CSC 595 - Research Skills

How to Recognize Great Ideas

Nishant Mehta

*For the overall structure and some content, heavy credit goes to
Nick Feamster and Alex Gray: https://noise-lab.net/research-course/

Exercise

- Think of one example of a great research idea from your area?
- What makes the idea great?

Signs of a great idea

- Originality
  - How original is the idea?
  - Does the idea involve new, original techniques/designs?
- Has the idea changed the trajectory of many researchers?
- Did the idea resolve a long-standing open problem?
- Paradigm shift (revolutionary science impact)
  - Has the idea led to changes in the “rules of the game” / assumptions (non-Euclidean geometry)
  - Did the idea obtain a surprising answer to an open problem? (future e.g., showing P = NP)
- Impact beyond the area:
  - Has the idea had real-world impact? (effect on practice, industrialization)
  - Has the idea had impact on other disciplines?

On Impact

- One of the most difficult things to measure
  - Inherently forward-looking
  - Requires understanding a trajectory (past, and predicted future); difficult for newcomers to an area to judge
- Revolutionary papers likely to have higher impact
  - New techniques, new language, hard-to-believe results
  - May be rejected many times due to reviewers’ disbelief or lack of time/patience to understand (famous examples abound)

Ways not to measure impact

- Prominence of first (or senior) author
- Number of citations
  - Something from an old paradigm that helped influence a new paradigm might not be cited anymore. Yet, it has been very influential…
  - A textbook (or to a lesser extent, a paper) might be so impactful that people don’t even see the need to cite it anymore (knowledge therein became so widespread that citation is besides the point) - no one cites Gauss/Laplace for the Gaussian distribution
- How aggressively the authors advertise their paper on Twitter / X …

From Problem and Solution to Magnitude

- Heuristic: **Magnitude = (Problem Size) x (Solution Goodness)**
- Can be difficult to determine which is better:
  - Making small progress on a major problem
  - Completely solving a small problem
- How to determine how major a problem is?
  - Consider size of community interested in problem
  - What is the rate of progress on the problem? Stalled for many years means new progress could be exciting
- Problem size and solution goodness don't tell the whole story
  - Consider also new perspective on problem (can lead to future impact) and new techniques present in solution (also can lead to future impact)

Two first, major questions when starting to review a paper

(a) Does the paper have a great promise? That is, if the authors do what they claim to do, is the paper awesome?

(b) Did the paper deliver on its promise?

- If paper delivered far below promise, probably headed to reject pile.
- If mostly yes and answer to (a) was yes, likely a very strong paper
- If mostly yes but answer to (a) was no, need to look more closely (see if there is some nugget in the paper that still warrants acceptance; this is tricky to formalize and may be highly subjective…)

why “mostly”? good to overlook small issues

Tips for evaluating papers (for any research area)

- Is the problem important?
  - Think about whether there is a community that would care about a solution to the problem (even if you yourself wouldn't care, possibly because you are in a different sub-area)
- How well/to what extent does the paper solve the problem?
  - If warm-assumptions were introduced, is the field in a better position to address original version of problem (without warm-up assumptions) after this paper? Consider perspective these assumptions provide (itself a contribution?)
- Is there a cool idea in the paper? Things like a new algorithmic technique, new proof technique, new type of visualization, new math trick, new reduction, way of vastly extending a previous solution to a larger class of problems, etc.
  - Careful: ingenuity must be for a purpose; fancy ideas should be justified. Do simpler approaches not work (determining this may require very high expertise)
- Is the main contribution(s) important? (related to "Does the paper have a great promise?")
- Does the paper actually achieve that contribution(s)? (related to "Did the paper deliver on its promise?")

