Parameter space exploration and tools for fast visualization in EM segmentation and MRI bias field correction in Slicer 3

From April 6th to Septemner 11th
At
Surgical Planning Laboratory (SPL)

SPL

Brigham & Women's Hospital
Harvard Medical School
75 Francis St.
Boston, MA 02115

Supervisors:

- Ron KIKINIS, Surgical Planning Laboratory, Harvard Medical School, Boston, MA - USA

- Sylvain JAUME, CSAIL, Massachusset's Insitut of Technologies, Boston, MA - USA

Referring teachers:

- Dominique MARATRAY, Institut Superieur de l'Electronique et du Numerique, Brest - FRANCE

- Christine CAVARO-MENARD, Universite de Angers, Angers - FRANCE

Student:

- Nicolas RANNOU, Institut Superieur de l'Electronique et du Numerique, Brest

Abstract

A couple of sentences on three or four lines to summarise your work.
This is a LaTeX template for undergraduate project reports.
Its detailed contents evolve to reflect FAQs.
Expectation-maximization is very popular for segmentation but it can be
tricky to understand and to use. A full description of the EMS algorithm
is done in this report. Different methods for fast parameters exploration
are described. As part of the research, preprocessing methods like MRI
bias field correction will be explained. The results obtained will be
presented. Following the new workflow should allow the user to segment
more datasets, more accurately.

Resume

Quelques phrases pour resumer mon travail.
C'est un template LaTeX pour les rapport.
Le contenu peut evoluer.

Contents

List of Figures

Chapter 1

Introduction

1.1 Context and motivation

Nowadays, medical image processing is becoming a major field of research in most of the laboratories. Indeed, because of the increasing complexity of the data they have to deal with, physicists need something to help. Help must be provided in many different ways. Before the surgery, to etablish a fast and accurate diagnosis. During the surgery to prevent the physicist from errors and to help him to proceed to more precise moves. After the surgery, to see if the surgery succeed, or to follow the pathology of a patent. Nevertheless, the informations bring to the physicist by the tool must be accurate, robust and provide a fast feedback.

In the context of pre and post operation, plenty of work as already been done. Thus, there is still a lot of work to achieve. Regarding data storage and exchange, the increasing among of informations NOUS OBLIGE to find other and more approriate methods. Another interesting contribution is images segmentation. New methods have to be developped for a better diagnosis, or to detect new pathologies. A lot of methods appears like levelset segmentation, region growing or texture based segmentation. Each one is adapted for a specific problem like vessels segmentation, tumors detection or SDFSF lungs detection. Another remarquable contribution is the segmentation based on expectation maximization (EM), which is very well suited for brain segmentation.

In the MR images segmentation context, the Surgical Planning Laboratory (SPL), Harvard Medical School, has developped an EM algorithm to segment brain's MR images. The results obtained are very good until we segment small structures. The approach used for the intensity inhomogeinities estimation appears less efficient for the particuliar purpose of large structures segmentation WHY. Moreover, the implementation is not widely used so far, regarding the complexity of the segmentation process. In this report we will present an approach to enhance the segmentation for

large structures, correctiong the intensity inhomogeinities in large structures and providing the end-user tools for an easier segmentation process.

1.2   Contents

The main body of this report is divided as follows.

Chap. 2 deals with the EM segmentation. Fundamentals will be reminded and the algorithm used will be described. Will will also present in this chapter the limitations of the current implementation. Chap. 3 explains our contribution. It explains the solution we choose to correct the intensity inhomogeinities. It also presents the tools we provide to the user for an easy segmentation worflow. Chap. 4 shows the results achieved and discussed about what have been done, and the future work.

Chapter 2

Expectation-maximization applied to brain segmentation

Here we get going with theory of the expectation-maximization, applied to brain segmentation and show firstly a simple approach of the problem. Then there's a more realistic approach of the problem with different constraints, followed by a presentation of the algorithm used in Slicer 3[1].

