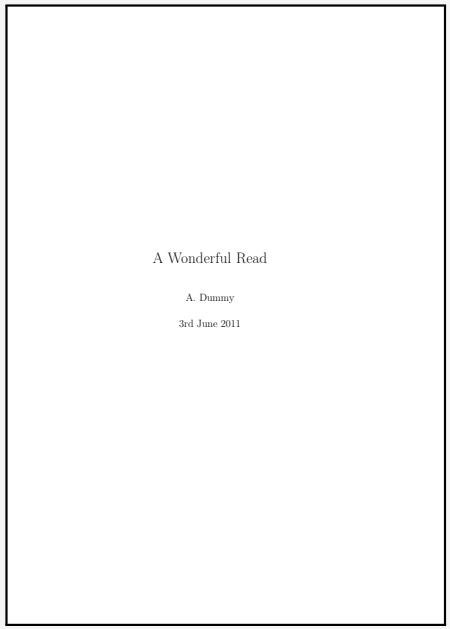
Basic Types

Books

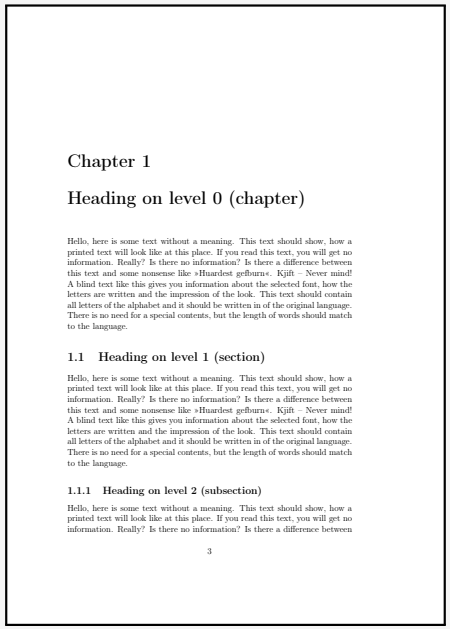
```
\documentclass{book}
\author{...}
\title{...}

\begin{document}
\maketitle
\chapter{...}
\section{...}
...
\subsection{...}
\end{document}
```



A Wonderful Read

A. Dummy

3rd June 2011



Chapter 1

Heading on level 0 (chapter)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

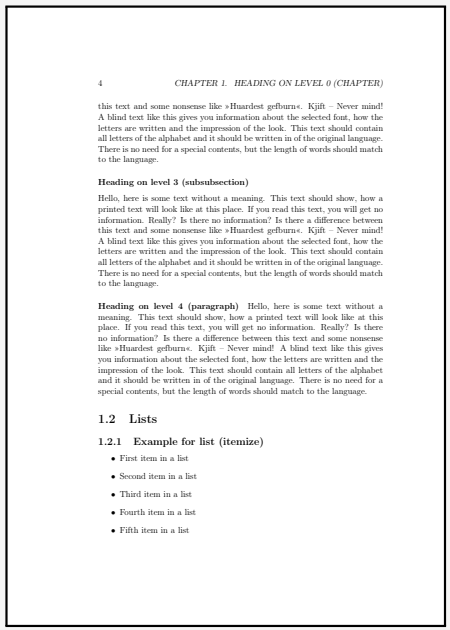
1.1 Heading on level 1 (section)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1.1.1 Heading on level 2 (subsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between

3



4 CHAPTER 1. HEADING ON LEVEL 0 (CHAPTER)

this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

Heading on level 3 (subsubsection)

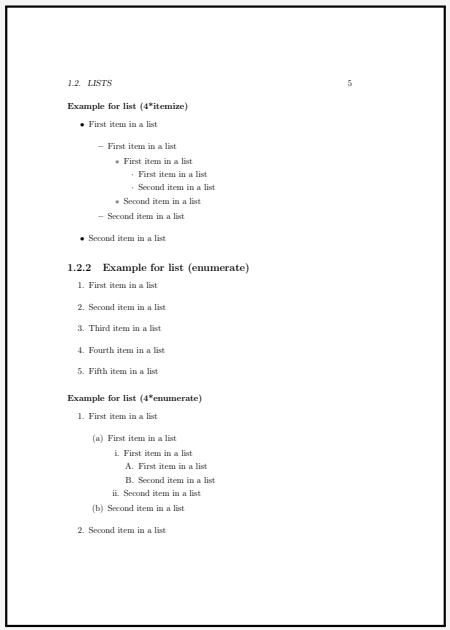
Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

