

Literature: Finding & Citing

Roderick Bloem

IAIK – Graz University of Technology
Roderick.Bloem@iaik.tugraz.at

Research

Goal: Generate new knowledge

How do you measure the quality of a researcher?

- Can you?
- By publications and citations

Publications

- Conferences
- Journals

Why do we find conferences and journals important?

Conferences and Journals

Organization

- Conferences: program chair and program committee (prestigious)
- Journal: Editor in Chief and editors (even more prestigious)

Procedure

- Submit paper
- Paper is reviewed by several colleagues (anonymous)
- Paper is accepted or rejected
- Paper is modified based on reviews

Bottom line: Academic publications have a quality standards

Which Papers are Accepted

What people look at:

- Is this new?
- Is it useful?
 - How is usefulness shown?
- Are all the claims in the paper supported?
- Is proper account made of prior work?

In Contrast: Newspapers and the Web

- Content not new
- Claims not backed up.
- **No way to check if the journalist is wrong!**

http://www.krone.at/krone/S6/object_id_109903/hxcms/index.html, 26 October 2008.

“Feng Shui im Kinderzimmer”:

“Elektronische Geräte, wie Fernseher oder Computer sollten aus dem Kinderzimmer ferngehalten werden, da deren Schwingungen das Kind in seinem Schlaf stören könnten. Wenn gewisse Geräte dann Einzug halten müssen, sollten sie mindestens zwei Meter vom Bett entfernt stehen.”

The question you always have to ask: **why do you think so?**

Another question: Would Krone have published an article saying “TVs have no influence on how well your children sleep”?

- Note: What Krone writes is not necessarily bad, it is just a different style
- Newspapers, Wikipedia, and Google are great places to start.
- Next step: find the original sources and check if they make sense!

The web, wikipedia, newspapers, magazines, etc are **not** citable!

Go find the original sources!

But: Using a source, even a dubious one, without proper reference constitutes plagiarism!

Why Cite?

- Your reader is critical
- To give credit
 - Not giving credit is stealing ideas

Reason I: the Critical Reader

You want to put forward an argument

Electronic equipment should not be put in a child's bed room.

Your reader (a 16 year old child) thinks:

Nonsense!

So you write:

As Jones has shown [reference 1], electromagnetic waves may interfere with sleep. Since it is well known that TVs and some other electronic equipment emit electromagnetic waves [reference 2,3,4], one should avoid putting such equipment in children's rooms.

Your reader thinks.

Hm, even if they are turned off?

You have a hole in your argument, but at least the reader has the chance to go to the library and **check whether [2,3,4] really provide compelling evidence**

I will tell you *how* to cite in a minute

Reason II: Plagiarism

What you **can** do

- Discuss or use someone's idea, while giving her credit

What you **cannot** do

- Discuss or use someone's idea without giving credit.
- Copy someone's words
- Copy someone's argument

Plagiarism: Ideas

If you write

In this paper, we show that the speed of any object is limited by the speed of light,

You are stealing an idea.

Stealing ideas or text is a capital crime. A scientist will lose her job over it.
(And plagiarism is always discovered, sometimes years later.)

So what do you do?

You write

In this paper, we discuss why the speed of any object is limited by the speed of light, as shown by Einstein's special relativity theory [refer to Einstein's paper].

Some ideas are so trivial that you do not need citations:

- Winter days are shorter than summer days.
- Children result from sexual intercourse.

When in doubt: cite.

Plagiarism: Words

Plagiarism is the unauthorized use or close imitation of the language and thoughts of another author and the representation of them as one's own original work.

PLAGIARISM: a verbatim copy from wikipedia!

You also commit plagiarism when you

- Translate the text to german
- Change the text

What you are allowed to do

- Read the source; go to bed. The next day, without opening the source, write down what *you* think plagiarism is. (And then check that it is not too close to your source.)

What you can also write:

Wikipedia defines plagiarism as “the unauthorized use or close imitation of the language and thoughts of another author and the representation of them as one's own original work.”

But try to avoid having more than one or two literal citations.

Authorization?

- Wikipedia is wrong, plagiarism is plagiarism even if authorized!

Plagiarism: Arguments

Never take a paragraph, and rewrite each sentence.

This is not plagiarism of words, but plagiarism of an argument.

Rather: understand what the paper tells you and write your paper the next day in a different room.

How to Cite

- Have at most one attributable idea per sentence, so that no confusion ensues.
- The exact syntax of the citation is unimportant. Imitate a paper you find.
- Example

As Jones has shown [1], electromagnetic waves may interfere with sleep.

And at the end of the paper:

[1] A. Jones, Electromagnetism and sleep, Journal of Neuropsychology 2:12-34, Springer, 1996. (Note: I made this reference up.)

- You typically include
 - Author names, paper title, name of Journal or conference, Journal volume, page numbers, publisher, year of publication
- Important: it must be easy for the reader to find the paper.
- Avoid citing more than single ideas. (Arguments are your business!) If you have to, be clear about it.

