The Structure and Layout of Technical Documents

Andreas Rüping

sd&m software design & management GmbH & Co.KG

Thomas-Dehler-Straße 27
D-81737 München, Germany

e-mail: rueping@acm.org

Introduction

Documentation is defined as the accumulation, classification, and dissemination of information [7]; a document consists of material so collected. A technical document is a document used in a technical domain. Technical documents include project reports, project proposals, product descriptions, market studies, research reports, and so on, in various technical disciplines. The background of this paper is software engineering, but the patterns may equally apply to other technical domains.

A technical document has one author or a small number of authors. However, a technical document is usually read or browsed through many times and, furthermore, is typically used by a number of readers. Therefore, effort placed on the quality of technical documents generally pays off well. Authors who place effort on the quality of their documents are amply repaid by many readers' ease.

Guidelines for the Readers

This paper presents a collection of patterns that help authors set up, develop, and improve technical documents. The focus of this paper is on the structure and layout of technical documents, and not on their contents. Guidelines concerning the contents of technical documents are to a large extent domain-specific and are beyond the scope of this paper. However, the fact that both a clear structure and a good layout support the comprehensibility of the contents is not to be ignored.

The process of writing and reviewing technical documents is also not within the scope of this paper. These issues are treated in a separate collection of patterns [9].

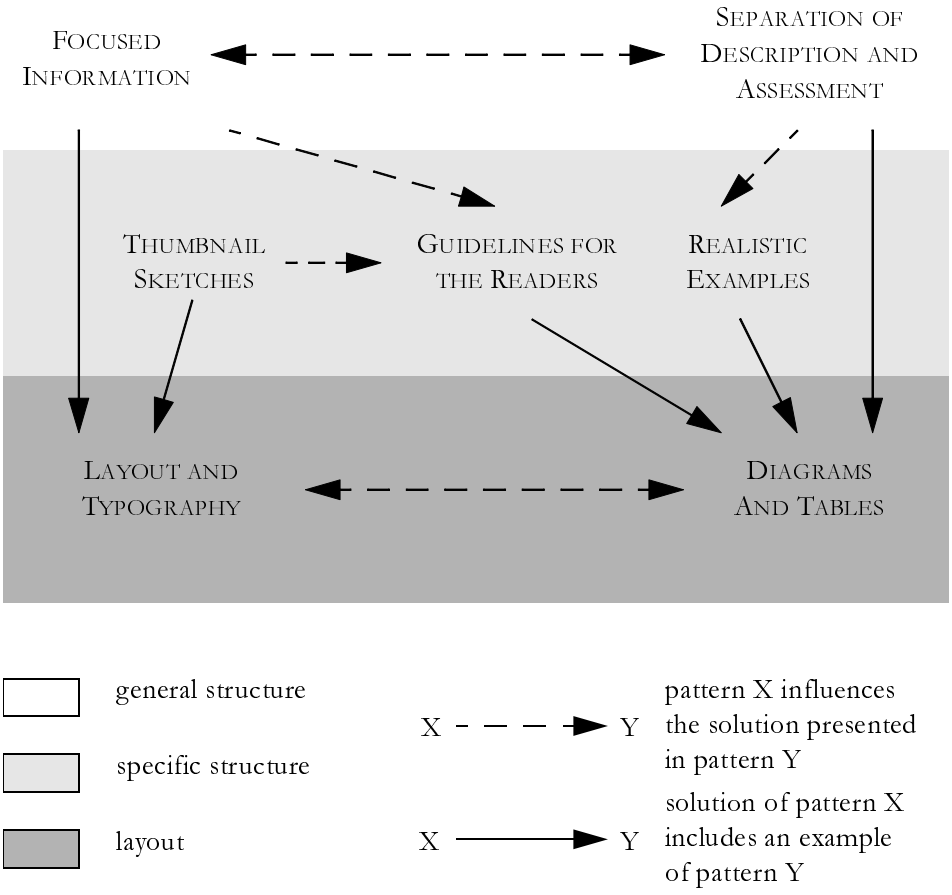
The patterns in this paper use a relatively easy pattern form consisting of only four parts. The problem section presents a question that typically arises when writing technical documents. The forces section describes the driving forces behind possible solutions. The solution section presents an answer to the question from the problem sections that resolves the forces as well as possible. The discussion section gives additional information, such as relationships to other patterns, and may also point out possible consequences of the pattern's application.

