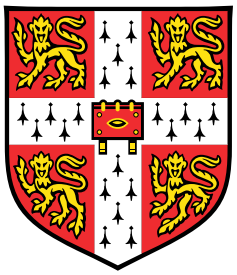
Writing your PhD thesis in LaTeX2e

Using the CUED template



Krishna Kumar

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

King's College

April 2017

I would like to dedicate this thesis to my loving parents ...

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Krishna Kumar
April 2017

Acknowledgements

And I would like to acknowledge ...

Abstract

This is where you write your abstract ...

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Wikipedia - the free online encyclopedia - is a web-based, collaborative, multilingual encyclopedia project supported by the non-profit Wikimedia Foundation. According with Alexa Internet, Inc.[1] the online enyclopedia is among the top five most visited worldwide websites.

Wikipedia is widely used as a reference source either for humans and for automatic processing language tools.

Some of its features that have made it a relevant online resource are:

- It is large size, with a constant growth rate,
- It has semi-structured content which is recognized as high quality,
- It is domain-independent,
- Is available in several languages,
- Is freely accessible, available for editing and use even offline.

Due to her characteristics it is considered as a very large lexical-semantic resource and has been used in Natural Language Processing tasks such as [10] :

- Information retrieval,
- Text categorization,
- Question answering,
- Automatic summarization,

[1] Alexa is an Amazon company that provides commercial web traffic data and analytics. http://www.alexa.com/topsites, December 2016.

- Named entity recognition,
- Wikification.

The present proposal is about automatic Wikification which is a task inspired in the manual Wikification process made by the Wikipedians [2]. These editors, do a wide variety of tasks, the one related to this work is the manually selection of relevant words or phrases in an article and link them to other Wikipedia articles whose titles correspond to these sentences.

The rest of the chapter will include the following topics:

1.1 Wikification

Lorem Ipsum is simply dummy text of the printing and typesetting industry (see Section 1.3). Lorem Ipsum [3] has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum [1, 4, 5].

The most famous equation in the world: $E^2 = (m_0c^2)^2 + (pc)^2$, which is known as the **energy-mass-momentum** relation as an in-line equation.

A *LaTeX class file* is a file, which holds style information for a particular LaTeX.

$$CIF: \quad F_0^j(a) = \frac{1}{2\pi\imath} \oint_\gamma \frac{F_0^j(z)}{z-a} dz \tag{1.1}$$

1.2 Why do we use loren ipsum?

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many

[2] The Wikipedians are the people who write, edit or translate articles for Wikipedia. They are volunteers part of the Wikipedia Community. https://en.wikipedia.org/wiki/Wikipedia:Wikipedians

web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

1.3 Where does it come from?

Contrary to popular belief, Lorem Ipsum is not simply random text. It has roots in a piece of classical Latin literature from 45 BC, making it over 2000 years old. Richard McClintock, a Latin professor at Hampden-Sydney College in Virginia, looked up one of the more obscure Latin words, consectetur, from a Lorem Ipsum passage, and going through the cites of the word in classical literature, discovered the undoubtable source. Lorem Ipsum comes from sections 1.10.32 and 1.10.33 of "de Finibus Bonorum et Malorum" (The Extremes of Good and Evil) by Cicero, written in 45 BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, "Lorem ipsum dolor sit amet..", comes from a line in section 1.10.32.

The standard chunk of Lorem Ipsum used since the 1500s is reproduced below for those interested. Sections 1.10.32 and 1.10.33 from “de Finibus Bonorum et Malorum" by Cicero are also reproduced in their exact original form, accompanied by English versions from the 1914 translation by H. Rackham

“Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum."

Section 1.10.32 of “de Finibus Bonorum et Malorum", written by Cicero in 45 BC: “Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?"