Heading on level 4 (paragraph) Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1.2 Lists

1.2.1 Example for list (itemize)

- First item in a list
- Second item in a list
- Third item in a list
- Fourth item in a list
- Fifth item in a list



1.2. LISTS 5

Example for list (4*itemize)

- First item in a list
  - First item in a list
    - First item in a list
      - First item in a list
      - Second item in a list
    - Second item in a list
  - Second item in a list
- Second item in a list

1.2.2 Example for list (enumerate)

1. First item in a list
2. Second item in a list
3. Third item in a list
4. Fourth item in a list
5. Fifth item in a list

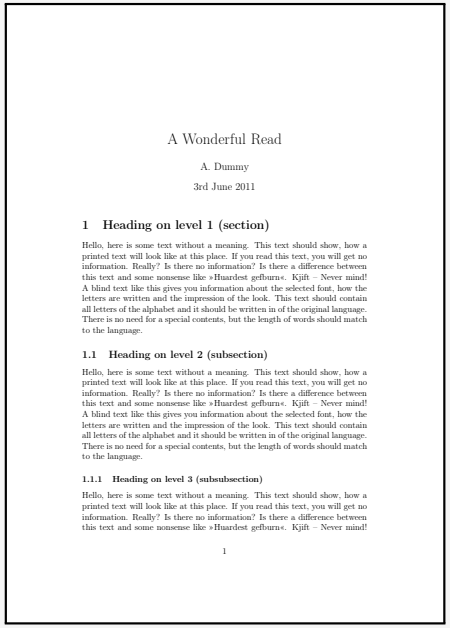
Example for list (4*enumerate)

1. First item in a list
   (a) First item in a list
      i. First item in a list
         A. First item in a list
         B. Second item in a list
      ii. Second item in a list
   (b) Second item in a list
2. Second item in a list

Basic Types (cont'd)

Articles

```
\documentclass{article}
\author{...}
\title{...}

\begin{document}
\maketitle
\section{...}
...
\subsection{...}
\end{document}
```


A Wonderful Read

A. Dummy

3rd June 2011

1 Heading on level 1 (section)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

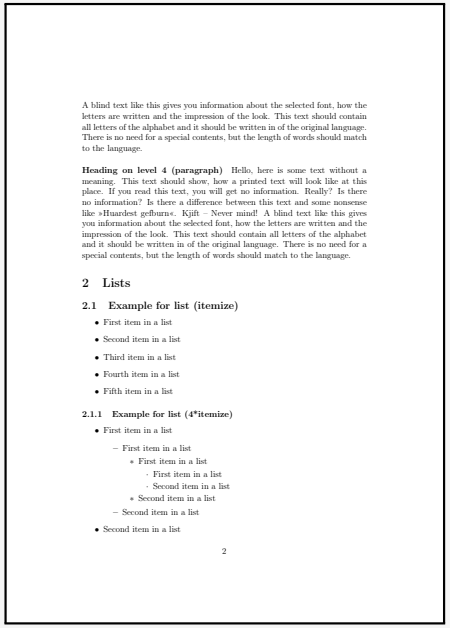
1.1 Heading on level 2 (subsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1.1.1 Heading on level 3 (subsubsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind!

1


A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

Heading on level 4 (paragraph) Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

2 Lists

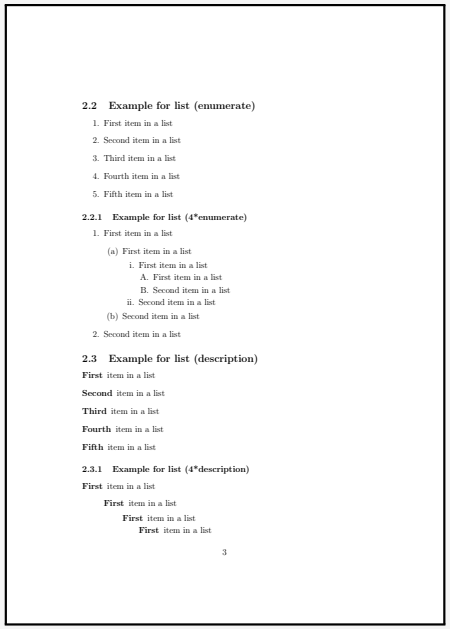
2.1 Example for list (itemize)

