

# How to Write a Term Paper or Thesis

Michael A. Covington  
Artificial Intelligence Center  
The University of Georgia  
Athens, Georgia 30602  
<http://www.ai.uga.edu/mc>

Revised June 3, 2005

## Abstract

This is a basic style guide for writing scientific papers under my direction. It was written for internal use in the Artificial Intelligence Center. However, other science departments may find it helpful. You are encouraged to share it with colleagues.

## 1 What is a term paper or thesis?

In the Middle Ages, in order to be admitted to a trade guild, a craftsman had to demonstrate his ability by producing a piece of work, called his *masterpiece*, for examination by the officers of the guild. He was then granted the title of Master of his trade.

When universities were founded, they immediately emulated this practice and began to grant the degrees of Master and Doctor to people who had proved their ability to do scholarly work.

A thesis or dissertation today serves the same purpose as a medieval craftsman's masterpiece. It is proof of ability to do a certain kind of work. Specifically:

- To get a doctor's degree, you must become familiar with current scientific knowledge of your subject, add to this knowledge by making an original discovery, and then report the results in a *dissertation*.

- To get a master's degree, it is sufficient to make a new synthesis or application of knowledge already available, and report the results in a *thesis*.
- To pass certain courses, you must write a *term paper*, which demonstrates that you can connect information and report it in your own words. It is not necessary to make a new discovery or a new synthesis.

To get a degree from an English-speaking university, you must prove that you can write scholarly papers, in English, that reflect state-of-the-art knowledge of your subject.

## 2 How is a term paper or thesis judged?

Presentation is more important than most students realize. Even the most brilliant piece of scientific work is useless if the resulting report is unclear or needlessly hard to read or leaves vital questions unanswered.

I grade term papers on how well they do the following four things:

- Define a clear topic and stick to it throughout the paper, addressing a consistently defined audience.
- Use the best available sources of information and acknowledge them appropriately.
- Display careful organization and clear wording.
- Follow scholarly standards for format, grammar, spelling, and other mechanical matters.

If the paper reports original research or technical work, a large part of the grade is based on the quality of the work itself, but *presentation still counts*; if you can't present your work well, you might as well not have done it.

## 3 Receiving and acknowledging help

Getting help from the right sources is a useful skill. Real scientists do it all the time.

When writing a term paper or thesis, you are permitted to receive any amount of help from anyone, as long as you *acknowledge* the help so that your instructors can distinguish your work from the work of others.

You can acknowledge help in an “Acknowledgments” section at the beginning of the paper, or in a footnote attached to the first sentence.

Be sure to *acknowledge grants* if you are supported by research grant funds; this is usually done by a sentence such as, “This work was supported by National Science Foundation Grant ZZZ-99-99999, Michael A. Covington, Principal Investigator.” It is not necessary to acknowledge scholarships or fellowships.

## 4 Planning a term paper or thesis

Before you start writing — in fact before you get very far gathering information — you must decide two things:

- What your paper is about (the *topic*);
- For whom it is written (the *audience*).

Normally, the audience of a scholarly paper consists of people familiar with the general area but not with the specific topic. For instance, if you are writing about implementation of a numerical equation solver in Prolog, you can assume a passing acquaintance with numerical methods and with Prolog.

To a considerable extent, the choice of audience is up to you. But once you have made it, *stick with it* — unless there is a good reason to do otherwise, write all parts of your paper for the *same* audience.

If you have a hard time visualizing the audience, try writing a paper that you would have understood if someone had given it to you a month ago, before you started researching the topic.

The topic of a paper is often expressed in the first sentence. To quote from some classic UGa ACMC research reports:

“This paper presents a formalism and an  $n^2$ -time parsing algorithm for unification-based dependency grammars” (M. Covington, 01-0022).

“Our tests show that the current CYBERPLUS is not an appropriate machine on which to implement Prolog or similar logic programming languages” (R. E. Stearns and M. Covington, 01-0019).

This reflects a general principle: *get to the point early*. State your conclusions at the beginning, and then state the reasoning that leads up to them. Never leave the reader wondering where you are heading with an idea.