Copyright © 1998 by Andreas Rüping. Permission is granted to EuroPLoP '98 to make copies for conference use.

The problem section and the first paragraph of the solution section are sufficient to get an overview of each pattern. Together, they form so-called pattlets — thumbnails that give an overall impression of the actual patterns. The pattlets are printed in bold face. The other sections explain the rationale behind the patterns and let readers gain a deeper insight into the patterns.

The patterns presented in this paper can be grouped into three groups: patterns concerning the general structure of technical documents, patterns concerning specific structural aspects, and patterns about the layout of documents. Furthermore, the patterns are related to each other. The following diagram shows the overall organisation of the patterns and the relationships between them. The diagram can serve as a roadmap to the collection of patterns in this paper.

Overview



1 Focused Information

Problem

How can a document be structured, and how can authors determine which information belongs into a particular chapter or section, and which doesn't?

Forces

The chapters of a document are more easily understood when they are self-contained and when not much additional information is necessary. First, a clear focus on one topic makes a chapter more easily comprehensible. Second, a clear focus helps readers look up a certain piece of information in a document, since it is clear from which chapter this piece of information can be retrieved.

However, some general aspects may be relevant in the context of many different chapters of a document. Authors are easily led to scatter these general aspects over the entire document. Yet, this way a lot of information is presented redundantly. Redundant information is undesirable since it tends to be annoying to the readers and it makes a document difficult to maintain.

In 1956, the psychologist George A. Miller observed that, in general, people are able to identify and to memorize about seven pieces of information at one time [6]. This observation applies to the overall structure of document. Documents consisting of significantly more than seven chapters are generally difficult to handle. On the other hand, documents consisting of significantly less than seven chapters seem to be poorly structured.

Solution

Each chapter of a document should have a clear and identifiable focus on one particular topic.

There are some signals that indicate whether a chapter has a clear focus:

- A chapter should be aptly titled; a clear title suggests that the focus of the chapter is also clear.
- A chapter should begin with a statement that clearly says what the focus of the chapter is.

When chapters are further structured into sections, subsections etc., the focus should be refined into more specific aspects which still need to be clearly identifiable.

At the highest level, documents should have about seven chapters (plus or minus two). The chapters should be of similar length in order to give the entire document a well-balanced structure where information is evenly distributed.

Discussion

When trying to focus on one particular topic in a chapter or section, using a certain format can be helpful. The pattern form that is used in this paper is one example of a format that was designed to bring across information as focused as possible. The pattern form was first used by the architect Christopher Alexander [1][2]; several variations have recently been used in software engineering, too [8].

Since introducing a chapter for each major topic avoids redundant information to a large extent, references to other chapters become sometimes necessary. References between chapters indicate dependencies that should be described in the GUIDELINES FOR THE READERS. The references themselves can be emphasized in the text by a special font style (see LAYOUT AND TYPOGRAPHY); and hyperlinks can be used if the document is available online.

2 Guidelines for the Readers

Problem

How can readers be informed whether they should read a document and if so, how they should read it and which parts they should read?

Forces

In the age of information processing there is rather too much information than too little information. The problem is to filter out the information that is really needed. Potential readers must therefore be able to find out quickly what purpose a document serves and which parts they have to read.

Normally, various people are going to read the same document, but with different intentions. Some readers might only want to get an overview of the topic, some might be looking for some very specific detail, while others might want to read the document in its entirety.

Some technical documents are updated whenever the subject of documentation changes. Therefore, different versions of the same document may exist. When a new version is released, readers might not want to re-read the entire document, but the modified parts only.

Solution

A technical document should include guidelines for the readers. The guidelines should be either a special chapter or they should become part of the introduction.

All different types of readers should be able to use the guidelines as a commented table of contents — a roadmap for browsing through a document and for finding the parts that are of interest.