1914 translation by H. Rackham: “But I must explain to you how all this mistaken idea of denouncing pleasure and praising pain was born and I will give you a complete

account of the system, and expound the actual teachings of the great explorer of the truth, the master-builder of human happiness. No one rejects, dislikes, or avoids pleasure itself, because it is pleasure, but because those who do not know how to pursue pleasure rationally encounter consequences that are extremely painful. Nor again is there anyone who loves or pursues or desires to obtain pain of itself, because it is pain, but because occasionally circumstances occur in which toil and pain can procure him some great pleasure. To take a trivial example, which of us ever undertakes laborious physical exercise, except to obtain some advantage from it? But who has any right to find fault with a man who chooses to enjoy a pleasure that has no annoying consequences, or one who avoids a pain that produces no resultant pleasure?"

Section 1.10.33 of “de Finibus Bonorum et Malorum", written by Cicero in 45 BC: “At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat."

1914 translation by H. Rackham: “On the other hand, we denounce with righteous indignation and dislike men who are so beguiled and demoralized by the charms of pleasure of the moment, so blinded by desire, that they cannot foresee the pain and trouble that are bound to ensue; and equal blame belongs to those who fail in their duty through weakness of will, which is the same as saying through shrinking from toil and pain. These cases are perfectly simple and easy to distinguish. In a free hour, when our power of choice is untrammelled and when nothing prevents our being able to do what we like best, every pleasure is to be welcomed and every pain avoided. But in certain circumstances and owing to the claims of duty or the obligations of business it will frequently occur that pleasures have to be repudiated and annoyances accepted. The wise man therefore always holds in these matters to this principle of selection: he rejects pleasures to secure other greater pleasures, or else he endures pains to avoid worse pains."

Chapter 2

State-of-the-Art

Wikification its an automatic process inspired by *Wikipedians*[1] Wikipedians manually select words or phrases that are considered relevant in an article and link it to another Wikipedia article which title are closely related to this words and phrases.

According to [6] **Wikification** is "*the automatically extraction of words or terms that are the most importants in a document and for each of this keywords in an article identify the most appropiate article in Wikipedia*".

For humans this in an easy process, but it is a hardly process to get it done automatically [9]. A Wikipedian links articles from Wikipedia according to the knowledge he or she has from a topic, that is a tasks done by her or his personal apreciation about a concept they consider must be explain.

. . . and some more

In this chapter we will provide an overview of some relevant approaches related to Wikification Process

2.1 Relevant approaches related to Wikification Process

This work is related to the next natural language processing issues:

Itemize

- Word sense desambiguation problem

[1] *Wikipedian*, also known as editor, is a contributor from the Wikipedia Community, who volunteers to write and edit Wikipedia's articles. Anyone who edit or find something that can be improved in the Wikipedia articles can become a Wikipedian.

- Machine learning problem
- Keyword extraction
- Document and text processing

Some approaches have been proposed with big influence in the selection and methodology used for the Wikification process about to be describe in the next sections.

Wikipedia has become a very large and rich source of information and is the input for the next related work.

Wikipedia as a Dataset

One thing that must be considered when working with Wikipedia data, is dealing with her size, specially the English-language edition[2] with approximately 5,360,487 articles on the site (live count) is the biggest one with the most articles of any of the other Wikipedias. In October 2015, the combined text of the English Wikipedia's articles totalled 11.5 gigabytes when compressed and on November 1st 2015, the English Wikipedia announced it had reached 5,000,000 articles and ran a special logo to reflect the milestone.