Outline the whole paper before you write it. If you find this difficult, make an *unordered* list or collection of ideas you want to include, and then sort it.

The first paragraph of a paper is the hardest to write, and it's a good idea to try writing it — or at least sketching it — long before you write the rest of the paper. Often, once you compose the first paragraph, the whole paper will fall into place.

You *do not need* a long introductory section. Many term papers wander around for a few pages before they reach the main point. Don't do this. If you have an introduction (necessary in a long paper), it should be an overview of the paper itself, not a disquisition on other “background” topics, nor a record of everything you looked at while starting to research the topic.

You *do not need* a “conclusions” section at the end unless the paper either (1) reports an experiment or survey, or (2) is rather long (thesis-length or more). When you come to the end of your argument, stop, ending, if possible, on a general point.

## 5 How to write clearly

Most people can write very clearly if they know what they are trying to say. If your paper is well planned, it should not be difficult to write.

When writing the first draft, do not worry about spelling, punctuation, or precise wording. Just get everything down on paper. Then go back and improve the organization and wording. Finally, once all the information is in place and clearly presented, check spelling, punctuation, and grammar.

Write as simply as possible. Your subject matter is complicated enough on its own merits; your task is to make it clear, not obscure. Use short words.

When you are explaining anything, ask yourself what the reader wants to know, and make the answers to those questions prominent. (Imagine a FAQ file about your subject; what would be in it?) Don't simply write down all the information you want to present in an arbitrary order (or even in a well-organized order that doesn't put the emphasis on the reader's needs).

You can improve your explanations by getting other people to read your paper and insisting that they tell you whenever they find it obscure. Make sure all of the steps in your logic are clearly presented and that you create a clear image in the reader's mind.

Formal notations (formulas, program listings, etc.) do not speak for themselves. They must be accompanied by plain-English explanations of how the formula or program contributes to your point.

At the same time, don't try to translate Prolog or algebra into English. When you need listings or formulas, use them.

Divide the paper into sections with section titles. Tell the reader how you are organizing your work: "This section will review previous results... There are three important results from earlier work... First... Second... Third..."

Say "I" if you mean "I." That is, do not call yourself "the present author" or use other awkward phrases. On the other hand, do not write about your personal experiences when you are supposed to be writing about your subject. Experiments are usually described by saying what was done, not who did it, because it doesn't matter who did it.

Notice the large number of one- and two-syllable words in this guide; emulate it. Do not imitate most of the scientific papers you read; they are not particularly well-written.

## 6 How to find sources

A crucial property of scholarly writing is that scholars *trace each idea to its source*. If you are writing about Weizenbaum's ELIZA program, you must use the actual report that Weizenbaum published. You are welcome to use later, secondhand accounts of it in addition to the original, but not in place of it.

Many bibliographical aids are available on line and in the Science Library.

In AI, however, your main problem is likely to be *too many sources* of information, not too few. How do you find the information you need without wading through fifteen popular articles and web pages that all say the same thing? The key is to distinguish between:

1. Books and articles that report original research (e.g., articles in *Artificial Intelligence* and *Journal of the ACM*), along with articles of this type published by researchers on their web pages;
2. Books and articles that survey research from a scholarly point of view (e.g., articles in *AI Magazine* and *Scientific American*);
3. Popular accounts designed for audiences other than working researchers (e.g., articles in *PC World*, *Popular Science*, and hobbyist web pages).

As far as possible, use (1) and (2); that is, use researchers' own accounts of their work. A type (3) source is acceptable if it is actually the only place an important piece of information is available.

## 7 How to cite sources

Your paper will consist of:

- The text, with references such as “(Chomsky 1957:192)” marking the information you have obtained from others;
- A *reference list* or *bibliography* at the end.

In the sciences, we normally *do not* use footnotes for citations. Nor do we use Latin abbreviations such as *ibid.* or *op. cit.*.

Cite sources by *author*, *date*, and, where applicable, *page number*. Give the page number if you are quoting exact words or referring to a specific page; otherwise leave it out.