Guide to Reviewing

Reviewing vs Reading

- Reviewing
  - Ultimately is about giving evidence that paper should be accepted (or should not be accepted) to a conference
  - Giving feedback to authors to improve paper (whether paper will be accepted or not)
  - Also, hidden goals: expand your expertise and impressing senior people (reputation)
- Reading
  - For your own research; many different types of reading (more on this soon)

Example Review Form (NeurIPS 2025)

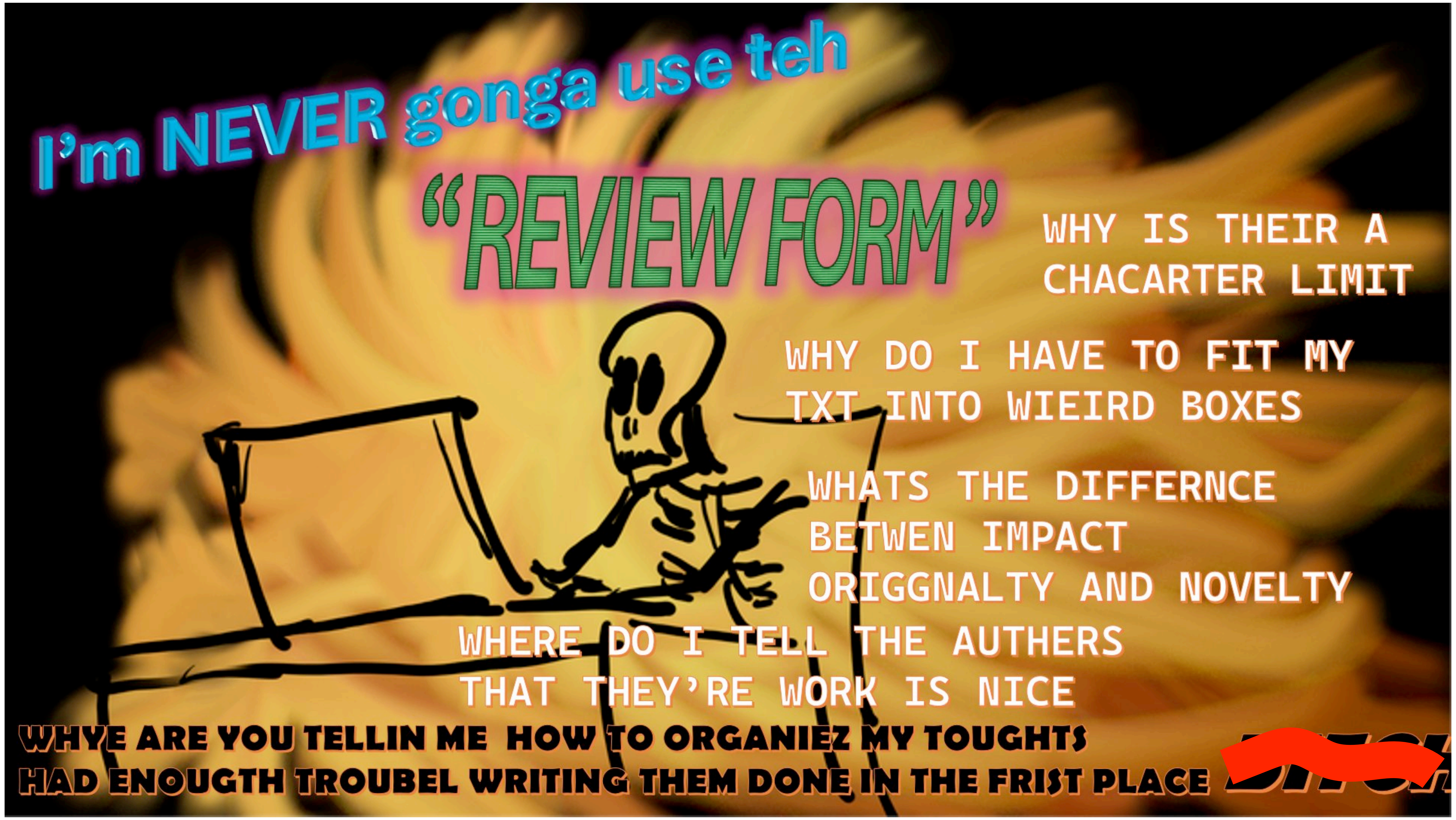
- Summary
- Strengths and Weaknesses
- Quality - score
- Clarity - score
- Significance - score
- Originality - score
- Questions
- Pros/Cons
- Limitations
- Rating - score
- Confidence - score
- Ethical concerns