2.1 Presentation of the EM segmentation

The EM algorithm was explained in 1977 by Arthur Dempster, Nan Laird, and Donald Rubin[1]. They generalized and developped a method used in several times by authors, for particular applications. It is widely used to solve problems where data aisre "missing". The EM algorithm is an iterative algorithm which works in two steps: Expectation and Maximization. It can be use to solve a lot of image processing's problems like classification, restoration[3], motion estimation[2], etc.. Since the generalization of the algorithm, a lot of related papers were proposed. Most of them bring algorithms derived from the original one to adapt it to particuliar problems using additional informations. Nowadays, EM algorithms are become a popular tool for classification problems. It is particulary well suited for brain MR images segmentation. A lot of algorithms exist. They present complex frameworks using spatial information, neighborhood or intensity inhomogeneities to enhance the classification.

In the SPL, the algorithm developped uses spatial, structural and intensity inhomogeneities informations to segment the brain.

---

[1]open source software developped in the SPL for biomedical engineering purpose

2.2 Fundamentals

you should read this part if..

2.2.1 Statistical model used for the brain

2.2.2 Gaussian mixture model

2.2.3 Maximum likelihood model

Here's an inline formula: $E = mc^2$, and here's the same thing displayed:

$$E = mc^2.$$

A matrix needs to be displayed:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc,$$

and here's a result that's displayed and labelled:

$$\lim_{a \to \infty} \int_0^a \exp(-x^2) \, dx = \tfrac{1}{2}\sqrt{\pi}. \tag{2.1}$$

Note the punctuation in eq. (2.1) *etc*, which recognises that displayed equations are parts of sentences.

Note also that a smaller (`\textstyle`) size of numerical fraction is often useful in a displayed equation. The appropriate `\newcommand` called `\fr` is defined in the 'preamble', in file `report.tex`.

More complicated limits on `\int` (and `\sum`) need to be enclosed in braces $\{\cdots\}$.

In an inline formula, avoid ugly fractions such as $v = \frac{s}{t}$ and $\frac{dy}{dx} = x + y$. It looks much better to write $v = s/t$ and $dy/dx = x + y$. That is, keep `\frac` for display:

$$\frac{dy}{dx} = x + y$$

where it belongs.

Beware of inadvertent blank lines in your `.tex` file; they often creep in immediately after a displayed equation. A blank line gives a new paragraph, usually indented, which may not be what you want. You can avoid blank lines altogether by breaking paragraphs with `\par` instead, as done here.

2.3 Expectation maximization algorithm

first simplified then real

Simplified algorithm

Expectation maximization algorithm

Now an equation on several lines using `eqnarray`:

$$\begin{aligned} |\vec{A}|^2 &= a_1^2 + a_2^2 \\ &= \sin^2\theta + \cos^2\theta \\ &= 1 \end{aligned}$$

— where the star on `eqnarray*` suppresses all line-numbering[2].

Again, with just one line numbered for reference:

$$\begin{aligned} |\vec{A}|^2 &= a_1^2 + a_2^2 &\quad (2.2) \\ &= \sin^2\theta + \cos^2\theta \\ &= 1 \end{aligned}$$

— where `\nonumber` suppresses numbering otherwise. Now you can refer to equation (2.2).

If you want to number the equations in Chapter 2 as (2.1), (2.2), etc, then[3] put '`\numberwithin`' into *Google* or see [MG, Sec 8.2.14].

Notice that in equations you use `\lim`, `\exp`, `\sin` and `\cos` and not plain *lim*, *exp*, *sin* and *cos*. See [NSS, Sec. 3.3] for a list.

For a lot of complicated multi-line formulas it's better to use the more sophisticated display environments provided by the package `amsmath` — see Sec. 8.2 of *The Companion* [MG]. For instance the `cases` environment is used to get

$$\theta(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x > 0. \end{cases}$$

And if you need continued fractions, instead of

$$x = \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \cdots}}},$$

a better-looking result is

$$x = \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}},$$

for which `amsmath` provides the command `\cfrac` [MG, Sec. 8.4.2].

---

[2] And re-phrase material where any displayed equation splits between pages.

[3] Top tip: easily find out about any LaTeX command by putting it *with* its backslash (and perhaps the word 'latex') into your favourite search engine.