Put the date and page number in parentheses. The author's name may or may not be in parentheses depending on whether it forms part of the sentence. Examples:

Chomsky (1957:23) showed that natural language cannot be generated by a finite-state Markov process.

“Natural language cannot be generated by a finite-state Markov process” (Chomsky 1957:23).

Markov processes are not powerful enough to generate natural language (Chomsky 1957:23).

Some scientists prefer to cite sources by number ([1], [2], etc.) rather than by name and date. I don't like this because it makes it hard to recognize authors. Chomsky is Chomsky wherever you meet him; if he is [24] in one paper and [35] in the next, you may never be aware of his identity.

The reference list at the end of your paper tells what the names and dates refer to. *It lists only the sources that are actually cited in the text*; it is *NOT* a list of everything you looked at while preparing the paper.

Here are guidelines for the content of reference lists:

- Everything has an author, date published, and information about where to find it. Even the most unusual sources can be cited if you keep this in mind.
- Authors' names are given as in the original source, but with the last name first.
- Titles of books (in English) are capitalized like sentences.
- Titles of series and of journals (in English) have every important word capitalized. An “important word” is one that can be spoken with full stress.
- In German, all nouns are capitalized. In most other languages, titles are capitalized like sentences.
- Always give complete numbers (e.g., 401–420, not 401–20).

## 8 What to cite

Cite all ideas or facts that can be obtained from only one source, or that reflect the opinion of one person, whether or not you are quoting the author's exact words. If you quote exact words, use quotation marks.

Do not cite every article that you looked at while preparing your paper. Cite only those that actually provided essential information.

Do not give citations for “common knowledge” (information that can be obtained from multiple sources that do not cite each other).

## 9 Examples of citations

The citation format used here is known as *Cambridge, Chicago, or Harvard style* and is quite similar to APA (American Psychological Association) style. In general, entries contain the same information as in APA style, but:

- Authors’ first names are given in full, not reduced to initials unless the author gave only the initials in the first place. (Why throw away information that might be useful?)
- There is no period after the parenthesized date.
- Some other details of punctuation are slightly different.
- Names of states are given in full (*Massachusetts, New Jersey*) or as traditional abbreviations (*Mass., N.J.*) rather than ZIP code abbreviations (*MA, NJ*).

### **Book:**

Chomsky, Noam (1957) *Syntactic structures*. The Hague: Mouton.

### **Book with 2 authors:**

Aho, Alfred V., and Ullman, Jeffrey D. (1972) *The theory of parsing, translation, and compiling*. Englewood Cliffs, N.J.: Prentice-Hall.

### **Book with 3 or more authors:**

Partee, Barbara H.; ter Meulen, Alice; and Wall, Robert E. (1990) *Mathematical methods in linguistics*. Dordrecht: Kluwer.

### **Book in a series:**

Blaser, A. (1988) *Natural language at the computer*. Lecture notes in computer science, 320. Berlin: Springer.

**Article in a journal:**

Rapaport, William J. (1986) Philosophy of artificial intelligence.  
*Teaching Philosophy* 9.2:103–120.

Here “9.2:103–120” means “Volume 9, issue 2, pages 103 to 120.” If the page numbering does not start afresh in each issue, do not give the issue number.

Note that you must *list all the pages the article occupies*, not just the pages you cited.

**Article in a book:**

Rapaport, William J. (2008) Philosophy of artificial intelligence. In John Doe and William F. Nusquam, eds., *Studies in the philosophy of artificial intelligence*, pp. 103–122. Dordrecht: Reidel.

**Article in proceedings of a well-known conference:**

Doe, John P. (1988) Prolog optimization systems. *Proceedings, AAAI-88*, 128–145.

If the conference is not well-known, handle the proceedings volume like a book of articles, identifying its editors and publisher.

**Unpublished paper:**

Doe, John P. (1964) *Giant computers of the future*. Department of Computer Science, University of Tasmania.

**Paper retrieved from a web site (and not otherwise published):**

Doe, John P. (2008) *Giant computers of the present*.  
<http://www.utan.edu.au/cs/doe/giant.pdf>.

