How Readers Annotate Textbooks and Manuals

Jakob Nielsen

How Readers Annotate Textbooks and Manuals

Jakob Nielsen

Aarhus University

Denmark

Abstract

An empirical investigation of how students annotate computer
science textbooks and programming language manuals is reported.
Also, a taxononomy of such annotations is presented. These
results may be used when designing systems for online represen-
tation of either computer documentation or literature in general,
including textbooks.

Keywords and Phrases
Note taking, Annotation, Online documentation, Study habits.

1. Introduction.

When reading textbooks or other literature, the reader often
makes notes or underlines important material. Underlining and
annotations in the margin of the text can easily be made in
present-day paper based books but are more problematical in com-
puter-based literature such as online documentation. Most
present systems (e.g. DOKUMENT (Girill and Luk 1983)) do not al-
low the reader to annotate online documents.

In the future much literature will exist in online form only, as
part of an online journal (Shackel 1983 and Senders 1984) or even
online textbooks. Therefore systems for such literature must in-
clude the possibility of individual annotations for each reader.
An interesting point about such computer based annotations is
that they may be automatically carried over into updates (i.e.
"new editions") of the text base, if the specific part of the
text they are commenting on is not changed.

However, several form of annotations exist, and it may be dif-
ficult for a computer system to support them all equally well.
To ensure a more firm understanding of the possibilities, we
present a taxonomy of annotations in section 2.

Of course, each person has his or her own style for making an-
notations, but several investigations have been made on which
study methods are most effective in general.

The advantage of using note taking is often emphasized, and (Kul-
havy et al. 1975) showed that students taking notes did in fact
have superior recall of a test text compared with students not

taking notes. However, other research (Hoon 1974) showed no sig-
nificant difference in comprehension between reading with note
taking, reading with underlining, and simply reading. It has been
hypothesized (Harken and Maland 1979) that the effect of a study
method is obscured by differences in subject's preferences for
each study method. So we should under all circumstances supply
the reader with a rich set of annotation possibilities.

It is also of interest how people actually annotate traditional
books when they do not take part of an experiment to determine
how they annotate. Therefore we have performed a more an-
trophologically flavored investigation of the notes made by com-
puter science students at Aarhus University in their own personal
books during normal study activities the year before the research
project. The results are reported in sections 3 and 4.

2. Taxonomy.

There are five major types of annotations possible in traditional
books. They are Highlighting, Structuring, Cross References, For-
mal Additions and Informal Additions. Each of these types are
further subdivided in the following.

Highlighting

Highlighting is the setting apart of some parts of the original
text from other parts of this text. It has traditionally been
done with underlining or overmarking by the use of "Magic Mar-
kers" in yellow or other colors. Highlighting usually serves to
distinguish parts of the text of percieved special importance or
to make keywords stand out for quick scanning of pages.

If highlighting is done by just one means it is <u>simple high-lighting</u>. If e.g. several colors have been used indifferently, we will also consider the resulting highlighting simple. But different types of highlighting may also be used purposely, resulting in either hierarchical or differential highlighting.

<u>Hierarchical highlighting</u> is the use of different types of high-lighting for different grades of importance. One could e.g. use a yellow Magic Marker for somewhat important parts of the text and a green one for very important parts.

<u>Differential highlighting</u> results when different types of information is given different types of highlighting. One could e.g. use red underlining for definitions and blue underlining for theorems in a mathematics textbook. Differential highlighting is thus superimposing a structure on the text.

A computer system would be suited for all kinds of highlighting, and it could include the possibility of hiding the highlighting when not needed so that it does not distract from the reading of the text. The user might have a means of denoting keywords and phrases and then having the computer show only those words and not the rest of the text, enabling a quick scan for some relevant keyword. Of course some text editing search operation could also be given the option of searching for only those occurrences of a word that were denoted keywords.

These observations hold also for hierarchical and differential highlighting in general. One could ask for only text that was highlighted with an "importance quotient" of more than, say 3 on a scale of 1-5. Or for only text marked as e.g. definitions.

Finally, we should note that computers will be able to display
highlighting using colors and that these colors may be changed
dynamically by the reader. One of the problems in writing with
color pens in paper books is that it is a write-once operation
that cannot be undone.