Smith argues that sleep deprivation is particularly serious for infants. Let us reconstruct Smith's argument [2]. First, Smith assumes that.... Then, she shows that.... This allows Smith to conclude that...

Cite Carefully

- Is the article trustworthy?
- Does the article really show what you want to say?
- Are there articles claiming the opposite?
 - Is cell phone use unhealthy?

Compare

- Chemicals in French fries cause cancer
- Chemicals in French fries seem to be a factor in causing cancer
- Chemicals in French fries correlate with a higher incidence of cancer

[Booth, Colomb, Williams]

Literature Search

1. Figure out what to write about.
Best: a question you find interesting that has not been addressed before.
Often, you find question and literature simultaneously.
Your paper should not describe one paper you find, but rather synthesize an argument from many sources.
 - What is the best way to debug a program?
 - How does security of a computer system depend on its age?
1. Read all you can. Goal: find out what is known, what the vocabulary is.
 - Library
 - Wikipedia
 - Ask people
 - Google
2. Read towards a goal
 - Back to the library!
 - Use academic search tools
 - follow citations in both directions.
 - Forget wikipedia

The Library

- Fachbibliothek Inffeldgasse
 - Small, close, you can touch the books.
 - Good for a first search
 - Search in catalogue is bad.

Go browse the stacks! Figure out the “Aufstellung” – similar books are next to each other.

How to Read

You can read anything in five minutes.

- Papers:
 - Read abstract
 - Read introduction, in particular “related work”. (Related work is sometimes found at end of paper.)
 - Check out the references.
 - If it looks interesting: write down a reference as shown before (authors, title, etc.) + short summary. Perhaps make a copy.
- Books:
 - Read the preface
 - Read the introduction
 - Read table of contents
 - Read introduction of interesting chapters.
 - See if there are “notes” sections with related work

When you know something is interesting, read it in more detail.

Using the Web

There is too much there, and not enough.

- First, figure out terminology
 - Bugs, faults, failures, errors, mistakes, incorrect software, software crash, ...
 - search in English! (Unless...)
- Read wikipedia page to get the background. (Don't expect to use it much – wikipedia is bad at referring to sources.)
- Google around.
- Use the library of congress

Using the Web II

When you know what you want:

- Use scholar.google.com
 - Indexes almost all of computer science
- When you are looking for a particular paper:
 - use DBLP: dblp.uni-trier.de
 - use scholar.google.com
 - use citeseerx.ist.psu.edu
 - use google
 - Search for authors, title, conference name
- To do a search, come to university
 - We pay a lot of money for access to many scientific databases – you will need them!
 - You can also log in from home using VPN.

The Power of Citations

- Citations tell you which papers are related!
- The related work section of a paper is very important
- But who refers to a paper? scholar.google.com knows.
 - scholar.google.com also orders papers by citation count, a rough indication of importance.

While Reading, Learn to Write

We present a novel approach for detection of tubular objects in medical images. Conventional tube detection / liness filters make use of local derivatives at multiple scales using a linear scale space; however, using a linear scale space may result in an undesired diffusion of nearby structures into one another and this leads to problems such as detection of two tangenting tubes as one single tube. To avoid this problem, we propose to replace the multi-scale computation of the gradient vectors by the Gradient Vector Flow, because it allows an edge-preserving diffusion of gradient information. Applying Frangi's vesselness measure to the resulting vector field allows detection of centerlines from tubular objects, independent of the tubes size and contrast. Results and comparisons to related methods on synthetic and clinical datasets show a high robustness to image noise and to disturbances outside the tubular objects.

[C. Bauer and H. Bischof, A Novel Approach for Detection of Tubular Objects and Its Application to Medical Image Analysis, Proc. Pattern recognition, 163-172, 2008]

While Reading, Learn to Write

We present a novel approach for detection of tubular objects in medical images. Conventional tube detection / liness filters make use of local derivatives at multiple scales using a linear scale space; however, using a linear scale space may result in an undesired diffusion of nearby structures into one another and this leads to problems such as detection of two tangenting tubes as one single tube. To avoid this problem, we propose to replace the multi-scale computation of the gradient vectors by the Gradient Vector Flow, because it allows an edge-preserving diffusion of gradient information. Applying Frangi's vesselness measure to the resulting vector field allows detection of centerlines from tubular objects, independent of the tubes size and contrast. Results and comparisons to related methods on synthetic and clinical datasets show a high robustness to image noise and to disturbances outside the tubular objects.

[C. Bauer and H. Bischof, A Novel Approach for Detection of Tubular Objects and Its Application to Medical Image Analysis, Proc. Pattern recognition, 163-172, 2008]

Further Reading

Booth, Colomb, Williams, The Craft of Research,
The University of Chicago Press, 2003.