APA’s preference is to also give the date on which the item was retrieved.

**Reprinted article:**

Doe, John P. (1987) Prolog optimizers. Reprinted in L. C. Moe, ed., *Prolog optimization*, pp. 101–105. New York: Columbia University Press, 1998.

You can get more information about Chicago/Cambridge/Harvard scientific citation format from the *Chicago Manual of Style*, available in most libraries. Do not expect different style guides to agree 100%; you can get away with following any of them as long as you are consistent.

## 10 How to type a reference list in L<sup>A</sup>T<sub>E</sub>X

L<sup>A</sup>T<sub>E</sub>X expects you to be citing references by bracketed numbers. Here's an example of how to trick it into printing a reference list without them. The key idea is to give zero width for the identifier (that's the meaning of {} in the argument list) put the author's name in place of the identifier; it is simply the first thing in each entry.

```
\begin{thebibliography}{}
\raggedright

\item[Covington, Michael A.] (1994)
\emph{Natural language processing for Prolog programmers.}
Englewood Cliffs, N.J.: Prentice-Hall.

\item[Weizenbaum, Joseph] (1976)
\emph{Computer power and human reason.}
San Francisco: Freeman.

\end{thebibliography}
```

## 11 “Cited by” references

Sometimes you will be unable to obtain an important source, but another source will tell you what is in it. In that case you must cite it secondhand. You can identify it as “cited by” when you cite it:

People started telling their innermost secrets to the computer  
(Weizenbaum 1964, cited by Doe 1988).

In this case Weizenbaum 1964 is not in your reference list.

Or you can put it in your reference list and identify the secondhand citation there:

Weizenbaum, Joseph (1964) An experience with ELIZA. Cited by Doe  
(1988).

## 12 Grammar and punctuation

Scientific journals (as well as thesis committees and employers) want *perfect* spelling and punctuation in papers submitted to them; so do we.

### 12.1 Parentheses

Parentheses enclose material that can be left out. If the parentheses and all the material within them were omitted, the sentence should still make sense and be correctly punctuated. Thus:

RIGHT: This is a sentence. (So is this.) This is, too.

WRONG! This is a sentence (So is this.) This is, too.

RIGHT: This is a sentence (and this is more of it). This is, too.

### 12.2 Commas

Commas *do not* denote pauses in speech. Their purpose is to make the meaning clearer. Use commas to keep sentences from being misunderstood. Often — not always — you will end up putting them where you would pause if speaking.

Commas are often used like parentheses to mark material that can be left out. Thus:

Chomsky, who invented transformational grammar, was a Harvard Junior Fellow in the early 1950s.

(Chomsky is Chomsky whether or not we know he invented TG.)

The student who left his books here should come and get them.

(No commas here because we cannot leave out “who left his books here;” if we did, the sentence would not reveal which student was meant.)

### 12.3 *Who* vs. *whom*

The difference between *who* and *whom* is exactly the same as the difference between *he* and *him*.

For *whom* does the bell toll? For *him*.

*Who* rings the bell? *He* rings the bell.

## 12.4 *That vs. which*

As we saw in the comma example, there are two kinds of relative clauses: *restrictive* clauses, which limit who or what is being talked about, and *non-restrictive* clauses, which merely add further information to something that has already been identified.

Nonrestrictive clauses begin with *who(m)* or *which*, but not *that*. They are set off by commas. Thus:

Ronald Reagan, who was our oldest President, is a Republican.

Blair's Theorem, which is also known as Murphy's Law, states that. . .

Restrictive clauses begin with *who(m)* or *that*. Some grammarians say that a restrictive clause cannot begin with *which*; others allow it. Thus:

A man who used to be an actor became President.

The theorem that [which?] I want to prove is. . .

## 12.5 Possessives

Possessives of nouns, singular or plural, are formed by the following sure-fire procedure:

1. Add an apostrophe.
2. If an *s* is not already present, add it.

Thus:

singular *boy*, possessive *boy's*  
plural *boys*, possessive *boys'*

singular *lady*, possessive *lady's*  
plural *ladies*, possessive *ladies'*  
singular *man*, possessive *man's*  
plural *men*, possessive *men's*  
singular *Athens*, possessive *Athens'*  
singular *Mars*, possessive *Mars'*

Add 's after an *s* if it forms a separate syllable: *Charles's*.