The package `amssymb` [MG, Chap. 8] extends the range of mathematical symbols (*eg.* $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{N}$ are available).

And its sister `amstext` provides the command `\text` to put words into a displayed equation

$$\text{like this: } E = mc^2.$$

Note that you need to put in by hand the spacing between the text and the mathematical symbols.

2.4 Expectation maximization algorithm used in Slicer 3

In this part, we present the algorithm which has been intergated in Slicer 3, and discuss of the limitations of this one. Some informations has been added to the algorithm to get the best and most automatic segmentation as possible.

2.4.1 Spatial information

2.4.2 Structure information

2.4.3 Intensity inhomogeneities information

Discussion

2.5 Nothing

Each chapter should end with a round-up of its contents and a link with the contents of the next.

After formulas, equations and theorems, the next important topics are graphics and tables.

Chapter 3

The contributions

Expectation-maximization is a very popular framework for different classification problems. It became extremly popular in segmentation of brain MRI during the last decade. However, it is not easy to understand the underlying theory.

In this chapter, a very simple approach of the theory will be given.

Finally, some solutions will be presented. Graphics files are inserted via the package `graphicx` which is loaded in the document preamble (in file `report.tex`).

Here we deal firstly with the inclusion of a single unlabelled diagram, and then with captioned, labelled, and multiple diagrams.

After that we consider tables of information.

3.1 MRI Bias Field correction

Here is a colour `.jpg` file, inserted in place without caption and without a number for cross-reference. It's suitable only for a simple self-explanatory diagram, used here and not again.

Bias correction is not efficient so we propose an other method, as part of the segmentation process.
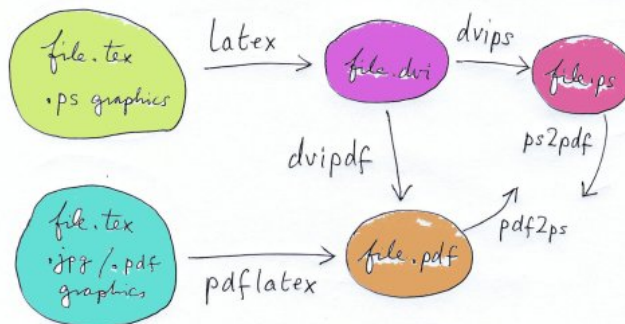
3.1.1 Interest

Kilian algorithm is only accurate for small region bias correction, because au the gaussian model he uses. Correction the bias in the whole volume first could enhance results.

3.1.2 Our approach

Gaussian distrib.... ITK filter Automatic

3.1.3 Results

corrected not corrected parameters explanation

It shows the normal routes from a LATEX source (`.tex` file) to `.ps` or `.pdf` output depending on the nature of the graphics files included[1]. *Confusion over these routes is a frequent source of grief.*

This template has diagrams and graphs in `.jpg` format[2] and so its source is compiled to a `.pdf` file with the command or menu button `pdflatex`.

Mathematical software such as *Maple* and *Matlab* can produce either `.ps` or `.jpg` diagrams, while cameras and scanners[3] usually produce `.jpg`. If accurate detail is vital, `.ps` is best.

With care, you can mix some graphics formats within a document. For instance, if you compile to `.pdf` then you can use an arbitrary mix of `.jpg`, `.png` and `.pdf` graphics, but not `.ps`. And a document to be compiled to `.dvi`, including `.ps` graphics, can include other formats too if they are explicitly given '*bounding boxes*'. But it's relatively tricky and you're better-off converting everything to one format.

Note that you can't include `.bmp` or `.gif` files at all, but conversion of graphics files — from `.bmp` to `.jpg` for example — is easily done with software like *PhotoShop* [PS] or *ImageMagick* [IM].

It's simplest to keep your image files in the same directory as your `.tex` files. Otherwise you can explore the mysteries of `\graphicspath` [MG, Sec. 10.2.5]. In any case, use only forward slashes in a directory specification. Names of directories and of graphics files *must not* include spaces!

---

[1]An IDE such as TeXnicCenter has menu buttons for each part of each route.

[2]File extensions must be `.jpg` and not `.JPG` or `.jpeg`.