Example Review Form (AISTATS 2025)

- Summary and Contributions
- Soundness
- Significance
- Novelty
- Non-conventional Contributions
- Clarity
- Relation to Prior Work
- Additional Comments
- Reproducibility
- Rating - score
- Confidence - score

Example Review Form (AISTATS 2025)

- Summary and Contributions
- Soundness
- Significance
- Novelty
- Non-conventional Contributions
- Clarity
- Relation to Prior Work
- Additional Comments
- Reproducibility
- Rating - score
- Confidence - score



Summary - Overall point

- Convince someone else that you read the paper (also authors can clarify if they disagree with the summary)
- Provide some understanding of the paper that can help the area chair (who usually doesn't have time to read all the papers in their stack)
- Summary is ***not*** meant to include critical feedback (save that for rest of review)
  - Avoid "The authors attempted to show (but there are errors) that…"
  - Avoid "The authors *tried* to solve…" (suggests you don't think the authors solved it…)

Summary - What to write

- Give some context if necessary (“Several papers have recently shown evidence of a double-descent phenomenon in machine learning.”)
- What problem do the authors tackle? (“This paper sets out to demonstrate that the double descent phenomenon disappears if we change the way we measure complexity.”)
- What are the main contributions?
  - Try to describe each one (briefly)
  - Don’t just copy-paste from the paper. Write according to your own perspective and understanding
- Try to briefly describe any new techniques

Two examples of summaries of the same paper

[Example 1]

This paper explores a new approach to replacing features in the forward pass of a neural network, for purposes of interpretability. The approach is, rather than zeroing out features or replacing them with some pre-specified constant or random variable, to optimize a replacement constant to minimize the loss of the model over the subtask of interest. So when ablating an attention head or MLP layer, the ablated component's outputs are replaced with a constant that is optimized to maximize model performance over some data of interest. This is done to be "minimally disruptive to model inference", e.g. avoid disrupting model inference in the way that OOD replacement values are known to disrupt model inference, as has long been lamented by past work in interpretability. This new feature replacement method is demonstrated in three case studies, focused on circuit analysis, factual association localization, and something to do with the tuned lens that I could not quite understand. Results suggest that this method reveals clearer/sharper/stronger phenomena in circuit analysis and factual association localization, a promising result for interpretability research.

Examples of 2 summaries for the same paper

[Example 2]

Different intervention techniques aim to "ablate" parts of the representation of the model to infer their causal function, e.g. by adding gaussian noise. the paper suggests to derive a notion of "optimal" ablation. Particularly, instead of zeroing out or replacing the ablated part with its mean, it is proposed to optimize with GD the constant value that minimizes the loss, I.e, the "perfect" ablation would be replacing the element with the "best" constant value. It is shown that the proposed techniques improves circuit discovery, identifying a subnetwork that reconstructs the original performance of the model in some task.

Detailed comments - Things to discuss

- Importance of the problem
- Novelty of design, system, techniques (algorithmic, proofs), etc.
- Novelty of perspective
- Quality of solution
    - For each contribution, try to assess how well authors achieved it and importance of contribution
    - Don’t just say “blah contribution is nice”; explain ***why*** you think it is nice
- Do authors overclaim? (originally claimed contributions are less than what authors show)
- Highly relevant related work that was omitted
- Clarity - good to explicitly tell authors what parts (if any) were unclear so they can improve
- Missing details (related to clarity) - experimental design too high-level, missing details in algorithm, proofs, etc.