## 12.6 *Its* vs. *it's*

The procedure above does not apply to pronouns.

*Its* and *it's* correspond exactly to *his* and *he's*.

That is, *its* is possessive, and *it's* means 'it is.'

## 12.7 Spacing after punctuation

In English, the marks ( [ come at the beginning of a word and are preceded by a space, thus:

one (two)  
one [two]

The marks . , : ; ) ] come at the end of a word and are followed by a space, thus:

One. Two.  
one, two  
one: two  
one; two  
(one) two

This is quite different from computer programming languages.

Commas, periods, colons, and semicolons "slide under" quotation marks in order to adhere to the preceding word:

“Come in,” he said.

Sometimes, when dealing with computers, we have to avoid this practice.

*Conventional punctuation:* Type ‘go,’ and then proceed.

*In computer manuals:* Type ‘go’, and then proceed.

Here the problem is whether the user should type the comma. It would be better to rearrange the sentence or simply leave the comma out.

There is a space after the periods in people’s initials:

M. A. Covington      *not*      M.A. Covington

However, there are no spaces in other abbreviations:

P.O. Box              F.B.I.              Ph.D.

**Extra space after periods:** Conventionally, the period after the end of a sentence is followed by two spaces, not just one. However, with word processors, this can sometimes produce awkward results, and you are encouraged to put only one space after every period.

L<sup>A</sup>T<sub>E</sub>X thinks all periods mark the ends of sentences, and puts extra space after them, unless you tell it otherwise. The L<sup>A</sup>T<sub>E</sub>X manual tells you how to deal with this problem. The simplest solution is to put the command `\frenchspacing` near the beginning of your document. Then there will be only one space after each period, regardless of its position.

## 12.8 Names of computer languages

Do not write a word in all capitals unless it is actually an abbreviation or is normally written that way. Thus:

DOS  
UNIX  
Windows      *not* WINDOWS

Prolog     *not* PROLOG  
BASIC *or* Basic  
LISP *or* Lisp

Programming languages designed before about 1975 usually have all-upper-case names.

## 12.9 Special typography

Underlining is not used in scholarly papers. In high school, they told you to underline things that should actually be in italics.

*Italics* are used for:

- Titles of books (*Computer power and human reason*);
- Titles of journals (*Journal of Irreproducible Results*);
- Foreign (human) words and linguistic examples (Spanish *ayer* ‘yesterday’);
- Emphasis.

Computer type (`\texttt`, `\verb`, and `verbatim`) is used for *anything in a computer language that is quoted in English context*. Just like italics, computer type is used to set off things that are not English. Be sure to use it consistently.

**Boldface** is sometimes used for new terms being defined.

SMALL CAPITALS are sometimes used for new terms being defined.

‘Single quotes’ are used for meanings. For example, in Spanish, *ayer* means ‘yesterday.’

“Double quotes” are used for quoting someone’s exact words.

- In Word, it is acceptable to use either "straight quotes" or “smart quotes,” but do not vacillate between one usage and the other.
- In `\LaTeX`, it is very important to distinguish opening and closing quotes in English:

“right”      ”wrong”      “wrong“      ”wrong“

This applies to *English and other human languages*. Always type *computer languages* exactly the same way in L<sup>A</sup>T<sub>E</sub>X that you do on the computer. Thus:

```
['This', 'is', 'a', 'tokenized', 'sentence', 'in', 'Prolog']
```

If you put `\usepackage{upquote}` near the beginning of your document, and if `upquote.sty` is installed on your computer (as it should be), then your `verb` and `verbatim` single quotation marks will look the same way in your L<sup>A</sup>T<sub>E</sub>X document as on your keyboard, and you'll get this:

```
['This', 'is', 'a', 'tokenized', 'sentence', 'in', 'Prolog']
```

## 13 Physical format of the paper

In L<sup>A</sup>T<sub>E</sub>X, use the `article` document class. Otherwise, start your paper with its title, your name, and other local information, something like the following:

Artificial Brains

John R. Doe

Artificial Intelligence Center  
The University of Georgia

July 1, 2004

The academic address is important because you may exchange copies of your work with people elsewhere. Also, when you use laboratory facilities, courtesy dictates that you indicate, on the paper, the institution at which it was written. (Your student ID number is not needed; we know you by name!)