[3]The figure itself is hand-drawn and scanned.

3.2 Intensity Normalization

3.2.1 Interest

usefull if you need to do a lot of segmentations enhance the rescaling..?

3.2.2 Our approach

background/brain

3.2.3 Our implementation

show results... corrected image...

It's awkward to flow text round a picture, but it can be done with packages such as `picinpar` or `wrapfig`, as *The Companion* explains [MG, Sec. 6.4.1].

And here's a `.jpg` image side-by-side with some explanation, using the `minipage` environment.

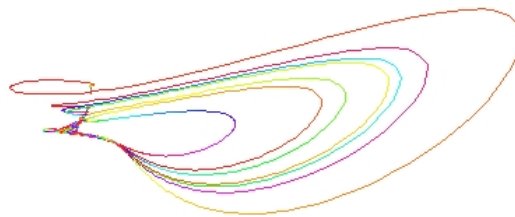Something more elaborate sometimes gives unexpected results.

Figure 3.1: This is a caption explaining the diagram clearly and fully.

That is, a picture 'floating' in a `figure` environment — with a caption and number as in Fig. 3.1 — might be placed by LATEX on a page well after the text meant to accompany it.

This happens if you have a relatively high density of graphics to text. You may be able to deal with it by re-distributing the pictures, changing the size of some of them, and the careful use of paragraph breaks. Otherwise you may resort to `\clearpage`, which forces insertion of floating objects waiting to go in.

The *Not so Short* book [NSS] has Sec. 2.12 on the issue of 'floating bodies' and *The Companion* has a whole chapter [MG, Chap. 6] on 'mastering floats'.

Note that inside a `figure` (or `table`) environment the `\label` must always *follow* the `\caption` — see [MG, p. 67].
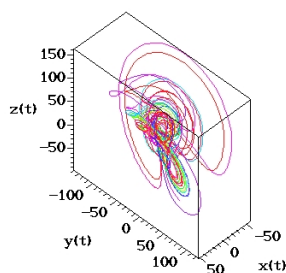

Figure 3.2: Caption to explain the diagram clearly and fully.
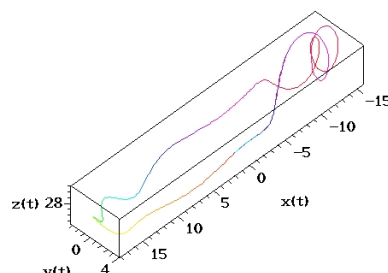
Figure 3.3: Caption to explain the diagram clearly and fully.

Fig 3.2 and Fig. 3.3 are two pictures floating side-by-side, each in a `minipage` and each with its own caption and label. This can be an economical way to insert multiple diagrams.

But don't get carried away and cram too many pictures together. Be aware of the size-reduction that's often necessary, and maintain legibility by choosing large enough font and weight for symbols, axis-labels and other text.

Figure captions should give plenty of information, because readers will first skim the Introduction and Conclusions — and then look at the pictures. If you want (really really want) to hook them, make sure that every important figure plus its caption is as self-explanatory as possible.

3.3   Class Distribution visualization

An important contribution is a tool which allows to visualize the classes to be segmented distribution, through 2 volumes.

3.3.1 Interest

As we discussed in the EMS algorithm requires variance and mean values for each class to be segmented. So far, the user defines these parameters by selecting samples on the dataset. exampleSDFSF A main issue is that user user can know if the pixels he selected are really representative or not of the class he wants to select. Thus, our objective was to find a way to give the user if his samples are reprenstative of not of a tissue.

3.3.2 Method used

Problem's description

Here is a brief description of the situation to let you understanding the problem. As input we have N images, K classes to segment, and for each class a mean value μ, a variance σ and covariance ζ. The covariance is estimated for each couple of images. We will come back on it later. where Using these informations, we had to find a way to give the user usefull informations about the class distribution.

Choices

Our first idea was to display gaussian curves in 3D, for each class. The X and Y axes would be the intensity ranges for the to images we want to use. The problem with this representation is that it does not contain any information about the number o f pixels . 3d too sophisticated...? To solve this problem, we decided to divide the problem in two parts. The first part will consist in creating a "temperature map" as background. The second part will consists in displaying ellipses over this map. This will give the user a good information about how good his selection was.