Detailed comments - Things to discuss

"E is the new P"

- Flaws in experimental setup (systems, machine learning, user studies)
- Issues with choice of datasets
- Issues that compromise statistical validity (reporting only best results, optional stopping, etc.)
- Assumptions (good? bad? ugly? be sure to explain your opinions)
- Correctness issues with proofs
  - good practice to mention what proofs (if any) you checked
  - think (at least a little bit) about whether there are easy fixes; don't reject paper based on little issues

Grading Criteria

1. Impact
2. Novelty
3. Clarity
    - Clearly stated problem, contributions, notation, and results?
    - Explained meaning of assumptions/theorems?
    - Reproducible?
4. Soundness
    - Flawed experimental design? (for papers with experiments)
    - Incorrect claims? (for theory papers)
    - Incorrect proofs/methodology? (for theory papers)

Getting started on a review

- Do a quicker read to get high-level details (problem, contributions, overview of techniques,, etc.)
  - Avoid getting bogged down in details (don’t dive into proofs yet or read experimental setup in detail)
  - Think about how you would evaluate paper if authors deliver on the promised contributions
- Do a longer, detailed read
  - Write down questions; see if they are addressed as you continue reading
  - These questions (if not addressed) can later be posed to authors
  - Write down criticisms ***and*** praise (I write notes in the margins)
  - Try to write the review the same day if you can (save yourself a context switch!). If you feel some questions might be resolved from another look, revisit the paper a later day, see what gets resolved, but then write the review

Super-important tips

- Be nice - use polite language
    - paper could have been written by a first year grad student
    - authors spent a lot of effort (may have worked on the paper for over a year)
    - research community as a whole benefits from constructive criticism
- Don’t write something like “Theorem 3 is wrong. The paper should be rejected.”
- Instead, write:

  “I believe there may be an issue with the proof of Theorem 3. Specifically, in the second paragraph of the proof, [INSERT DETAILED EXPLANATION]. I would be very interested if the authors could propose to fix this issue (or explain why it isn’t a real issue).”

Super-important tips

- Don’t attack the authors - “In what world do the authors live where it is ok to have an algorithm with $O(n^6)$ complexity?”
- Until the paper is publicly available, don’t use information from the paper for your own research (also, don’t distribute the paper)
- Protect yourself: take care not to reveal your identity
- Protect yourself from getting scooped: take care not to reveal what problems you’re planning to work on

Recommended Reading

How NOT to review a paper
The tools and techniques of the adversarial reviewer

Graham Cormode
AT&T Labs–Research
Florham Park NJ, USA
graham@research.att.com*

ABSTRACT

There are several useful guides available for how to review a paper in Computer Science [10, 6, 12, 7, 2]. These are soberly presented, carefully reasoned and sensibly argued. As a result, they are not much fun. So, as a contrast, this note is a checklist of how *not* to review a paper. It details techniques that are unethical, unfair, or just plain nasty. Since in Computer Science we often present arguments about how an adversary would approach a particular problem, this note describes the adversary's strategy.

1. THE ADVERSARIAL REVIEWER

In Computer Science, we often form arguments and proofs based around the concept of an 'adversary'. Sometimes, this adversary can be malicious; in cryptography they are often "honest but curious". However, the most commonly encountered adversary in Computer Science is the adversarial reviewer: this reviewer uses of the paper while packed into coach class on an intercontinental flight with a small child kicking the seat from behind. Even in favorable conditions, such as a Lazy Boy recliner [1], the adversarial reviewer feels no compulsion to refer to external sources, or find a technical report containing the elusive "full details"[2]. It may be wise for for authors ensure that their work is as readable as possible in worst-case settings.

2. ADVERSARIAL REVIEWING TECHNIQUES

The adversarial reviewer does not reject every paper that they review. In fact, it is often easier to accept a paper (with a short review to the effect of "looks good to me") than to reject one. But, when the situation demands it—say, if the reviewer has submitted a paper to the same venue and wants to even up the odds a bit—a review must be crafted to force the desired outcome. Simply scrawl-

Using LLMs to write reviews

- Don’t do it!
- It’s lazy
- It prevents your growth as a researcher
- It helps our future AI overlords
- It’s ***banned*** from many conferences (ICML 2025 policy: reviewer caught using LLM to write reviews can have their submissions rejected, which also hurts co-authors)

