

RESEARCH PAPER GUIDELINES SHEET

by James H. Faghmous

INTRODUCTION

For many of you, this might have been the first time you have written a scientific research paper. The writing portion of this course has been taught, in most cases, in the same manner than computer science, that is, *learning by doing*. Your first drafts were a good way to get your hands dirty and get a sense of how much work needs to be done in order to write a scientific research paper. This article will assist you with revising your first drafts (in terms of content and style) to produce sound scientific research papers.

In the feedback you received on your first drafts, you might have seen a comment along these lines: “*You introduction could be expanded. See II-b for help.*” This is referring to item II-b on this feedback document. Although, this feedback sheet might have been referenced only a few times in your first draft, we suggest reading through the whole document for completeness.

This paper is divided into three sections: I- Research; II- Structure; and III- Style. In the *Research* section we will discuss aspects dealing with the clarity and validity of your research topic. In the *Structure* section, we will review how to best structure your paper to present your findings effectively. Finally, the *Style* section will guide you through the process of writing publishable content. Before we dive into the content of your research papers, it is important to highlight our expectations, in case you were not aware of them.

EXPECTATIONS

As we have emphasized throughout the course, *you must always write your assignments with the grading rubric in mind*. To ensure fairness in grading, we will evaluate *every paper* on the same set of criteria. These criteria are listed on the assignment page, but we will add them here for your reference:

1. Contains accurate grammar and spelling, and complies with the ACM style for citations
2. Forwards an intriguing and relevant thesis, is not merely a restatement of known arguments
3. Research is thorough and covers topic from many angles; the writing is not based on a greatly limited set of information
4. Research presentation is well structured, with good writing transitions from point to point; no extraneous or incomplete points distract the reader.

This document will help you meet these requirements, but first we will reformulate these criteria into written English, for your convenience:

When evaluating your papers, we will first and foremost look if you have

- A. a *clear, concise research thesis*
- B. That you used *strong, logical arguments to support it* (your thesis). We will discuss in greater detail what makes a clear thesis and strong, logical arguments throughout this article.
- C. Your paper should not gather facts from sources that only support your thesis but from opposing views as well.
- D. Your sources should include *peer-reviewed articles* and not only blogs and/or web articles that are not subject to the peer-review process.
- E. Overall, *your paper should be clear and focused* (aka not wordy!) and *follows a logical*

structure.

Now that our expectations are clarified, we will start by helping you formulate a clear research thesis in our *Research* section.

I- RESEARCH

Many of you used your first drafts as an opportunity to learn more about a topic in computer science, rather than researching a specific question. It is important at this stage, however, that we have a clear distinction between: (1) a research topic; (2) a research question; and (3) a research thesis.

Research topic: Your research topic is simply the discipline or field of computer science you are interested in researching.

Example: Artificial intelligence (AI), Programming methodologies, etc.

Research question: Your research question is generally an aspect of your research topic that you would like to investigate.

Example: If your research topic was AI, then this may be your research question:

“How can we learn about our users' health based on their web surfing habits?”

Or another question might be:

“How can we automatically detect an epidemic outbreak?”

Research thesis: Your research thesis should be a clear, concise statement that focuses your research question to guide your arguments and narrow the scope of your paper.

Example: Based on the previous research question, one possible thesis could be:

“We can effectively predict epidemic outbreaks based on search engine queries and the geographical location of the user.”

Here is another example in the topic of computer science education:

“After researching Active Learning techniques, we show that using Active Learning in computer science education can increase (1) student engagement, (2) retention (in the CS major), and (3) performance compared to the traditional lecturing methods.”

To help illustrate the difference between a research topic, question and thesis, please see Figure 1 below

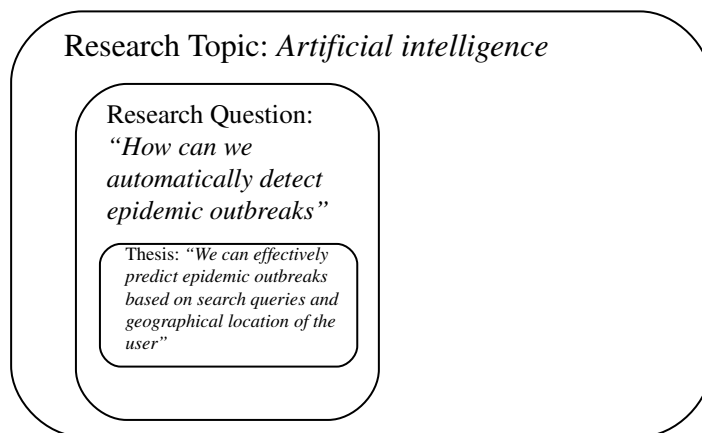


Figure 1 – An illustrative example of the difference between a research topic, question, and thesis. The size of the rectangle represents the “breadth” of each aspect.