Structuring

Structuring is the explicit annotation of some structure in the
original text. A traditional structuring annotation is list num-
bering in the margin, perhaps using keywords instead of numbers.
This is used when the author is enumerating several points or al-
ternatives concerning some issue.

Also the reader may impose his or her own structuring by adding
margin keywords that do not form an enumeration list.

Finally the reader may note that only some parts of the text are
of relevance, and that only those parts should be presented the
next time if no counterorder is given. An example could be a
student only wanting to read those parts of a textbook that are
study assignments. One could also note that certain parts of the
text are useful for some specific purpose.

Cross References

A cross reference is a reference to some other text. We may
distinguish between references to some other part of the same
page, to some other page or section of the same book, and finally
to some other book.

Cross references within one page are probably best implemented
as graphical arrows directly pointing to the revant spot. This

is also how they are drawn in many cases in paper books. Other cross references within the same book may be implemented in a computer system as pointers enabling the user to get the system to fetch the reference automatically instead of having to leaf through a book. In this way the system would be an efficient support for browsing.

Of course references to other literature can only be automated when that literarture is also online and accessible through a shared hypertext system (Nelson 1973).

Formal Additions

A formal addition is an addition having some specific structure. There are several different kinds, including quantitative diagrams such as bar charts, formal diagrams such as flowcharts, and formal tables of numbers or words.

A computer system would have knowledge of the syntax of several different types of Formal Additions, perhaps including some not mentioned here. The system could aid the user in building the annotation by supplying a structure editor with this knowledge which would result in both a speed up of construction and modification and a more pleasingly looking result.

The system could also have knowledge of the semantics of the annotation, resulting in e.g. a table build using an underlying spreadsheet. The system could automatically include data from the underlying text in such a spreadsheet or in a quantitative diagram specified by the user. In the last case the reader would always see an updated e.g. piechart of exactly the numbers of interest when consulting the information.

Informal Additions

Informal additions are annotations without any specific struc-
ture. Typically they consist of natural language text or free
hand drawings. They may be subdivided into the following
categories:

Corrections of errors in the text (typos) are often needed in
traditional books, but ought not to be needed in computer based
literature that may be updated instantly by the author when an
error is discovered. Correction of information made obsolete by
the passage of time is presently also done by the reader but
should not be necessary in an online system.

Word explanation is needed when the author uses either little
known words or when the text is in some foreign language. A com-
puter system could include a dictionary to supply such ex-
planations automatically when needed and could keep track of the
words for which explanations are requested and then automatically
supply the explanation for the next few occurrences of those
words.

Exclamation and question marks in the margin are often used to
signify either agreement or disagreement with the author and also
to denote material that is importance or difficult to understand.

Answers to exercises are often written in margin next to the ex-
ercise for easy reference later on.

Interpretations and elaborations of existing text. This is a type
of annotation that is directly tied to some specific part of the
basic text and which could not be transferred to an updated ver-

sion. An example from a Pascal manual is the following: The text is "const maxint = 281474976710655;" and the annotation is "= 2.8 * 10^14".

Informal drawings of objects or e.g. of the partitioning of computer memory in an operating system.

New ~~independent~~ ~~text~~, including new examples. Such annotations may in many cases be carried over to new versions of the underlying text, but for safety the reader should in all cases be notified that the note was made in an earlier version. The reader may then cancel either the warning or the annotation or may edit the annotation to reflect the new situation.

Other Media

It should be noted that the subclassification of the two Addition types - Formal and Informal - is to a large extent due to the media used (text versus graphics). This is because a person will often use different media to express different thoughts, but it is also due to the different editing principles used to manage the information. This last difficulty may be partly removed by the use of a virtual editor (Maxemchuck and Wilder 1982) with a uniform handling of several types of information.

On a piece of paper the only possible media are text and pictures, but in a computer we may have also e.g. sound, animations and executable programs (including such semiprograms as spreadsheet models). Information in some of these new media should be added to the taxonomy under either Formal Additions or Informal Additions as the case may be.

3. Inspection of books.

Six graduate computer science students were asked each to supply
an undergraduate computer science textbook and a Pascal program-
ming language manual for investigation. Five students did so, but
three claimed that their Pascal manual contained no annotations
whatsoever and did not hand it in. Therefore five textbooks and
two manuals were investigated for annotations. The results are
reported in Table I.

It should be noted that the annotations registered were made by
the students as part of their normal study activities one or more
years before this research.