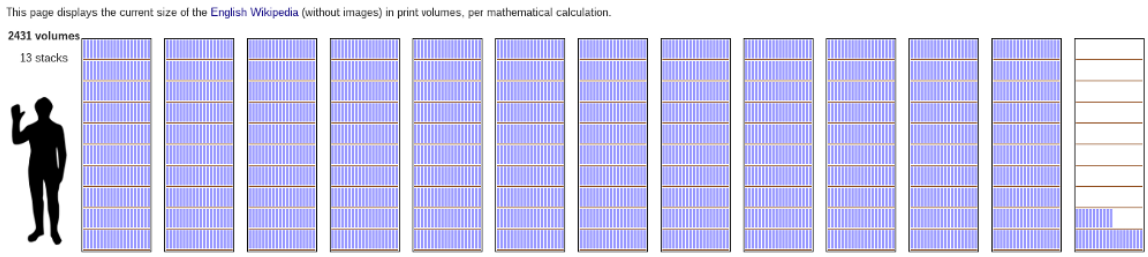


Fig. 2.1 Representation of the current size of the English Wikipedia (without images) in print volumes, per mathematical calculation. https://en.wikipedia.org/wiki/Wikipedia:Size_in_volumes

2.1.1 David Milne and Ian H. Witten - Learning to link with Wikipedia

cross-reference documents with Wikipedia by identifying significant terms within unstructured text using machine learning and enrich it with links to the appropriate Wikipedia articles.

[2]Size of Wikipedia, https://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia

They describe their work as a link detector and disambiguator, able to enrich any unstructured fragment of text which is automatically recognized and linked with structured knowledge, in this case with Wikipedia articles.

The uses machine-learning approach to disambiguate links for the wikification process.

For the experiments Milne and Witten, worked with a Wikipedia snapshot from November 20, 2007 that contained under two million articles. They selected articles containing at least 50 links, avoided list and disambiguation pages because they were considered not representative unstructured text.

A total of 700 articles were randomly selected and set aside for developing the disambiguation algorithm: 500 were used for training, 100 for configuration, and a further 100 for final evaluation.

The 500 training articles contain more than 50,000 links. Each link represents several training instances, they built an anchor tree with the possible senses, only one sense is a positive example and the rest are negative. As explained before, an anchor are phrases that link to a target.

They used the same approach than Mendelyan et al. to balance the commonness of a sense with its relatedness to the sorrounding text. The commonness or prior probability is computed by the number of times it is used as a destination in Wikipedia, but they compare each possible sense with its sorrounding context and also use every unambiguous link in the document as context to the disambiguation process. Each candidate sense and context term is represented by a single Wikipedia article, so they solve selecting the sense article that has most in common when comparing with all of the context articles, the relatedness.

They used a method they also already developed [7] to measure the semantic similarity of two Wikipedia pages known as Wikipedia Link-based Measure, which compares the incoming and outgoing links, expressed in the next formula:

$$relatedness(a,b) = \frac{\log(\max(|A|,|B|)) - \log(|A \cap B|)}{\log(|W|) - \log(\min(|A|,|B|))} \tag{2.1}$$

Where a and b are two articles, A and B are sets of all the articles that link to a and link to b and W is the set of all the articles in Wikipedia.

As mentioned before the approach use the information from the context to determine how closely related to the central thread the terms are, and is determine by two main features (1) link probability proposed by Mihalcea and Csomai that provides the commonness of each sense and (2) semantic relatedness. A third feature is involve —the context quality—which ss given by the sum of the weights that were previously assigned to each context term. This three features are used to train the classifier. A diference from other approaches in the desambioguation process is that Milne and Witten can be done in two ways:

1. considering each sense independently with an assigned probability, this implies that it could be not the best sense but a valid one.

2. using the sense with the highest probability or the set of senses that may be useful, and with the higher probability of being valid than not.

2.1.2 Rada Mihalcea and Andras Csomai - Wikify! Linking Documents to Encyclopedic knowledge

The autors assert that Wikipedia http://en.wikipedia.org can be used to achive state-of-the-art results on both, 1) keyword extraction and 2) word sense disambiguation.

The Wikify! system was designed and implemented with two visions:

1. Semantic Web, it could be used for automatically enrich online documents with references to semantically related information,

2. As a tool that could be used for education porpouses, capable to link important terms to encyclopedic pages as a gateway between for example lecture notes, teaching materials, assigments.