Background creation

First of all, the background has to be created. This background is a "temperature map". We can present it as a two dimensions array. The height and the lenght of the array depend of the intensity range of the 2 volumes we want to process. Then, we fill the array in an easy but useful way. Each case of the array contains the number of occurence of an intensity pair at the same position, through the 2 volumes.

Algorithm 1: BACKGROUNDCREATION(V1,V2)
---

define RangeI1 ← intensity range in V1
define RangeI2 ← intensity range in V2
define A(RangeI1,RangeI2) ← background
for each pixel **i** in V1
for each pixel **j** in V2
▷A(i,j)=A(i,j)+1
return A

Ellipse representation

We will first start we a brief rappel.. The covariance matrix is
Representing ellipse associated to this covariance matrix.
Maximum variance axe.Eigen values
Algorithm used. Big and small axes of the ellipse

3.3.3 Results

show result

First a simple table . . .

row 1	it's	just	as	easy
row 2	as	this	$E = mc^2$	what
row 3	could	be	simpler	?
row 4	as	easy	as	Π

Here it is again, now floating in a `table` environment, with a caption and a label to identify it as Table 3.1. The remarks about giving full information

row 1	it's	just	as	easy
row 2	as	this	$E = mc^2$	what
row 3	could	be	simpler	?
row 4	as	easy	as	Π

Table 3.1: Floating table, with a caption to explain it clearly and fully.

in figures and their captions apply equally to tables, of course. So do the remarks about unexpected placement of floating objects.

If you need a long table, occupying more than a page, and which can't reasonably be divided, then you need to delve into *The Companion* [MG, Chap 5] to find the secret.

3.4 Global Prior Estimation

The last contribution to the EMS is a tool which provides the user an easy and fast way to estimate the global prior weights (GPW).(ref ch2 ...)

3.4.1 Presentation of the problem

This contribution is usefull in many different ways. When you run the segmentation process,at the 6th step of the process, you have to provide to the algorithm an estimation of the GPW for each node in the tree. First of all if there are a lot of structures to segment, they user can spend a lot of time during this step. Indeed, for each part of this tree strcuture, they have to define the GPW. Moreover, the user may not know at all which weights to choose. This new approach will provide the users a good estimations of the weights to use. We must also keep in mind that the end users are physicists. They might don't understand what the parameters meanings and providing them a visual feedback could help them a lot.

3.4.2 Our approach

We divided the problem in two parts. The first part will be about providing the user a real-time feedback regarding the GPW estimation. The second part will consist in developping an algorithm which fills automatically the tree.

Fast user feedback

We can divide the feedback part in 3 steps: the histogram computation and utilisation, the multicolumn list and the labelmap generated. The histogram allows the user to manual segment classes based on intensity. The multicolumn list allows the user to change the order of the classes in the histogram. The labelmap provides to the user a visual feedback, base on the segmentation realized in the histogram. Using these three complementary tools, the user, even if he is not initiated can estimate easily, accurately and rapidly the GWP.
schema

Global priors evaluation

The algorithm used to estimate the weight of each node is iterative. It starts from the root and goes to the leaves. It evaluates the weight of the childs of the active node at each iteration. Here is a description of the algorithms used to compute the GPW of each node.
DEscirption algo 1

Algorithm 1: TREEWEIGHTESTIMATION(R, W)
---
define C = CHILD(R) ← set of childrens of root R

define LEAF(C) ← set of leaves of tree with roots C

define H ← set of structure-specific information defined by LEAF(C) for each leaf

update W in childrens of root R with the results of WEIGHTESTIMA-TION(C,LEAF(C),H)

for each node R' in CHILD(R) that is not a leaf

▷TREEWEIGHTESTIMATION(R', W)

Descripton algo2

The algorithm used estimates the global prior of the leaves of the current node, based on the number of pixel which belong to the child classes. This number of pixels is calculated from the segmentation computed in the histogram.(CF ..)

