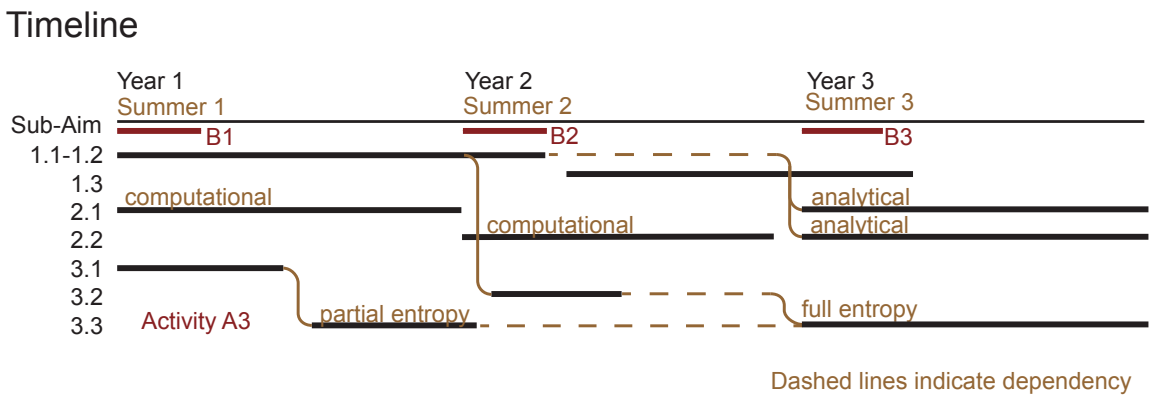
4.1 Advancement and defence

One of the most fruitful, energizing and empowering activities in science is getting a bunch of smart people in a room and talking about a research project with them. One of my favorite memories from recent years is one time when two senior faculty had a project they were excited about, and they got a few other faculty in a conference room to discuss their project idea. Advancement to Candidacy is one of your opportunities to do this! It is both a time to get smart people's feedback. It is also stressful because it involves a tonne of work, and because it is a checkpoint to discuss challenges. Here are some tips.

4.1.1 Advancement tips

- Projects don't need to be related.
- Co-authored papers can be excerpted, provided you add a suitable Intro and Discussion, and a Statement of Co-authorship
- Common pitfall in Advancement presentation: Be clear to distinguish the state of the field and your contribution.
- Unlike for papers, the title doesn't need to actively contain the result.
- Executive summary is self-contained. See "Project summary" examples.
- For published/submitted papers, copy the tex source and dump it in your Advancement write-up. What was Supp Mat or Appendix can be part of main body.
- For unpublished projects, be sure to include
  - Target results
  - Preliminary results: at least one figure
  - How you hope the results address the specific question

  - Timeline



- Use bibtex and a modern bibliographic tool for References (Mendeley, ReadCube Papers, . . .).

4.1.2 The prototypic timeline

- T-minus-6 weeks: First draft of write-up to Jun
- Once Jun approves draft:
  - Invitation sent to prospective Committee members
  - Scheduling
- T-minus 2 weeks:
  - Write-up sent to Committee
  - Presentation draft sit-down with Jun
- T-minus 1 weeks: Presentation practice with group
- Weekend or evening before: Final run-through with Jun
- T-minus 0: Advance!

How to pick a committee and schedule the exam: Pick a 1-2-week interval, and make a ranked list of potential committee members, then try to find a committee that fits into the time interval. This is much easier than setting a committee then finding a time.

- Levine "Doctor's advice" Nature 2016 on choosing a Committee

4.2 Quarter reports

It's easy to get lost in details and short-term milestones, forgetting about the big picture, so we put this in place to force ourselves out of the "urgent-vs-important" trap. The quarter report is an opportunity to think about the big picture. This is a short summary of what you've done and are planning to do. It can be as short as 7 sentences, but can be longer.

1. This quarter I planned to…
2. I generated the following results…
3. The main challenges were/are…
4. I presented my work by/at… (paper submissions, conferences, …)

5. In addition to my research, this quarter I (took classes, organized a seminar series, TAed, mentored undergraduate or rotation student…)
6. Next quarter, I plan to…
7. (If <100%) My percent-effort on these projects was… [This is so we can both keep track of time off for personal reasons, projects with other PIs, classes etc.]

Put your report in an editable format (Google Doc, latex, MS Word doc), and we will schedule a special slot to go through it together.

Bonus topics you're welcome to include (in addition to anything else you want to):

1. For one of my current projects, of all the things that might happen, here is one cool thing, one weird thing, and one bad thing that we could discover:…
2. A skill or technique I want to learn and teach the group is…
3. In the own-learn-teach-delegate axes, here is something I want to delegate to Jun/someone else, and something I want to take ownership of:…

CHAPTER

FIVE

THE GROUP!

5.1 Computing

5.1.1 Local machines

There is a document with account information including IP addresses. (For best-practices reasons, this information is not in a sharable resource.) The document is in the Google Drive AllardLab/Machines.