The guidelines should make the following clear:

- What is inside the scope of the document and what is beyond the scope of the document?
- How is the document organised?
- What are the relationships and dependencies between the different parts of the document? Which parts have to be read before which other parts?
- How can readers get a quick overview of the contents?
- Is the document completed or does it describe a work in progress? Are updates to be expected? If the document is an update of a former version, which parts have changed?

Discussion

A diagram (see DIAGRAMS AND TABLES) is often the method of choice to describe the overall organisation of a document as well as the relationships and dependencies between its parts.

Pointing out how readers can get a quick overview is particularly easy when the individual chapters of the document are provided with THUMBNAIL SKETCHES.

3 Thumbnail Sketches

Problem

How can readers get an overview of the topics dealt with in a document?

Forces

Technical documents can be rather long. It is likely that some readers will not have to read the entire document but only parts of it. These readers must be able to find out what each chapter is about.

More specifically, some readers will not plough through many technical details. Rather, an overview of the material presented is sufficient for them.

Some other readers will only look for a specific topic. In order to access the desired information, these readers must be able to find the chapter that includes the material they look for.

Solution

Each chapter should provide a thumbnail sketch — a brief description of the chapter's general goals as well as its major ideas.

There are two ways to set up thumbnail sketches:

- Each chapter can be provided with an extra paragraph that serves as an abstract of the chapter.
- Sometimes, one or two paragraphs from each chapter can be chosen to form the thumbnail sketches. In this case, it must be made clear to the reader which paragraphs form the thumbnail sketch so that readers know which paragraphs are sufficient to get an overview.

In any case, the thumbnail sketches must provide readers with the information that is necessary to decide whether or not they want to read the rest of the chapter.

Discussion

Choosing one or two paragraphs of each chapter as thumbnail sketches is particularly useful when all chapters follow the same format. Because the thumbnail sketches are at the same place inside each chapter, they can be easily identified. The pattern form used in this paper is an example. For each pattern, the thumbnail sketches consist of the problem and the solution. Readers can therefore quickly find out what the main idea behind each pattern is.

Thumbnail sketches are particularly important when readers navigate through a document a lot. Pattern languages are an example [5].

It can be desirable to use a special font style for thumbnail sketches, for instance italics (see LAYOUT AND TYPOGRAPHY). However, care should be taken not to use too many font styles in one document.

Since the thumbnail sketches help readers navigate through a document, they will probably be referenced from the GUIDELINES FOR THE READERS.

If a document is available online, thumbnail sketches can increase the efficiency, since readers can read the (relatively short) thumbnail sketches and then decide whether they want to load the (relatively long) chapter, too.

4 Realistic Examples

Problem

How can the material included in a document be explained in a comprehensible way?

Forces

Technical material is sometimes abstract and difficult to understand. Furthermore, not all readers of a technical document are necessarily experts in the field. Examples can help readers to understand the material.

Material is usually more successfully presented when it is accompanied by convincing examples. Readers are sometimes sceptical about the general conclusions drawn in a document. Examples can give evidence that what is said in a document is substantial information rather than the author's mere opinion.

However, toy examples can have quite the opposite effect on readers. When a major point is explained only with a toy example, readers are led to believe that the point is not substantial, and that a suggested solution might not work in practical cases.

Solution

Each major point in a document should be explained with an example. Whether a document describes existing work or work in the planning stage, all examples must be realistic.

- If a document describes an existing system or an ongoing project, then that is where the examples have to be taken from.
- If a document describes a system in the planning stage or a project proposal, the examples can be taken from the future system or from the project to be, provided that sufficient information for meaningful examples is available. Otherwise former systems or previous projects may supply realistic examples.
- If a document describes conceptual work or research, realistic examples from an application domain should be chosen which have been or will be used to validate the concepts.

Often, realistic examples are too large to be presented entirely. In this case it is fine to use only an extract or to ignore irrelevant details.

Discussion

Sometimes it is useful to present statistics taken from examples. Statistics are best presented using tables (see DIAGRAMS AND TABLES).

When using examples it is particularly important to have a clear SEPARATION OF DESCRIPTION AND ASSESSMENT. Typically, examples or figures taken from examples represent facts while the conclusions drawn from the examples are the author's assessment of the topic. Readers must be able to tell which is which.