Algorithm 2: WEIGHTESTIMATION(C,LEAF(C), W,H)
---
define T ← set of total weight of leaves in LEAF(C). Leaves weights are contained in H

define E ← set of weight for each node of C

for each node of C

▷E=E+H : Get the total weight of each node

W=E/T

return (W)

3.4.3   Results and discussion

As we can see, the results obtained are good and accurate.The time of processing is fast and most of the people are now able to understand these parameters, even they are not familiar with EMS. This is an intuitive way to get an estimation of the GPW parameters.

Of course some sfsdfs must be done. If there is a strong bias in the images we process, the pixels of a same class will have different values. Then, a segmentation based on intensity will provide bad results. Moreover, if the user wants to segment 2 classes which have the same color, just using the spatial information for example, they won't be able to use it properly.Thus, the user must always keep in mind that it is just an estimation and that he has to check if the values are accurate.

Chapter 4

Results and discussion

The last chapter of a well-structured report reviews what's been done and mentions the main open questions. It's the part a typical reader looks at second, after the Introduction, and it should be just as appealing.

This template is provided only to help you start with LaTeX and to exemplify good structure. When you've got some practice, you'll want to alter it to suit your topic and your taste — and you are free to do that.

- `\documentclass[a4paper,twocolumn,10pt]{article}` gives an attractive alternative framework, for example.

  Or you may prefer `\documentclass{memoir}` which has an excellent manual [MEM] that includes a general introduction to typography.

- Try to resist the temptation to make margins narrower and lines longer — it may save a sheet or two of paper but it can make the text hard to read. Indeed, typesetting professionals quote a maximum of 66 characters per line as an ideal for easy reading [NSS, Sec. 5.2.2].

- proof-read carefully — not only to get the best from LaTeX but also to polish grammar, spelling and punctuation [ESL]. App. E alerts you to common pitfalls.

Good layout and use of LaTeX, plus fluent and grammatical writing, will together give a very good impression.

Top tip — write what you think is the final version, then put it away in a drawer and come back to it a week or so later. With fresh eyes you'll see many potential improvements. This works for anything you prepare, not just a project report. It needs forward planning of course!

Acknowledgements

Here's the place for thanks to anyone who particularly helped you. Don't go too far OTT — try to keep some dignity. This isn't the Oscars.

Bibliography

[1]   A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm", *Journal of the Royal Statistical Society: Series B*, vol. 39, pp. 1-38, November 1977.

[2]   Y. Weiss, "Bayesian motion estimation and segmentation", *PhD thesis*, Massachussets Intitute of Technology, May 1998.

[3]   R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence undersampled images", *IEEE Transaction on Image Processing*, vol. 6, pp. 1621-1633, December 1997.

[AC]   A Cottrell, *Word Processors: Stupid and Inefficient*,
                              `www.ecn.wfu.edu/~cottrell/wp.html`

[BR]   Visit `www.dur.ac.uk/library/using/guides/` and click on '*Writing your bibliography and citing references*'.

[ESL]   L Truss, *Eats, Shoots and Leaves*, Profile Books 2003
                              (ISBN `1-86197-612-7`).

[GRM]   M Goossens, S Rahtz and F Mittelbach,
                  The LaTeX Graphics Companion, Addison-Wesley, 1997
                              (ISBN `0-201-85469-4`).

[IM]   *ImageMagick*, `www.dur.ac.uk/its/software/application/...`
                              `...?application=ImageMagick`

[LAT]   *LaTeX stuff*, `maths.dur.ac.uk/Ug/projects/resources/latex/`

[MEM]   *Memoir document class*,
            `www.ctan.org/tex-archive/macros/latex/contrib/memoir/`

[MG]   F Mittelbach and M Goossens *et al.*, *The LaTeX Companion*,
                  Addison-Wesley, 2nd ed. 2004 (ISBN `0-201-36229-6`).