All of our local lab machines (iMacs, Linux boxes) should have a standardized administrative account and standardized password. All of you should not use this for any actual work. Instead, you each have a user account on your primary machine. Sometimes we ask each other for access to each other's primary machines, for access to files or computing power. If so, make your own account.

5.1.2 High-performance computing: hpc3

The Allard Lab is allocated a certain number of computing hours on UCI's shared high-performance computing facility, Google Drive AllardLab/Machines., under `jallard_lab`.

Our rough guideline is that, if you are using *>1000 core-hour equivalents / week*, we need to make a [budget, e.g., using Google Sheet example. If you are doing exploratory stuff, try to keep the jobs <1000 core-hours/week. If you need help being clever about how to do more with less, let me know!

5.2 AllardLab G Drive

There is a shared Google Drive at AllardLab that can be used for file storage. Each of you can make a dedicated folder under LabMembers, and every project can have a dedicated folder in Projects. Some things, like code, belong in a git repository. Some things, like gigabyte data, belong on one of our external drives in RH274 instead.

Should we make a group Discord?

5.3 Resources

What shared resources should we develop? What outside experts should we call in for this?

- hpc3 tips (Private link, see Jun and Read Lab)
- Adobe Illustrator

Within-team wisdom-sharing has been the single biggest superpower of the Allard Lab. It leads to exponential improvement, like compound interest. How do we maintain that in an asynchronous world? A remote world? My worst concern is that the advantages of asynchronous Work From Home are *multiplicative*, while the advantages of synchronous, in-person work are *exponential* (because they compound). If true, at first it seems like WFH is better, but you pay the price later.

5.4 Rowland 274

- The printer
- The coffee maker
- What else should we do?
- How do we maximally exploit RH274 in the post-covid era?

5.5 Activity

Should we do a weekly Starbucks/Peet's coffee break?

CHAPTER

SIX

WORKING WITH JUN / "A USER GUIDE TO JUN"

6.1 Old e-mails

6.1.1 An old e-mail about receiving peer reviews

"…it is possible we get our referee reports back …

Even when the reviews are good, they are often worded in a way that sounds harsh. It is difficult to tell whether we should resubmit and rebut, or move to another journal, without careful thinking.

My rule is that, when I get review reports, I read them fully, then I set them aside for 24 hours for my emotions to cool before I think about next steps.

The history of scientific progress is contained in all the papers that were, at first, rejected. The following advice I was given by my friends in high school: If you are not getting rejected some of the time, you are not aiming high enough!"

6.1.2 An old e-mail about accepting a conference invitation

"Many times, I've made decisions to spend time with family and friends rather than scientific opportunities, and I have never ever regretted it. A few times, I've tried to juggle too many things simultaneously, and then I've sometimes regretted it – feeling like I gave both family & science my "leftovers". I have also had events where the most valuable thing was simply showing up, to show support for family or friends, e.g., to give one specific hug, and it didn't matter whether it lasted 1 minute or 3 days. If it were me making this decision, I would think of it as speaking at the conference and giving my family my leftovers (which might be enough) versus going to the family event, and giving the conference my leftovers (in which case we can't commit to speaking)."

6.2 "A user guide to Jun"

1. How long should you stay stuck on something, trying to figure it out for yourself, before coming for help? About one week.
2. The best way to reach me is e-mail.
3. Under normal operations, the turnaround time for e-mails is within 48hrs of "business time" (e.g. weekdays). Sometimes it is crunch time or something logistical (we're at a conference, we're hosting a visitor), and turnaround needs to be faster. Sometimes it gets slower, e.g., when you're on vacation, or weekends or planned time-off.
4. **It is more important for me to know when you're on reduced effort than that your effort be maximal.**
5. I take working with collaborators seriously. By far the most impactful stuff we do is collaborative, but collaboration means it's not just about you. I take my role here seriously and expect you to as well.

6. I take student privacy seriously, and I take openness with students seriously.
7. Time is the single most important resource, both materially and symbolically. Not wasting other people's time means showing up on-time to meetings and not going over in a formal time allotment.

6.3 Jun's notes on finding productivity

- I came up Jun's 3 rules for productivity way way back when I was an undergrad at Queen's. They are not universal. But even after all this time, they still ring true.
    1. Break before you're forced to. Break often and break well.
    2. As a corollary to the first, Push when you don't have to. Find projects that you *want* to push when you don't have to.
    3. And finally, Act as if nothing more is going to get done next week, next quarter, next year (what I now call "the Fallacy of Next Quarter").
- If you look at your agenda and you feel overwhelmed, subdivide your tasks into subtasks. If you're still overwhelmed, cut them up again. Repeat, until each task fits in half-hour chunk. Then start doing. When rock climbing, sometimes you look up from your current position, you see no holds, no path up. It is amazing how different the wall looks when you pull yourself up even by a few inches – new holds and pathways are revealed.
- How much time should you spend making yourself faster, eg by automation? See XKCD Productivity tip