For the experiments Mihalcea and Csomai, worked with a Wikipedia download from March 2006, with approximately 1.4 million articles and more then 37 millions hyperlinks.

Text Wikification

The first task **keyword extraction** is about the identification of those words and relevant phrases that best describe the subject of a document [3].

The second task **link a candidate keyword with the correct Wikipedia article**, for what word sense disambiguation must be perform and context is taken in consideration and the *disambiguation pages* play an important rol in the process.

Some of the decisions taken for this task were supported by the Wikipedia manual of style, 1) like selecting links to provide deeper understanding of the topic or particular terms, such as technical terms, names, places that means only links relevant to the context [4], 2) Avoid terms unrelated to the main topi, 3) Proper amount of keywords in an article, too many links could obstruct the reading.

[3] Keyword extraction, https://en.wikipedia.org/wiki/Keyword_extraction

[4] Wikipedia:Manual of Style/Linking, https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style/Linking

They aware the similarity between this recommendations and the keyword extraction problem, so they address the solution as a keyword extraction task, the recommendations derived in decisions for the system, like constructing controlled vocabulary with 1,918,830 terms including the Wikipedia article titles and extended to include morphological variations taking in consideration all the surface forms collected from all the Wikipedia articles and discounting all the ocurrences that were used less than five times.

The controlled vocabulary included acceptable phrases, the next steps deal with *candidate extraction* in the input document, which is done taking all possible n-grams that are also present in the controlled vocabulary and *ranking keywords* for reflecting the likelihood given to a keyphrase, they used:

- *tf.idf*
- x^2 *independence test*
- *keyphraseness*

Enumeration

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed vitae laoreet lectus. Donec lacus quam, malesuada ut erat vel, consectetur eleifend tellus. Aliquam non feugiat lacus. Interdum et malesuada fames ac ante ipsum primis in faucibus. Quisque a dolor sit amet dui malesuada malesuada id ac metus. Phasellus posuere egestas mauris, sed porta arcu vulputate ut. Donec arcu erat, ultrices et nisl ut, ultricies facilisis urna. Quisque iaculis, lorem non maximus pretium, dui eros auctor quam, sed sodales libero felis vel orci. Aliquam neque nunc, elementum id accumsan eu, varius eu enim. Aliquam blandit ante et ligula tempor pharetra. Donec molestie porttitor commodo. Integer rutrum turpis ac erat tristique cursus. Sed venenatis urna vel tempus venenatis. Nam eu rhoncus eros, et condimentum elit. Quisque risus turpis, aliquam eget euismod id, gravida in odio. Nunc elementum nibh risus, ut faucibus mauris molestie eu. Vivamus quis nunc nec nisl vulputate fringilla. Duis tempus libero ac justo laoreet tincidunt. Fusce sagittis gravida magna, pharetra venenatis mauris semper at. Nullam eleifend felis a elementum sagittis. In vel turpis eu metus euismod tempus eget sit amet tortor. Donec eu rhoncus libero, quis iaculis lectus. Aliquam erat volutpat. Proin id ullamcorper tortor. Fusce vestibulum a enim non volutpat. Nam ut interdum nulla. Proin lacinia felis malesuada arcu aliquet fringilla. Aliquam condimentum, tellus eget maximus porttitor, quam sem luctus massa, eu fermentum arcu diam ac massa. Praesent ut quam id leo molestie rhoncus. Praesent nec odio eget turpis bibendum eleifend non sit amet mi. Curabitur

placerat finibus velit, eu ultricies risus imperdiet ut. Suspendisse lorem orci, luctus porta eros a, commodo maximus nisi.