Double-spacing, required on typewriters, simply wastes space in academic work. I suggest  $1\frac{1}{2}$ -spacing your lines if not using L<sup>A</sup>T<sub>E</sub>X.

## 14 The most common word processing errors

### 14.1 Microsoft Word

- Absence of page numbers.
- Unequal amounts of spacing between different words.
- Inconsistent layout (margins or type size changing for no apparent reason).
- Failure to italicize mathematical variables.

Write  $a = b + c$ , not  $a=b+c$ .

- Awkward handling of illustrations. (I strongly recommend L<sup>A</sup>T<sub>E</sub>X for its ability to handle floating figures and tables. “Floating” means that they float to the top of the page rather than interrupting the text at a fixed position.)

### 14.2 L<sup>A</sup>T<sub>E</sub>X

- Listings or other material that runs off the right-hand edge of the page. If you tell L<sup>A</sup>T<sub>E</sub>X to go past the edge of the paper, it will do so. Look at your document and make sure it’s OK. An “overfull hbox” error message means something didn’t fit within the margins.

You can deal with excessively wide program listings like this:

```
{\small\begin{verbatim}
your material goes here
\end{verbatim}}
```

In extreme cases you may have to say `footnotesize` rather than `small`.

- Excessive space before a footnote number. Remember that there must not be a space between the `\footnote` command and the preceding word. Remember also that footnote numbers go after commas and periods, not before them.

- Excessive space after periods that are not ends of sentences. See the L<sup>A</sup>T<sub>E</sub>X manual, or simply give the command `\frenchspacing` near the beginning of your document, and the problem will go away.
- Failure to distinguish text italics (`\emph`) from math mode (`$. . .$`).

These are italics: *See the difference?*

This is math mode: *Seethedifference?*

- Mistaken handling of the following characters:

<i>Character</i>	<i>Typed as</i>
&	<code>\&amp;</code>
%	<code>\%</code>
\$	<code>\\$</code>
~	<code> \$\sim\$ </code>
\	<code> \$\backslash\$ </code>

- Failure to distinguish opening and closing quotation marks in English text. (Do not try to distinguish them in computer languages if the computer language doesn't.) See page 15.
- Mixing up `tabular` (which arranges items in rows and columns) with `table` (which creates a floating, labeled place to display information). (Every `tabular` needs to be within a `center` or `flushleft`, which in turn may or may not be within a `table`.)

## 15 Proofreaders' marks

Figure 1 on page 19 shows some “proofreaders' marks” that are standard in the printing industry. I use them for marking typing errors. You will find that practically all typists and publishers understand them.

— END —

<i>Error marked</i>	<i>Corrected</i>	<i>Meaning of mark</i>
<u>english</u>	English	Upper case
E <del>g</del> lish	English	Lower case
<u>this</u>	<i>this</i>	Italics
<u>THIS</u>	THIS	Small capitals
<b>this</b>	<b>this</b>	Boldface
tihs	this	Transpose (swap characters)
th <del>i</del> s	this	Delete character
th <sup>i</sup> <sub>^</sub>	this	Insert character (the mark ^ is called a <i>caret</i> )
and <del> </del> this	and this	More space (vertical or horizontal)
and <del> </del> <sup>#</sup> this	and this	More space (vertical or horizontal) (In the printing industry, # often denotes space)
Mc <sup>^</sup> Arthur	McArthur	Less space (vertical or horizontal)
<u>this</u> stet	this	<i>Stet</i> = "let it stand" (do not make the change after all)
this and that	this and that	Align
¶ this. Next, we show that...		Start a new paragraph

Figure 1: Proofreaders' marks for marking typographical errors.