Guide to Reading

Step 1: *What* to Read

- Purpose: You have a new idea and are wondering if it has already been done
- Approach: Consider any paper that comes close to doing your idea. Is there a fundamental paper (or two) that any such paper is extremely likely to have cited? Search for all papers that cite that paper.
    - If you can, include extra keywords to limit search (if looking for theory papers, include “theorem”, and if hardcore theory papers, also include “lemma”).
    - May result in 100 to 500 papers. Once the number is down to 200, do some manual inspection (first by title, then abstract). Eventually, may have ~30 papers that require a closer look. From Introduction (and Problem Setup section, if it exists), try to get down to around 5 papers (those 5 papers may need to be scrutinized closely)

Step 1: *What* to Read

- Purpose: Get up to speed in a sub-area within your area, either out of interest, to have coherent discussion with someone, to write an upcoming research proposal (fellowship application?), etc.
- Approach: Look for recent strong, well-written papers (with lengthy, well-written related work sections). See what papers they cite. For some of the cited papers, *again* follow citations from related work section.
    - Try to include some very recent (strong) papers, somewhat older papers, and a few classics (well-cited old papers, or decently cited old papers by famous people which may contain under-appreciated/unnoticed intellectual nuggets)

Step 1: *What* to Read

- Purpose: Learn about trends in an area (to keep up to date, think about problems to work on, etc.)
- Approach 1: Look at oral sessions (usually organized by sub-area). Select a few papers from each oral session.
  - At some conferences, papers that are presented orally are the very top papers, so this is a reasonable filter. For papers where *everything* gets an oral session, the total number of papers may not be that large so this is feasible.
- Approach 2: Ask expert colleagues

Step 2: *How* to Read

- Consider the Abstract optional - authors usually write this last, and once you've decided to read a paper, the abstract might be necessary to read anymore
- Introduction - in particular, figure out what problem the authors are addressing, and hone in on "contributions" (hopefully they are presented in one place)
- For technical, theory works, look at Problem Setup (or similar name) to get a clearer idea of the problem (very important when the goal is to identify if someone did what you are planning to do)
- Look at Conclusion/Discussion - usually easier to read than Introduction, may hint at extensions and future work (especially important if you are searching for problems to work on, but consider what authors may already be working on…)
- Then, as necessary, try to jump to specific sections of interest (e.g., new algorithm and analysis, or hardness result)

How to Conduct a Literature Survey

- First rule: **Don't write a survey paper if you aren't currently an expert**. Why?
  - It's a *lot* of work. People may not appreciate your survey unless you can provide a unique and valuable perspective (which usually requires a high level of expertise). Consider trade-off with doing your own original research (new project)
- Second rule: see first rule

researching in an area
not write a survey paper

"If this is your first time at ~~Fight Club~~, you have to ~~fight~~."

How to Conduct a Literature Survey

- So, why are you doing a Literature Survey? Because you have your own idea and want to write a related work section (plus improve your ideas by being informed of related work).
- Start with a seed paper (could be multiple seed papers).
    - Example: A new restaurant reviewing system that incentivizes reviewers to post reviews honestly.
    - Look for a recent, well-cited paper that is close to this topic, ideally by someone reputable.
    - Look at this paper’s related work section (gives a reasonable number of papers, probably not more than 20), and also look at papers that cite this paper (could be large; need some way to filter); this gives an expanded set of papers.
    - For some of those papers, repeat this process.
    - When looking for papers that cited the original seed paper, try using additional keywords to filter out things that are irrelevant for your purposes.

How to take notes

- My approach:
  - For a particular literature survey / related work investigation, keep a single LaTeX file
  - For each paper, write the title and citation (citation via BibTeX), and write down a few sentences (your notes, which might not be a summary; could be as simple as “does not contain any theoretical results”)
- My approach isn’t ideal, but for me it works ok
- Grad school is a good time to experiment with different approaches
- Try asking senior labmates what they do (ideally, senior PhD students)
- Try asking your advisor what they do