5 Separation of Description and Assessment

Problem

How can authors ensure that readers can tell facts and observations from authors' opinions?

Forces

Often, technical documents present an analysis, describe a system design, or report on a project. The nature of many technical documents is therefore, to a large extent, descriptive.

Sometimes however, authors draw conclusions, make assessments, or even come up with personal opinions. Although this is sometimes inappropriate, it is necessary in some cases. Strategy papers or market studies in which the recommendation of one specific strategy or one particular product is what is expected from the document are two examples.

In the realm of journalism, it is a good rule of thumb to make clear whether an article in a newspaper or a journal presents facts or whether it expresses the author's opinion [4]. This rule of thumb applies equally to technical documents. It is not good style to irritate readers by confusing description and assessment; readers might doubt the contents of a document that seems to be suggestive.

Solution

Documents should clearly separate description from assessment.

- Description includes facts, observations, statistics, etc.
- Assessment includes comments, conclusions, interpretation, opinions, judgement, recommendations, etc.

A document must make this separation of its contents clear to the readers.

First of all, the separation should be reflected by the structure of a document. Individual sections should be assigned either description or assessment, but should not mix both.

The separation of description and assessment should also be supported by the use of the language. The descriptive parts of a document should not implicitly include a judgement or assessment of certain facts; adjectives such as *good*, *desirable*, *reasonable*, *useful* or *bad*, *problematic*, etc. should be used very carefully when describing facts or observations.

Discussion

The separation of description and assessment contributes to the general goal of presenting FOCUSED INFORMATION. Each chapter of a document is supposed to have a clear focus; not to confuse description and assessment is a precondition for this.

Sometimes layout techniques can be used to support the separation of description and assessment. Often, tables (see DIAGRAMS AND TABLES) can be used to present facts, such as figures, statistics, etc., while the corresponding text gives an interpretation and hence an assessment of the figures.

6 Diagrams and Tables

Problem

How can information be presented that is difficult to express in words?

Forces

Sometimes, one picture is worth more than a thousand words. Pictures can be a very helpful means of explanation. Yet, pictures lack the nuances that a language offers. Unless they are very complicated, pictures contain less detail than text.

Technical documents often contain material that is best presented in a table. Examples include systematic overviews, classification schemes, and statistical figures. However, the mere contents of a table provide only incomplete information since they need interpretation.

Diagrams and tables are two-dimensional; they add one dimension to the linear organisation of mere text. According to the psychologist George A. Miller, adding dimensions is one way to extend the natural limit of processing information [6].

Readers tend to get bored with long texts. Texts that include diagrams and tables are less monotonous. However, when there are too many diagrams or tables, and when there is no text to connect the diagrams and tables, the text flow suffers and the document is hard to follow.

Solution

Diagrams and tables should be used to describe things which would be more difficult to explain verbally.

- Diagrams should use only a small number of graphical elements in order to be easy to understand. A legend should explain the meaning of the graphical elements used.
- Tables should always include a heading line that names the table cells. Text should be used to define the precise meaning of the table cells and of their relationships.

In general, tables and diagrams should never complicate, but should always complement the text.

Diagrams and tables should be placed as close as possible to the text from where they are referenced.

There should not be too many diagrams or tables. A good rule of thumb is to have at most one diagram or table on one page. If more diagrams or tables become necessary, for instance when a graphical notation is defined, they are better placed in the appendix.

Discussion

The GUIDELINES FOR THE READERS can almost always make good use of a diagram to give an overview of a document and to explain the relationships between its chapters.

Having facts go into tables and the interpretation of these facts go into the accompanying text supports the SEPARATION OF DESCRIPTION AND ASSESSMENT.

7 Layout and Typography

Problem

How can authors ensure that a document looks good, and that its layout supports the structure of the contents?

Forces

The art of printing books dates back to the year 1436 when it was invented by Johannes Gutenberg (1397 - 1468). Typesetting rules have been developed and have matured since then. They say, basically, what letters look like, how letters are arranged into lines, and how lines are arranged into pages. Applying these rules makes a document look better and makes it easier to read [10].