Furthermore is should be noted that there is no tradition in Den-
mark for students selling their old textbooks to second hand
bookshops. Therefore this inspection was possible, and therefore
Danish students may tend to make more annotations in their books
than students in some other countries where writing in the book
will diminish its resell value.

4. Questionnaire.

A questionnaire was handed out to all 107 students taking the
second year computer science course at Aarhus University and to
all 67 students taking the third year course (a total of 174
students). The questionnaire was totally anonymous and par-
ticipation in the survey was voluntary. 81 responses were
received (47 %).

To get a resonably high percentage of returned questionnaires it

Annotations per 100 pages

	Textbooks					Manuals	
Student No.	1	2	3	4	5	1	2
Highlighting							
- simple	2	–	–	52	120	–	–
- hierarchical	–	174	64	–	–	–	–
- differential	–	30	6	–	–	–	–
Structuring							
- list numbering/keywords	0	–	–	1	–	–	–
- margin keywords	–	–	2	–	0	–	1
- relevance (e.g. assignment)	–	–	3	5	1	–	–
Cross References							
- on same page	–	0	3	1	–	1	–
- to other pages/sections	–	2	3	3	–	1	–
- to external literature	–	0	–	–	–	–	–
Formal Additions							
- quantitative diagram	–	–	–	–	–	–	–
- formal diagram (e.g. flowchart)	–	–	1	1	–	–	–
- formal table	–	–	–	–	–	–	–
Informal Additions							
- correction of errors in text	–	2	–	–	–	3	1
- word explanation/translation	2	5	2	–	0	–	4
- exclamation/question marks	–	0	3	1	–	1	–
- answers to exercises	0	–	1	2	0	–	–
- interpretations/elaborations	2	3	10	10	1	1	1
- informal drawing	0	–	–	2	0	–	–
- new independent text	1	0	1	7	0	–	2

Table I.
This table shows the count of the different types of annotations
resulting from an inspection of students' textbooks and manuals.
The figure "0" signifies some use of an annotation type (but less
than 0.5 times per 100 pages), while a "-" means that the an-
notation type in question was not found in the book.

was decided not to ask the students to actualy count the number
of annotations in their old textbooks and manuals. Instead they
were asked to estimate whether they used each annotation form
"much", "a little" or "not at all". We defined "much use" as
being more than once for every 20 pages, which will correspond
to more than 5 annotations per 100 pages.

The results are given in Table II. Students were also asked to
what degree they made notes on separate sheets of paper that they
would keep for later use together with the textbook or manual.

Both the questionnaire reported in Table II and the inspection
of books reported in Table I show that manuals are significantly
less annotated that textbooks. There is however a rather strong
correlation of 0.52 between annotations in textbooks and an-
notations in manuals, so there is a tendency for some students
to be 'note-takers' in both textbooks and manuals and for some
other students to make very few notes at all.

The reason manuals are less annotated may be because students
feel that computer documentation is 'sacred' in some sense, so
that they do not have the same natural relationship towards it
as towards ordinary books. But it may also be simply because a
manual is used in another way than a textbook. It is used only
as a work of reference to look up answers to specific questions
while the textbook is read from cover to cover with an ex-
amination in mind.

Students were also asked whether they would agree to statements
claiming that it would be an advantage to have certain types of
text available as online literature. The results are given in

| | Use of the different annotations (in % of respondents) | | | | | |
| | Textbooks | | | Manuals | | |
	much	some	none	much	some	none
Notes on separate paper	18	31	30	–	21	75
Highlighting						
– simple	44	38	16	9	37	52
– hierarchical	12	12	74	–	7	90
– differential	12	10	77	3	3	93
Structuring						
– list numbering/keywords	14	40	46	–	11	86
– margin keywords	20	33	44	3	10	83
– relevance (e.g. assignment)	21	62	14	4	12	79
Cross References						
– on same page	16	44	38	–	21	77
– to other pages/sections	16	56	27	1	24	73
– to external literature	3	24	73	–	5	93
Formal Additions						
– quantitative diagram	3	24	73	–	–	98
– formal diagram (e.g. flowchart)	3	25	72	1	6	90
– formal table	1	20	77	–	3	94
Informal Additions						
– correction of errors in text	68	28	3	42	35	19
– word explanation/translation	36	52	11	16	54	26
– exclamation/question marks	31	44	24	10	20	67
– answers to exercises	11	57	30	–	6	90
– interpretations/elaborations	28	49	21	4	31	63
– informal drawing	21	48	30	3	14	80
– new independent text	9	41	49	1	19	78

Table II.
This table shows the restuls of a survey with 81 respondents.
The results are given in percent. "Much use" means that the an-
notation type in question is used by the respondent more than on-
ce every 20 pages on the average (corresponding to 5 or more an-
notations per 100 pages). "Some use" corresponds to less that
5 annotations per 100 pages in table I, while "none" corresponds
to a '-' in table I.