A) A Clear and Direct Thesis Statement

Your thesis is perhaps *the most important sentence* of your research paper. An unclear thesis will almost certainly result in an unfocused paper that lacks direction and clarity.

Here are a few examples of unclear research theses and how one could focus them:

- “*Agile development is a new style of development. It focuses on communication and teamwork.*” This is not a thesis. The statement doesn't do anything beyond stating mere facts. It does not give the reader any sense of the direction of the paper.

Here is an improved example:

“*Agile is a new style of software development that will improve customer satisfaction compared to the waterfall model, because of its focus on communication and teamwork*”

- “*I'll review programming styles and show which ones are good and which ones aren't*” This is extremely broad and vague. The author does not specify *which methodologies* will be explored and how are they considered *good* or *bad*.

A more focused thesis could be formulated as follows:

“*We will present the Agile and Waterfall development methodologies and compare them in terms of customer satisfaction and software reliability.*”

B) Data and Argumentative Analysis

As stated earlier, we will evaluate your paper based on the strength of your arguments. Most of the questions you will research have no right/wrong answers, and therefore the strength of your arguments are crucial.

The first, and possibly most important, thing you need to know is that *data is vital to making a strong argument*. Take the second thesis about computer science education, for example. If you were to prove that Active Learning improved student performance compared to standard lecturing methods, you must:

- Define what student performance means (e.g. Is it their overall grade in the course? Their ability to retain the information taught in the course beyond graduation? etc.)
- Provide data showing improvement of the students' performance (as defined in your paper) using Active Learning methods vs. regular lecturing.
- Analyze the data by clearly explaining how it was generated (gathered).

Common mistakes:

- Not providing *any* data to support your thesis
- Comparing two applications of the same method A and concluding that it is better than method B, while never comparing methods A and B
- Using vague explanations/interpretations of the data such as: “The authors ran an algorithm and got the results.” Instead of clearly explaining the workings of the algorithm or experiment.

C) Wide pool of sources

When conducting secondary research—that is, research done by others- we expect you to thoroughly review the literature on your chosen topic. This means you must review the literature that supports and opposes your position.

Common mistakes:

- Not presenting all views on the subject matter.
- Not giving the opposing view of your argument the same emphasis and depth as the favoring view.
- When defining certain terms, students used a single definition when there are several accepted ones. In this case, you should present all definitions and then state which one you will follow throughout your article.

D) Clarity

Communicating ideas clearly will always enhance the quality of your paper. A clear paper starts with a clear thesis and outline and should be self-contained (you can understand the contents of the paper without reading any other references). One way to achieve such clarity is to *define all technical terms you use repetitively*. Depending on how important these terms are, you may spend anywhere from a sentence to a whole section presenting/defining them.

Although the assignment states that your audience are your peers, you must keep in mind that they might not be familiar with every aspect of your research topic and, furthermore, certain technical terms have numerous definitions; therefore providing such definitions is necessary.

E) Quality of References

Since most of the information you will include in your papers will be based on someone else's work, the quality of your references will heavily influence the quality of your final paper. Most of you will find an abundance of information on the Internet and might be tempted to use any sources that fit your point of view. Using web articles and blogs as the primary sources of information for your paper has (amongst others) three major drawbacks:

1. You have little or no information about the author and his/her credentials.
2. These types of articles do not list their references so you have no way of going back and checking the validity of their claims.
3. Such articles are not subject to cross-examination or review so they could be highly biased and/or inaccurate.

Although peer-reviewed publications are not error proof, they are widely accepted as a sound source of information for research papers.

The best way to utilize the content-rich nature of the Internet is to use web articles and blogs to learn more about your topic and formulate a research question. After that, you are better off reading peer-reviewed research papers since your scope is significantly narrower than you could easily find materials on your thesis.