Nunc et dolor diam. Phasellus eu justo vitae diam vehicula tristique. Vestibulum vulputate cursus turpis nec commodo. Etiam elementum sit amet erat et pellentesque. In eu augue sed tortor mollis tincidunt. Mauris eros dui, sagittis vestibulum vestibulum vitae, molestie a velit. Donec non felis ut velit aliquam convallis sit amet sit amet velit. Aliquam vulputate, elit in lacinia lacinia, odio lacus consectetur quam, sit amet facilisis mi justo id magna. Curabitur aliquet pulvinar eros. Cras metus enim, tristique ut magna a, interdum egestas nibh. Aenean lorem odio, varius a sollicitudin non, cursus a odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae;

1. The first topic is dull
2. The second topic is duller
    (a) The first subtopic is silly
    (b) The second subtopic is stupid
3. The third topic is the dullest

Morbi bibendum est aliquam, hendrerit dolor ac, pretium sem. Nunc molestie, dui in euismod finibus, nunc enim viverra enim, eu mattis mi metus id libero. Cras sed accumsan justo, ut volutpat ipsum. Nam faucibus auctor molestie. Morbi sit amet eros a justo pretium aliquet. Maecenas tempor risus sit amet tincidunt tincidunt. Curabitur dapibus gravida gravida. Vivamus porta ullamcorper nisi eu molestie. Ut pretium nisl eu facilisis tempor. Nulla rutrum tincidunt justo, id placerat lacus laoreet et. Sed cursus lobortis vehicula. Donec sed tortor et est cursus pellentesque sit amet sed velit. Proin efficitur posuere felis, porta auctor nunc. Etiam non porta risus. Pellentesque lacinia eros at ante iaculis, sed aliquet ipsum volutpat. Suspendisse potenti.

Ut ultrices lectus sed sagittis varius. Nulla facilisi. Nullam tortor sem, placerat nec condimentum eu, tristique eget ex. Nullam pretium tellus ut nibh accumsan elementum. Aliquam posuere gravida tellus, id imperdiet nulla rutrum imperdiet. Nulla pretium ullamcorper quam, non iaculis orci consectetur eget. Curabitur non laoreet nisl. Maecenas lacinia, lorem vel tincidunt cursus, odio lorem aliquet est, gravida auctor arcu urna id enim. Morbi accumsan bibendum ipsum, ut maximus dui placerat vitae. Nullam pretium ac tortor nec venenatis. Nunc non aliquet neque.