[MKT]   *MikTex Project Page*, `www.miktex.org`

[NSS] T Oetiker, H Partl, I Hyna and E Schlegl,
                          The Not So Short Introduction to LaTeX 2ε,
                          `www.ctan.org/tex-archive/info/short`

[PS]   *Photoshop*, `www.dur.ac.uk/its/software/application/...`
                          `...?application=Adobe+Photoshop`

[TXC] *TeXnicCenter*, `www.toolscenter.org`

[WDT] *WinEdt*, `www.winedt.com`

[WL]   *Wikibook on LaTeX*, `en.wikibooks.org/wiki/Latex`

[WO]   *Controlling widows and orphans*,
                          `www.tex.ac.uk/cgi-bin/texfaq2html?label=widows`

[WSH] *WinShell*, `www.winshell.de`

Prepared in LaTeX 2ε by RCJ

Appendix A

About references

Your report will be put together in your own style, mostly using your own words. Much of it will be standard material that you've read and digested, but you may have a fresh example, application or calculation that you've done yourself.

You must make clear what's not your own work by referring suitably to your sources — books or articles or web-pages, etc. Not to do so can count as plagiarism, which is cheating.

This appendix aims to amplify the advice given in the Library's guide [BR], which you should read.

A.1 Organisation

Follow the style of this template — that is, put an orderly list of your sources ('*The Bibliography*') at the end of the main text, and refer to ('*cite*') items on the list by suitable codes placed appropriately in the report's text.

For your bibliography the golden rule is that each listed item must give enough detail to allow readers to follow it up themselves and find the precise part of the book, article, web-site or whatever without ambiguity or delay.

For instance it's no use saying just '*The Times newspaper*' unless you also give the date and page, or just '*Wikipedia*' unless you give the complete URL of the specific page, . . . and so on.

It's always better to give slightly too much information — e.g. you might include the ISBN of a book too. A good level of information is exemplified here and recommended in the Library guide [BR].

If you need to cite different parts of e.g. a book for different things, then list it (say AB) in the bibliography and give the different citations as [AB, page 32] and [AB, Sec. 4.7] and so on. The LaTeX \cite command allows for this (look it up!).

Now you see why a bibliography list is better than lots of footnotes — it neatly allows such multiple citations. Too many footnotes make a mess.

- This bibliography uses short letter-code keys in alphabetical order. It's one of several standard possibilities [BR] and is often preferred for economy and because the letter codes[1] can be chosen to be helpful mnemonics. Numerical codes, for instance, can't.

- Long lines (with URLs usually [IM, PS]) may need to be split. The LaTeX hack used here to manage such line-breaks may not appeal to everyone.

A.2 Doing it

What things need a reference? The answer is — anything you didn't work out or invent yourself, but took from someone or somewhere else. You may do that, so long as you say so *and also make it clear you understand it.* Don't copy blindly! Here are some examples.

- '*the following explanation is taken from [LH]*' — if you've copied word-for-word from the article by Laurel & Hardy, which you list as item LH in the bibliography. Direct quotation should be used sparingly, and the text emphasised — perhaps by use of the `quotation` environment in LaTeX.

  Don't be tempted to copy from a book, or `ctrl-C/ctrl-V` from the web, *without* clearly admitting it. It's easy to detect, and it's suicide.

- '*the following proof is in many textbooks, eg. [LH, page 16]*' — if it is, and you can't think of a different proof, except perhaps for notation.

- '*the material in this section is based on Sec. 2.6 of [CL] and Chap. 3 of [SH]*' — if you've combined into your own words the account by Cagney & Lacey with that by Starsky & Hutch.

- '*the following proof is adapted from Chap. 4 of [SJ]*' — if (say) you've filled in the gaps in Smith & Jones' proof, or perhaps changed it from the case of general n to your case $n = 2$.

- '*these calculations were done with Matlab [MAT] using the m-file listed in App. C*' — where the reference MAT is to the *MathWorks* website.

- '*these calculations were done with the TISEAN package [TIS]*' — if you've downloaded this specialised package from a web-page whose URL (and author's name, if available) is item TIS in the bibliography.

- '*the graph is taken from Morecambe & Wise [MW, page 16]*' — if you've scanned the figure from their book. Similarly, if you've downloaded a `.jpg` file from a web-site, then give in the bibliography the

---

[1]Note the punctuation of 'letter-code' and 'letter codes' here.