II- FORMAT

Although the assignment did not specify any particular layout for your paper, this section provides basic suggestions to make your paper easy to read by putting some thought into the overall structure of the final product. This section and the next (III- Style) are crucial for a writer because when you are in the business of publishing (or any business for that matter,) designing products that meet the customer's needs should be your primary goal. These sections will help you highlight the content of your product to help the customer decide if it's right for him/her not.

A) Abstract

In the academic publishing industry, the main source of advertising your published research is through online databases. Generally, only your abstract appears on such databases, then the customers chooses to read the full-length article or not based on the quality of your abstract.

An abstract should be a complete, clear and concise review of your research.

Your abstract should include the following:

- Motivations: Why is this research important?
- Problem statement: What is the problem you are trying to solve?
- Methodologies: How did go about solving your problem?
- Results: What did your research produce?

Example of a good abstract from Mitchell T. et al. (2008):

The question of how the human brain represents conceptual knowledge has been debated in many scientific fields. Brain imaging studies have shown that different spatial patterns of neural activation are associated with thinking about different semantic categories of pictures and words (for example, tools, buildings, and animals). We present a computational model that predicts the functional magnetic resonance imaging (fMRI) neural activation associated with words for which fMRI data are not yet available. This model is trained with a combination of data from a trillion-word text corpus and observed fMRI data associated with viewing several dozen concrete nouns. Once trained, the model predicts fMRI activation for thousands of other concrete nouns in the text corpus, with highly significant accuracies over the 60 nouns for which we currently have fMRI data.

Common mistakes:

- Not including any/all of the four components of an abstract (motivations, problem statement, methods, results.)
- Not using the information in the abstract in the paper. Think of the abstract as an independent entity. It would not be repetitive to use the information in your abstract in subsequent sections of your paper
- Citing sources in the abstract. You do not have to cite any sources in your abstract.

B) Introduction

Your introduction is the backbone of your research paper. It assists the reader in understanding the direction the paper will take and should strive to make the reading experience as pleasant as possible. In many instances a reader might never go past your introduction if it is unclear and too broad.

Generally speaking, your introduction should have the following components:

- Problem statement (might be a more elaborate version than the one in your abstract)
- Motivations
- Your research thesis
- An outline of your paper

Here is a good example of introduction from Koller, D. and Friedman, N. (2007):

In the previous chapter, we described the basic principle of decision making under uncertainty — maximizing expected utility. However, our definition for a decision-making problem was completely abstract; it defined a decision problem in terms of a set of abstract states and a set of abstract actions. However, our overarching theme in this book has been the observation that the world is structured, and that we can obtain both representational and computational efficiency by exploiting this structure. In this chapter, we discuss structured representations for decision making problems, and algorithms that exploit this structure when addressing the computational task of finding the decision that maximizes the expected utility.

We begin by describing *decision trees* — a simple yet intuitive representation that describes a decision making situation in terms of the scenarios that the decision maker might encounter. This representation, unfortunately, only scales up to fairly small decisions tasks; however, it provides a useful basis for much of the later development. We then describe *influence diagrams*, which extend Bayesian networks by introducing decisions and utilities. We then discuss algorithmic techniques for solving and simplifying influence diagrams. Finally, we discuss the concept of *value of information*, which is very naturally encoded within the influence diagram framework.

This particular introduction is adapted from the draft of an upcoming book and it might have minor grammatical and/or spelling errors. But it provides a good example on the contents of an introduction and is of good length.

C) Flow of your article

Keeping your paper focused on defending your thesis can be difficult. Losing focus or clarity can cause your arguments to seem disorganized and weak. Here are a few suggestions to ensure that your paper is coherent:

- One way to keep your paper on target is to first outline it in your introduction.
- Use section headings (titles) so your reader can always situate themselves within your paper (or argument).
- Use transition sentences at the end (or beginning) of each section to connect your current argument to your previous point.

III- FORMAL EXPECTATIONS & STYLE

A) ACM Citations

Citations occur in two instances in a research paper: in-text and references.