Description

The first topic is dull

The second topic is duller

- **The first subtopic** is silly
- **The second subtopic** is stupid

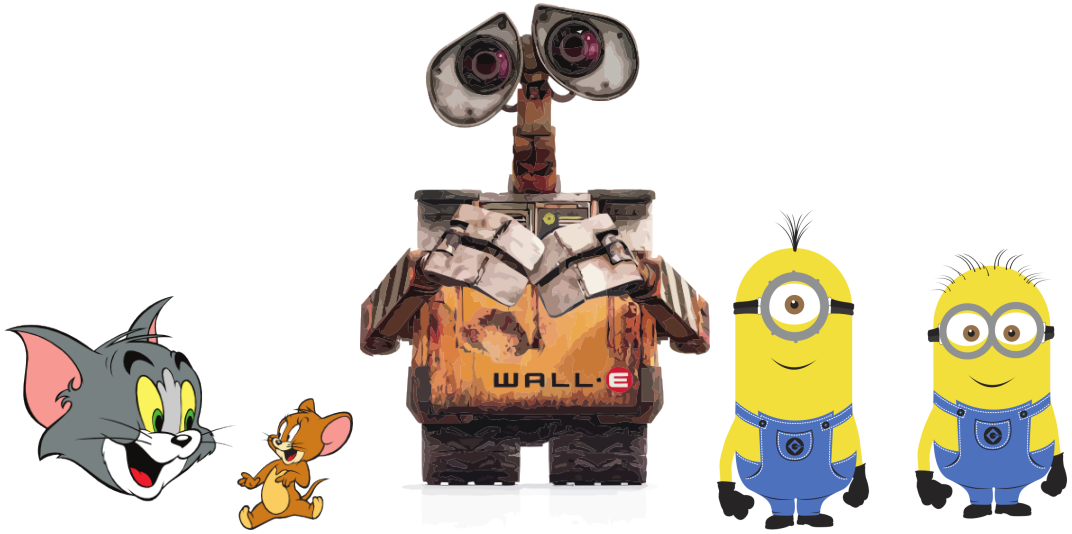
The third topic is the dullest

2.2 Hidden section

Lorem ipsum dolor sit amet, *consectetur adipiscing elit*. In magna nisi, aliquam id blandit id, congue ac est. Fusce porta consequat leo. Proin feugiat at felis vel consectetur. Ut tempus ipsum sit amet congue posuere. Nulla varius rutrum quam. Donec sed purus luctus, faucibus velit id, ultrices sapien. Cras diam purus, tincidunt eget tristique ut, egestas quis nulla. Curabitur vel iaculis lectus. Nunc nulla urna, ultrices et eleifend in, accumsan ut erat. In ut ante leo. Aenean a lacinia nisl, sit amet ullamcorper dolor. Maecenas blandit, tortor ut scelerisque congue, velit diam volutpat metus, sed vestibulum eros justo ut nulla. Etiam nec ipsum non enim luctus porta in in massa. Cras arcu urna, malesuada ut tellus ut, pellentesque mollis risus.Morbi vel tortor imperdiet arcu auctor mattis sit amet eu nisi. Nulla gravida urna vel nisl egestas varius. Aliquam posuere ante quis malesuada dignissim. Mauris ultrices tristique eros, a dignissim nisl iaculis nec. Praesent dapibus tincidunt mauris nec tempor. Curabitur et consequat nisi. Quisque viverra egestas risus, ut sodales enim blandit at. Mauris quis odio nulla. Cras euismod turpis magna, in facilisis diam congue non. Mauris faucibus nisl a orci dictum, et tempus mi cursus.

Etiam elementum tristique lacus, sit amet eleifend nibh eleifend sed [5]. Maecenas dapibu augue ut urna malesuada, non tempor nibh mollis. Donec sed sem sollicitudin, convallis velit aliquam, tincidunt diam. In eu venenatis lorem. Aliquam non augue porttitor tellus faucibus porta et nec ante. Proin sodales, libero vitae commodo sodales, dolor nisi cursus magna, non tincidunt ipsum nibh eget purus. Nam rutrum tincidunt arcu, tincidunt vulputate mi sagittis id. Proin et nisi nec orci tincidunt auctor et porta elit. Praesent eu dolor ac magna cursus euismod. Integer non dictum nunc.

[5] My footnote goes blah blah blah! ...



(a) Tom and Jerry (b) Wall-E (c) Minions

Fig. 2.2 Best Animations

Subplots

I can cite Wall-E (see Fig. 2.2b) and Minions in despicable me (Fig. 2.2c) or I can cite the whole figure as Fig. 2.2

Chapter 3

My third chapter

3.1 First section of the third chapter

And now I begin my third chapter here . . .

And now to cite some more people Read [8], Ancey et al. [2]

3.1.1 First subsection in the first section

. . . and some more

3.1.2 Second subsection in the first section

. . . and some more . . .

First subsub section in the second subsection

. . . and some more in the first subsub section otherwise it all looks the same doesn't it? well we can add some text to it . . .

3.1.3 Third subsection in the first section

. . . and some more . . .

First subsub section in the third subsection

. . . and some more in the first subsub section otherwise it all looks the same doesn't it? well we can add some text to it and some more and some more and some more and some more and some more and some more and some more . . .

Second subsub section in the third subsection

. . . and some more in the first subsub section otherwise it all looks the same doesn't it? well we can add some text to it . . .

3.2 Second section of the third chapter

and here I write more . . .

3.3 The layout of formal tables

This section has been modified from "Publication quality tables in LaTeX*" by Simon Fear.