- First item in a list
- Second item in a list
- Third item in a list
- Fourth item in a list
- Fifth item in a list

2.1.1 Example for list (4*itemize)

- First item in a list
  - First item in a list
    - First item in a list
      - First item in a list
      - Second item in a list
    - Second item in a list
  - Second item in a list
- Second item in a list

2


2.2 Example for list (enumerate)

1. First item in a list
2. Second item in a list
3. Third item in a list
4. Fourth item in a list
5. Fifth item in a list

2.2.1 Example for list (4*enumerate)

1. First item in a list
   (a) First item in a list
      i. First item in a list
         A. First item in a list
         B. Second item in a list
      ii. Second item in a list
   (b) Second item in a list
2. Second item in a list

2.3 Example for list (description)

First item in a list

Second item in a list

Third item in a list

Fourth item in a list

Fifth item in a list

2.3.1 Example for list (4*description)

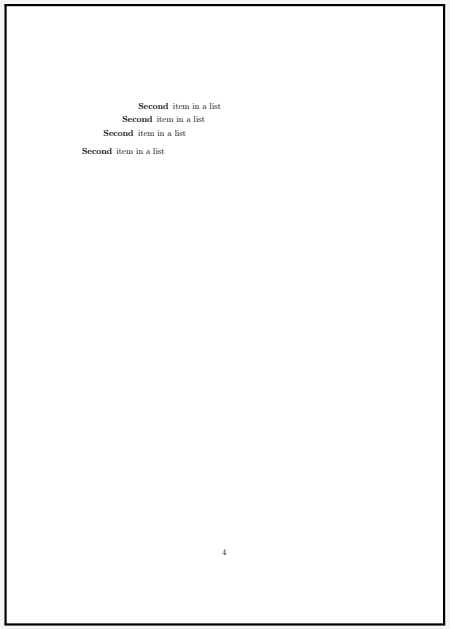
First item in a list

First item in a list

First item in a list

First item in a list

3


Second item in a list

Second item in a list

Second item in a list

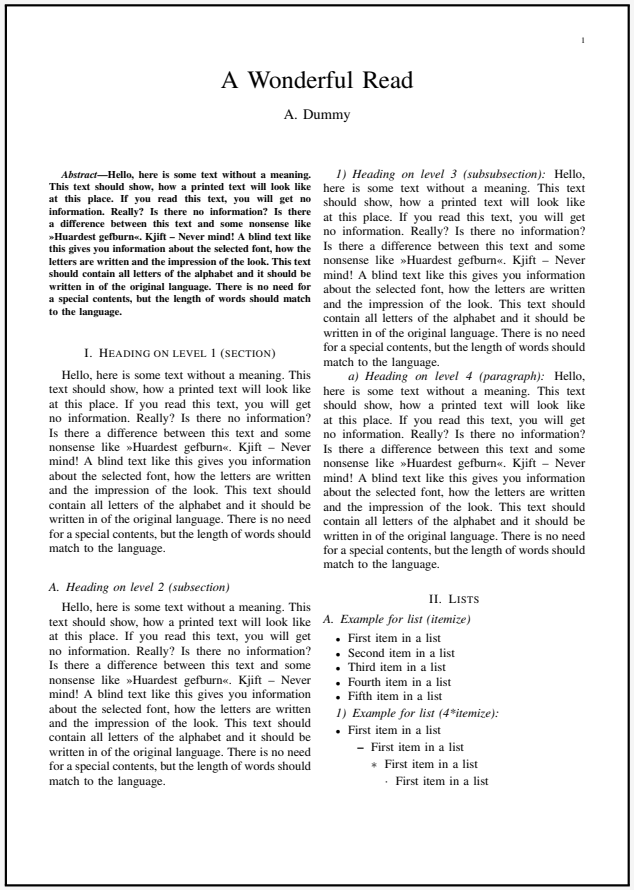
Second item in a list

4

Journal and Conference Proceedings Articles

IEEE

`\documentclass{IEEEtran}`



A Wonderful Read

A. Dummy

***Abstract*—Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.**

I. Heading on level 1 (section)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

A. Heading on level 2 (subsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1) Heading on level 3 (subsubsection): Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

a) Heading on level 4 (paragraph): Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

II. Lists

A. Example for list (itemize)

- First item in a list
- Second item in a list
- Third item in a list
- Fourth item in a list
- Fifth item in a list