A "-" in this table means that nobody gave that answer.

The rows do not add up to 100% as some respondents did not answer
all questions.

Table III. Many people would like to have manuals available on-line, but almost nobody wants electronic fiction.

Students were asked to state their opinion on the advantage of online literature both with and without the possibility for the addition of the readers own annotations. There was a slight correlation of -0.31 between the number of annotations used by a respondent in paper books and the change in agreement when annotations were assumed to not be included in an online literature system. So people who themselves made many notes intheir paper books also had a stronger tendency to require annotation facilities in online books.

5. Conclusions.

The two surveys of students actual annotation pattern are in resonably good agreement and show a rather high use of high-lighting. This is the case even though (Senders 1982) claims that "the yellow Magic Marker is probably the greatest obstacle to efficient learning ever sold in a college bookstore, for the passive process of underlining is no substitute for the active one of taking notes".

The surveys show relatively few annotations of the Addition types. This is probably due to the limited space for individual notes in the margins of most books. People choose instead to make their notes on separate pieces of paper. This limitation will not exist in a computer based annotation system so we might expect an integration of all notes with the original text.

Hopefully online literature will therefore tend to alleviate the

Would online literature be an advantage?

	Agree much	Agree little	Neutral	Disagree little	Disagree much	DON'T KNOW
Textbooks						
with annotations	12	17	14	32	21	3
without annotations	6	4	12	24	51	3
Manuals						
with annotations	32	26	11	12	15	3
without annotations	14	27	15	11	30	3
Fiction						
with annotations	–	6	5	19	69	–
without annotations	4	5	3	12	73	1

Table III.

Answers to questions on whether the respondent would agree that
it was an advantage to have certain forms of text as online
literature. The questions were asked both for the case where it
was assumed to be possible for readers to add their own an-
notations and for the case where there was no such annotation
feature.

Percentages do not add up to 100% since a few respondents did not
answer these questions.

problem pointed out in the above citation by Senders.

The goal must be some kind of hypertext system (Nelson 1973)
where we can not only view our own annotations but also those of
collegues and others who have placed their comments in the public
domain.

Acknowledgement

References

Girill, T.R.; and Luk, Clement H. DOCUMENT: An interactive, on-
     line solution to four documentation problems. Commun. ACM
     26, 5 (May 1983), 328-337.
Hoon, Peter. Efficiacy of three common study methods. Psy-
     chological Reports 35 (1974), 1057-1058.
Kulhavy, Raymond W.; Dyer, James W.; and Silver, Linda. The ef-
     fects of notetaking and test expectancy on the learning of
     text material. J. Educational Research 68, 10 (July-Aug.
     1975), 363-365.
Marken, Richard; and Maland, Joan. Single-subject analysis of
     three study methods. Psychological Reports 44 (1979),
     765-766.
Maxemchuck, N.F.; and Wilder, H.A. Virtual editing: I. The con-
     cept. In: Naffah, N. (Ed.) Office Information Systems,
     North-Holland, Amsterdam 1982, 161-172.
Nelson, Theodor H. A conceptual framework for man-machine
     everything. Proc. AFIPS NCC Vol. 42 (1973), pp. M21-M26.
Senders, John W. Design criteria for the electronic journal.
     Proc. IFIP INTERACT'84 (London 4-7 September 1984), pp.
     1.401-1.403 (in the preprint edition).
Senders, Virginia. How to get to work. IEEE Potentials (Winter
     1982), pp. 6-9.
Shackel, Brian. Interview with Professor Brian Shackel on the
     BLEND electronic journal project. Computer Compacts 1, 3
     (June 1983), 149-153.

Author's present address: Jakob Nielsen, Computer Science
Department, Aarhus University, Ny Munkegade, DK-8000 Aarhus C,
Denmark.