In the computer age almost all texts, and technical documents in particular, are produced using desktop publishing systems. These systems let users define their own layouts. Users can choose the font style and font size, indentation, margin size, and so on. The flexibility offered by desktop publishing systems can be desirable since it makes a wide range of typographic styles possible.

However, this flexibility is easily misused. Clearly, not everybody knows about typography — in fact, book printing used to be a profession of its own not too long ago.

Solution

Technical documents should respect some general rules concerning layout and typography.

Many recommendations are available from the literature [3], including the following:

- At most three different font styles may be used in a document: a serif font, size 10 - 12, for normal text, a corresponding sans-serif font, size 14 - 18, for headings, and possibly an additional font for the likes of code fragments or formulae.
- Underlines should only be used when no other font styles are available; normally italics and small caps are preferable since they do not interrupt the text flow.
- The page geometry should give a document a well-proportioned feel; side margins are sufficient when they use about 40 to 50 percent of the space on a page.
- Normally, chapters should start on a new page. Extra space is necessary between the sections and subsections on one page. Space between paragraphs is optional when paragraphs are indented.

Most of these rules are well known from traditional books. Modern technical documents should use a similarly simple layout style as good traditional books do, enriched only by additional styles for those features that are specific to technical documents such as references to other sections, inclusion of code fragments, etc.

Discussion

When writing technical documents using desktop publishing systems, document templates should be used that incorporate the layout and typography rules.

Special font styles, such as italics or small caps, can be used to emphasize THUMBNAIL SKETCHES or references to other sections, made necessary for instance by the FOCUSED INFORMATION pattern.

Conclusions

The patterns presented in this paper do not prescribe a specific writing style or a standard template for technical documents. In fact, many authors are unable or unwilling to give up their individual writing styles. These patterns do not suggest that authors should do that. Rather, these patterns offer suggestions that can help authors improve their own, individual writing styles.

The seven patterns in this paper are certainly not new; they have matured as long as documents have been written. This paper can therefore serve as a guideline for authors to provide their documents with some well-established qualities, with the overall intention to make the documents comprehensible and easy to read.

Acknowledgements

Some of the patterns included in this paper first got my attention a couple of years ago when I worked at FZI / Forschungszentrum Informatik Karlsruhe (Research Centre for Computer Science). Thanks are due to my colleagues who influenced my thoughts in many discussions throughout several software engineering projects.

In its software engineering projects, sd&m software design & management, München encourages a so-called document-oriented methodology. Some of the ideas included in this paper can also be found in sd&m's documentation guidelines and have been successfully applied in several software projects.

Last but not least, thanks are due to Ken Auer who, as the EuroPLoP shepherd for this paper, provided many valuable suggestions for improvement.

References

[1] Christopher Alexander, Sara Ishikawa, Murray Silverstein. *A pattern language: Towns, buildings, construction*. Oxford University Press, 1977.

[2] Christopher Alexander. *The timeless way of building*. Oxford University Press, 1979.

[3] Theodore E. Conover. *Graphic communications today*. West Publishing Company, 1985.

[4] Theodore L. Glasser. "Objectivity and news bias", in: Elliot D. Cohen (Ed.), *Philosophical issues in journalism*. Oxford University Press, 1992.

[5] Gerard Meszaros, Jim Doble. "MetaPatterns: A pattern language for pattern writing", in: Robert C. Martin, Dirk Riehle, Frank Buschmann (Eds.), *Pattern Languages of Program Design*, Vol. 3, Addison-Wesley, 1997.

[6] George A. Miller. "The magical number seven, plus or minus two: Some limits on our capacity for processing information", in: *The psychological review*, Vol. 63, No. 2, American Psychological Association, March 1956.

[7] *The Oxford English Dictionary*, second edition. Clarendon Press, 1989.

[8] *Pattern Languages of Program Design*, Vol. 1-3. Addison Wesley, 1995 - 1997.

[9] Andreas Rüping. "Writing and reviewing technical documents", *Proceedings of the European Conference on Pattern Languages of Program Design (EuroPLoP)*, 1998.

[10] Suzanne West. *Working with style: Traditional and modern approaches to layout and typography*. Watson Guptill, 1990.