*1) Example for list (4*itemize):*

- First item in a list
  - First item in a list
    - First item in a list
      - First item in a list

ACM

`\documentclass{sig-alternate}`

A Wonderful Read

A. Dummy

ABSTRACT

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1. Heading on level 1 (SECTION)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1.1 Heading on level 2 (subsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1.1.1 Heading on level 3 (subsubsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

Heading on level 4 (paragraph).
Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

2. Lists

2.1 Example for list (itemize)

- First item in a list
- Second item in a list
- Third item in a list
- Fourth item in a list
- Fifth item in a list

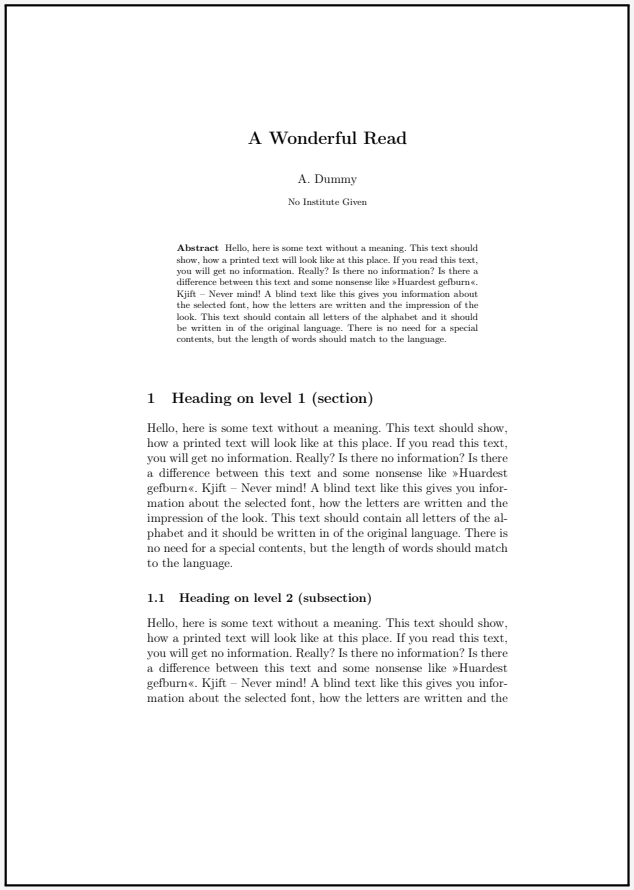
*2.1.1 Example for list (4*itemize)*

- First item in a list
  - First item in a list
    - First item in a list
      - First item in a list
      - Second item in a list
    - Second item in a list
  - Second item in a list
- Second item in a list

MOSC 2011, July 3–5, 2011, Penang, Malaysia.

LLNCS

`\documentclass{llncs}`



A Wonderful Read

A. Dummy

No Institute Given

Abstract Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1 Heading on level 1 (section)

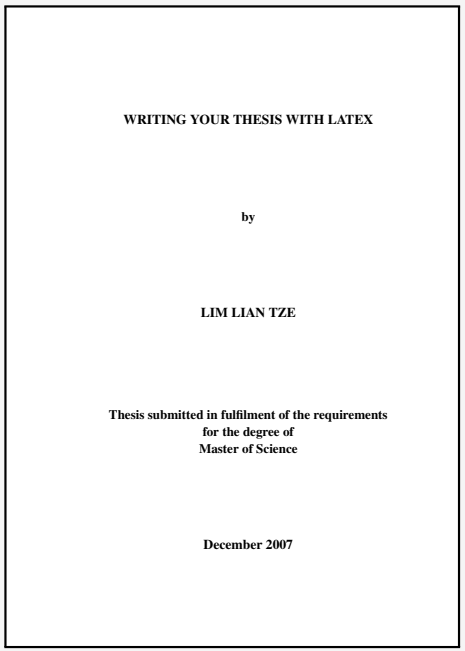
Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.