The layout of a table has been established over centuries of experience and should only be altered in extraordinary circumstances.

When formatting a table, remember two simple guidelines at all times:

1. Never, ever use vertical rules (lines).
2. Never use double rules.

These guidelines may seem extreme but I have never found a good argument in favour of breaking them. For example, if you feel that the information in the left half of a table is so different from that on the right that it needs to be separated by a vertical line, then you should use two tables instead. Not everyone follows the second guideline:

There are three further guidelines worth mentioning here as they are generally not known outside the circle of professional typesetters and subeditors:

3. Put the units in the column heading (not in the body of the table).
4. Always precede a decimal point by a digit; thus 0.1 *not* just .1.
5. Do not use 'ditto' signs or any other such convention to repeat a previous value. In many circumstances a blank will serve just as well. If it won't, then repeat the value.

A frequently seen mistake is to use '\begin{center}' . . . '\end{center}' inside a figure or table environment. This center environment can cause additional vertical space. If you want to avoid that just use '\centering'

Table 3.1 A badly formatted table

	Species I		Species II	
Dental measurement	mean	SD	mean	SD
I1MD	6.23	0.91	5.2	0.7
I1LL	7.48	0.56	8.7	0.71
I2MD	3.99	0.63	4.22	0.54
I2LL	6.81	0.02	6.66	0.01
CMD	13.47	0.09	10.55	0.05
CBL	11.88	0.05	13.11	0.04

Table 3.2 A nice looking table

Dental measurement	Species I		Species II	
	mean	SD	mean	SD
I1MD	6.23	0.91	5.2	0.7
I1LL	7.48	0.56	8.7	0.71
I2MD	3.99	0.63	4.22	0.54
I2LL	6.81	0.02	6.66	0.01
CMD	13.47	0.09	10.55	0.05
CBL	11.88	0.05	13.11	0.04

Table 3.3 Even better looking table using booktabs

Dental measurement	Species I		Species II	
	mean	SD	mean	SD
I1MD	6.23	0.91	5.2	0.7
I1LL	7.48	0.56	8.7	0.71
I2MD	3.99	0.63	4.22	0.54
I2LL	6.81	0.02	6.66	0.01
CMD	13.47	0.09	10.55	0.05
CBL	11.88	0.05	13.11	0.04

References

[1] Abramovich, Y. A., Aliprantis, C. D., and Burkinshaw, O. (1995). Another Characterization of the Invariant Subspace Problem. *Operator Theory in Function Spaces and Banach Lattices.* The A.C. Zaanen Anniversary Volume*, Operator Theory: Advances and Applications*, 75:15–31. Birkhäuser Verlag.

[2] Ancey, C., Coussot, P., and Evesque, P. (1996). Examination of the possibility of a fluid-mechanics treatment of dense granular flows. *Mechanics of Cohesive-frictional Materials*, 1(4):385–403.

[3] Aupetit, B. (1991). *A Primer on Spectral Theory*. Springer-Verlag, New York.

[4] Conway, J. B. (1990). *A Course in Functional Analysis*. Springer-Verlag, New York, second edition.

[5] Ljubič, J. I. and Macaev, V. I. (1965). On Operators with a Separable Spectrum. *Amer. Math. Soc. Transl. (2)*, 47:89–129.

[6] Mihalcea, R. and Csomai, A. (2007). Wikify!: linking documents to encyclopedic knowledge. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, CIKM '07, pages 233–242, New York, NY, USA. ACM.

[7] Milne, D. and Witten, I. H. (2008). An effective, low-cost measure of semantic relatedness obtained from Wikipedia links. In *AAAI 2008*.

[8] Read, C. J. (1985). A solution to the invariant subspace problem on the space l_1. *Bull. London Math. Soc.*, 17:305–317.

[9] Roris, A. et al. (2014). Information Extraction in Semantic, Highly-Structured, and Semi-Structured Web Sources. *Polibits*, 49:69–75.