ACM in-text citations:

- Last name name of the first author and year of publication in square brackets [Collins 2008]
- In the case of two authors both last names and year [Collins and Johnson 2008]
- In the case of three authors or more list first author's last name followed by et al. and year of publication [Collins et al. 2008]

If you are citing more than one paper, list them according to how they should be listed individually separated by a semi-colon [Collins 2008; Collins and Johnson 2007; Collins et al. 2006]

ACM references:

- Scientific journal:
LAST NAME, FIRST. Year. Title of article. *Name of journal* (italicized) volume number, pp-pp.
Example:
ABDELBAR, A.M., AND HEDETNIEMI, S.M. 1998. Approximating MAPs for belief networks in NP-hard and other theorems. *Artificial Intelligence* 102, 21-38.
- Conference proceedings:
LAST NAME, FIRST. Year. Title. *Name of conference*, city, state, date of conference, pp-pp.
Example:
MAREK, W., AND TRUSZCZYNSKI, M. 1989. Relating autoepistemic and default logics. In *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, Canada, May 1989, 276-288.
- Book:
LAST NAME, FIRST. Year. *Title*. Publisher, city, state.
Example:
GINSBERG, M. 1987. *Readings in Nonmonotonic Reasoning*. Morgan Kaufmann, Los Altos, CA.
- Web article:
LAST NAME, FIRST. *Title*. Title of online publication. (Date of publication). Date of access <url>.
Example:
BERNSTEIN, M. *10 Tips on Writing The Living Web. A List Apart: For People Who Make Websites*. (16 Aug. 2002). 4 May 2006 <<http://alistapart.com/articles/writeliving>>.
- Weblog post:
LAST NAME, FIRST. "Title of Entry." Weblog Entry. Title of Weblog. Date Posted. Date Accessed <URL>.
Example:
HAWHEE, D. "Hail, Speech!" Weblog entry. Blogos. 30 April 2007. 23 May 2007 <http://dhawhee.blogs.com/d_hawhee/2007/04/index.html>.

B) Proofreading

As stated in the rubric your papers must contain accurate grammar and spelling. We are going to be less lenient when grading your final submissions. It is, therefore, crucial your papers are free from any spelling and punctuation errors. A good strategy is to proofread your paper on your computer once, then proofread it again in hard copy. You should also have someone else read it to ensure that it is clear and error-free.

C) Conciseness & Clarity (based on suggestions from Markel, M. Technical Communication (2004))

Producing a clear, concise final paper is key to obtaining a good grade. You must display your ability to communicate technical content accurately and effectively. Conciseness means writing only what is necessary for the reader to clearly understand your message. Here are a few suggestions on how to stay brief while still communicating your central idea:

- **Avoid overly short sentences:**

Sometimes we confuse concise with short and this can result in choppy sentences that contain too little information and might seem ambiguous if taken out of context.

Example:

I have experience working with various programming languages. Some of these programming languages are C++, Java and Perl.

Improved example:

I have experience working with various programming languages, including C++, Java and Perl.

- **Don't bury your "real" subject or verb:**

The "real" subject of a sentence should be its grammatical subject. Sometimes students bury the "real" subject preceding it by a weak subject. This creates overly wordy sentences with buried messages.

Example: The use of this method would eliminate the problem of miscommunication.

Improved example: This method would eliminate the problem of miscommunication.

Similarly, the "real" verb can be buried as well when nominalizing a verb (transforming it into a noun) and coupling it with a weaker one.

Example: Each computation of the solution is done twice.

Improved example: Each solution is computed twice.

- **Be specific:**

Being specific is key to producing clearly written papers. Being specific involves using precise words, providing adequate detail, and avoiding ambiguity.

Example: An engine on the plane experienced some difficulties

Which engine? What plane? What kind of difficulties?

Improved example: The left engine on the Martin 411 lost power during flight

- **Avoid obvious statements:**

Writing gets too wordy if it over-explains or states the obvious.

Example:

The market for the sale of flash memory chips is dominated by two chip manufacturers: Intel and Advanced Micro Systems. These two chip manufacturers are responsible for 76% of the \$1.3 billion market in flash memory chips last year.

Improved example:

The market for flash memory chips is dominated by Intel and Advanced Micro Systems, two companies that claimed 76% of the \$1.3 billion industry last year.

- **Avoid filler words:**

Sometimes students tend to use filler that is more suited for speech. Using such words creates bloated and vague sentences. Here are few examples of filler words:

basically kind of

certain sort of

essentially various

Example: I think that, basically, the board felt sort of betrayed, in a sense, by the kind of behavior the president displayed

Improved example: The board felt betrayed by the president's behavior