1.1 Heading on level 2 (subsection)

Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like »Huardest gefburn«. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the

University Theses

Universiti Sains Malaysia `\documentclass{usmthesis}`



WRITING YOUR THESIS WITH LATEX

by

LIM LIAN TZE

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

December 2007

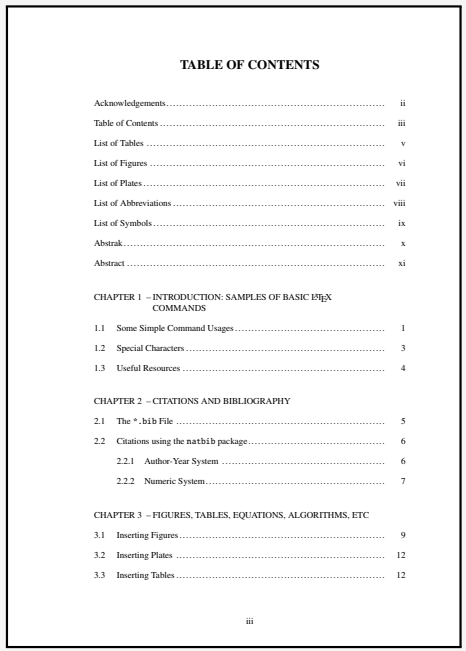
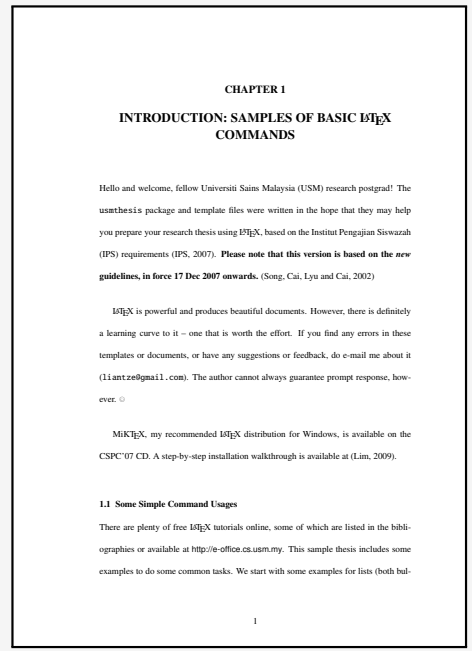


TABLE OF CONTENTS

iii



CHAPTER 1

INTRODUCTION: SAMPLES OF BASIC LaTeX COMMANDS

Hello and welcome, fellow Universiti Sains Malaysia (USM) research postgrad! The `usmthesis` package and template files were written in the hope that they may help you prepare your research thesis using LaTeX, based on the Institut Pengajian Siswazah (IPS) requirements (IPS, 2007). **Please note that this version is based on the *new* guidelines, in force 17 Dec 2007 onwards.** (Song, Cai, Lyu and Cai, 2002)

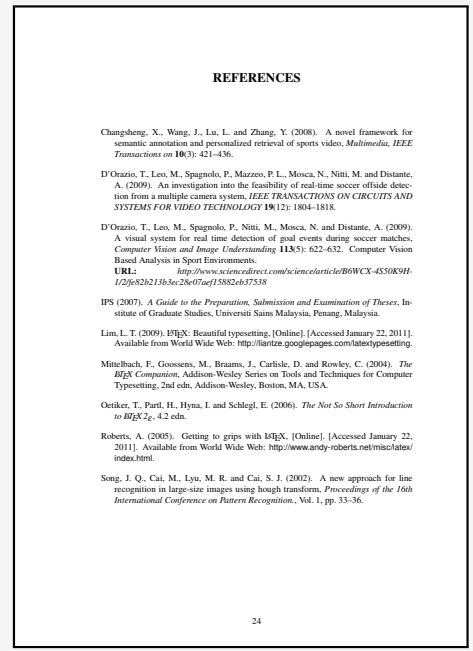
LaTeX is powerful and produces beautiful documents. However, there is definitely a learning curve to it – one that is worth the effort. If you find any errors in these templates or documents, or have any suggestions or feedback, do e-mail me about it (`liantze@gmail.com`). The author cannot always guarantee prompt response, however. ☺

MiKTeX, my recommended LaTeX distribution for Windows, is available on the CSPC'07 CD. A step-by-step installation walkthrough is available at (Lim, 2009).

1.1 Some Simple Command Usages

There are plenty of free LaTeX tutorials online, some of which are listed in the bibliographies or available at http://e-office.cs.usm.my. This sample thesis includes some examples to do some common tasks. We start with some examples for lists (both bul-

1



REFERENCES

Changsheng, X., Wang, J., Lu, L. and Zhang, Y. (2008). A novel framework for semantic annotation and personalized retrieval of sports video, *Multimedia, IEEE Transactions on* **10**(3): 421–436.