23

URL and author (if known). Such citations could go either in the figure caption or in the associated text.

- '*the data are taken from [XY]*' — where XY gives the source of the numbers you've analysed. This might be a journal or an online data-bank, *etc.* Generally the numbers themselves should be left out — a summary table or a graph or two (plotted with *R* or *Matlab* or *Maple*) is often enough. Occasionally raw data might be supplied separately — e.g. on a CDrom.

Often your supervisor or someone else gives you something — such as a set of data, or a useful *Maple* worksheet, or help with a proof — when the appropriate form of bibliography item is '*Dr I Newton, private communication, April 2007*'. Similarly for printed course material — '*Dr I Newton, lecture notes for module MATH5033, Durham University, Epiphany Term 2007.*'

Generally, don't refer to things you haven't read. Textbooks may well cite the Serbian-language journal where the result you want was originally published. But you are not in the business of ascribing credit for discovery. In *your* report *you* must cite the book where *you* actually got it from and not give any possibly false impression that you fluently read technical Serbian.

Likewise, you may want to quote from Euclid or Archimedes. But cite the place where you found the English words with something like '*Archimedes said, 'Eureka!' (as quoted by [AB])*'.

Summarising — tell the truth (where you got it from), the whole truth (give full details), and nothing but the truth (you don't read Serbian).

Finally, if in doubt — ask your supervisor.

Appendix B

Long proofs

An Appendix is a good place to put lengthy proofs that must be included but would impede the flow if placed in the main text.

B.1   Proof of theorem 2.4

By inspection.

Appendix C

Computer programs

An Appendix is the place to list computer programs that you've written, using the **verbatim** environment [NSS, Sec. 2.11.4].

For example —

```
10 PRINT "HALLO SAILOR!"
20 GO TO 10
```

Appendix D

Using a PC

You may want LaTeX on your own computer.

A popular version of TeX/LaTeX for a Windows PC is '*MikTex*' [MKT]. *MikTeX* is free, and is available online for download [LAT] or locally on a DVD from `bob.johnson@dur.ac.uk`, room CM315. It's also installed on the ITS Networked PC Service under `Programs | Miscellaneous`.

To use *MikTeX* easily you need a dedicated editor or IDE[1]. A good one is '*WinEdt*' [WDT], which runs under all recent versions of Microsoft Windows and integrates well with *MikTeX*. *WinEdt* is free for 31 days; to use it thereafter costs about $40.

Completely free rivals include '*TeXnicCenter*' [TXC], which is provided on the *MikTeX* DVD and also set up on the ITS Networked PC Service under `Programs | Miscellaneous | MikTeX`.

Others include *eg.* '*WinShell*' [WSH], but are untried.

All these editors/IDEs have a familiar style of graphical user-interface with a toolbar and pull-down menus for all the common tasks involved in editing source files, running LaTeX and viewing the results.

---

[1]Computer scientists say 'Integrated Development Environment'.

Appendix E

Typography and grammar

Some errors crop up time and again, and a few of the commonest — mostly relating to use of LaTeX — are mentioned in the body of this template. Some others are given here, aiming to help you avoid them.

Layout errors include starting a sentence with a symbol, or leaving a single symbol that ends a sentence hanging alone on a line. Related to this are '*widows*' and '*orphans*' [WO]. Can you spot any in this template?

Grammar — illiteracy makes good mathematical work look like rubbish. Of course, everyone makes mistakes, but a final version with many small errors in fact signals '*I am ignorant and I just don't care*'. Read, mark, learn, and inwardly digest the wisdom of the great goddess Truss [ESL].

Apart from spelling blunders, the top five usual suspects in recent project reports are (in random order) —

- '*sentences*' with no verb;

- mis-use of apostrophes;

- mis- or non-use of hyphens;

- writing '*comprises of*';

- unpunctuated displayed equations.

A tip for proof-reading — a friend or relative may spot things you've become blind to. And remember in practise their are a massive number of errors in grammer and stile, and so fourth its only to easy to make that you're spell checker on it's own wont find (14 in this sentence).