[10] Zesch, T., Müller, C., and Gurevych, I. (2008). Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary. In *Proceedings of the Conference on Language Resources and Evaluation (LREC)*.

Appendix A

How to install LaTeX

Windows OS

TeXLive package - full version

1. Download the TeXLive ISO (2.2GB) from
   https://www.tug.org/texlive/
2. Download WinCDEmu (if you don't have a virtual drive) from
   http://wincdemu.sysprogs.org/download/
3. To install Windows CD Emulator follow the instructions at
   http://wincdemu.sysprogs.org/tutorials/install/
4. Right click the iso and mount it using the WinCDEmu as shown in
   http://wincdemu.sysprogs.org/tutorials/mount/
5. Open your virtual drive and run setup.pl

or

Basic MikTeX - TeX distribution

1. Download Basic-MiKTeX(32bit or 64bit) from
   http://miktex.org/download
2. Run the installer
3. To add a new package go to Start » All Programs » MikTex » Maintenance (Admin)
   and choose Package Manager

4. Select or search for packages to install

TexStudio - TeX editor

1. Download TexStudio from
   http://texstudio.sourceforge.net/#downloads

2. Run the installer

Mac OS X

MacTeX - TeX distribution

1. Download the file from
   https://www.tug.org/mactex/

2. Extract and double click to run the installer. It does the entire configuration, sit back and relax.

TexStudio - TeX editor

1. Download TexStudio from
   http://texstudio.sourceforge.net/#downloads

2. Extract and Start

Unix/Linux

TeXLive - TeX distribution

Getting the distribution:

1. TexLive can be downloaded from
   http://www.tug.org/texlive/acquire-netinstall.html.

2. TexLive is provided by most operating system you can use (rpm,apt-get or yum) to get TexLive distributions

Installation

1. Mount the ISO file in the mnt directory

   ```
   mount -t iso9660 -o ro,loop,noauto /your/texlive####.iso /mnt
   ```

2. Install wget on your OS (use rpm, apt-get or yum install)

3. Run the installer script install-tl.

   ```
   cd /your/download/directory
   ./install-tl
   ```

4. Enter command ‘i’ for installation

5. Post-Installation configuration:
   http://www.tug.org/texlive/doc/texlive-en/texlive-en.html#x1-320003.4.1

6. Set the path for the directory of TexLive binaries in your .bashrc file

For 32bit OS

For Bourne-compatible shells such as bash, and using Intel x86 GNU/Linux and a default directory setup as an example, the file to edit might be

```
edit $~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/i386-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

For 64bit OS

```
edit $~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/x86_64-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
```

```
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

Fedora/RedHat/CentOS:

```
sudo yum install texlive
sudo yum install psutils
```

SUSE:

```
sudo zypper install texlive
```

Debian/Ubuntu:

```
sudo apt-get install texlive texlive-latex-extra
sudo apt-get install psutils
```

Appendix B

Installing the CUED class file

LaTeX.cls files can be accessed system-wide when they are placed in the <texmf>/tex/latex directory, where <texmf> is the root directory of the user's TeXinstallation. On systems that have a local texmf tree (<texmflocal>), which may be named "texmf-local" or "localtexmf", it may be advisable to install packages in <texmflocal>, rather than <texmf> as the contents of the former, unlike that of the latter, are preserved after the LaTeXsystem is reinstalled and/or upgraded.

It is recommended that the user create a subdirectory <texmf>/tex/latex/CUED for all CUED related LaTeXclass and package files. On some LaTeXsystems, the directory look-up tables will need to be refreshed after making additions or deletions to the system files. For TeXLive systems this is accomplished via executing "texhash" as root. MIKTeXusers can run "initexmf -u" to accomplish the same thing.

Users not willing or able to install the files system-wide can install them in their personal directories, but will then have to provide the path (full or relative) in addition to the filename when referring to them in LaTeX.