D'Orazio, T., Leo, M., Spagnolo, P., Mazzeo, P. L., Mosca, N., Nitti, M. and Distante, A. (2009). An investigation into the feasibility of real-time soccer offside detection from a multiple camera system, *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY* **19**(12): 1804–1818.

D'Orazio, T., Leo, M., Spagnolo, P., Nitti, M., Mosca, N. and Distante, A. (2009). A visual system for real time detection of goal events during soccer matches, *Computer Vision and Image Understanding* **113**(5): 622–632. Computer Vision Based Analysis in Sport Environments.
URL: *http://www.sciencedirect.com/science/article/B6WCX-4S50K9H-1/2/fe82b213b3ec28e07aef15882eb37538*

IPS (2007). *A Guide to the Preparation, Submission and Examination of Theses*, Institute of Graduate Studies, Universiti Sains Malaysia, Penang, Malaysia.

Lim, L. T. (2009). LaTeX: Beautiful typesetting, [Online]. [Accessed January 22, 2011]. Available from World Wide Web: http://liantze.googlepages.com/latextypesetting.

Mittelbach, F., Goossens, M., Braams, J., Carlisle, D. and Rowley, C. (2004). *The LaTeX Companion*, Addison-Wesley Series on Tools and Techniques for Computer Typesetting, 2nd edn, Addison-Wesley, Boston, MA, USA.

Oetiker, T., Partl, H., Hyna, I. and Schlegl, E. (2006). *The Not So Short Introduction to LaTeX 2ε*, 4.2 edn.

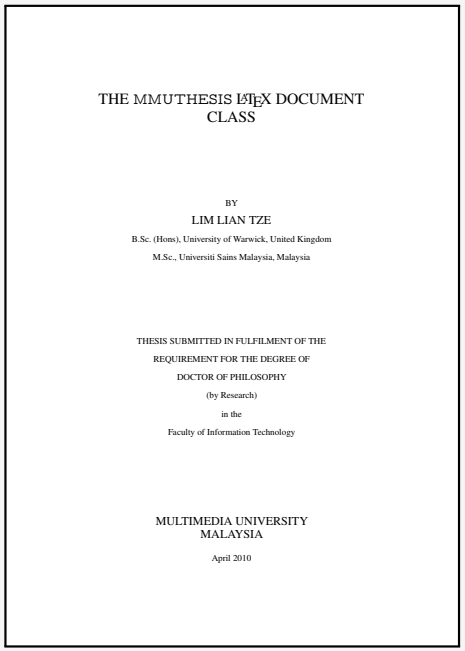
Roberts, A. (2005). Getting to grips with LaTeX, [Online]. [Accessed January 22, 2011]. Available from World Wide Web: http://www.andy-roberts.net/misc/latex/index.html.

Song, J. Q., Cai, M., Lyu, M. R. and Cai, S. J. (2002). A new approach for line recognition in large-size images using hough transform, *Proceedings of the 16th International Conference on Pattern Recognition.*, Vol. 1, pp. 33–36.

24

University Theses (cont'd)

Multimedia University `\documentclass{mmuthesis}`



THE MMUTHESIS LaTeX DOCUMENT CLASS

BY

LIM LIAN TZE

B.Sc. (Hons), University of Warwick, United Kingdom

M.Sc., Universiti Sains Malaysia, Malaysia

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

(by Research)

in the

Faculty of Information Technology

MULTIMEDIA UNIVERSITY MALAYSIA

April 2010

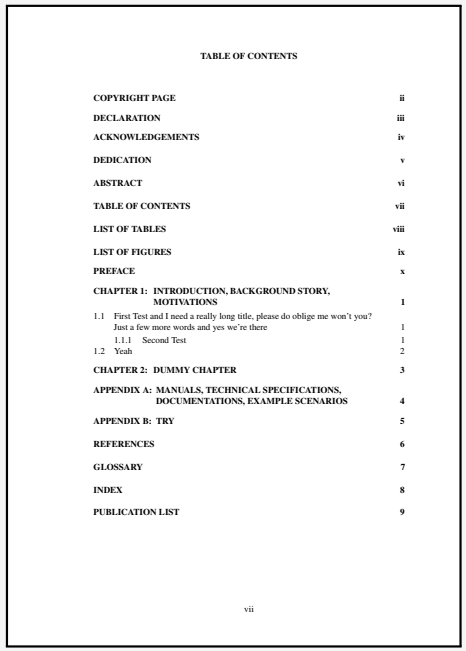
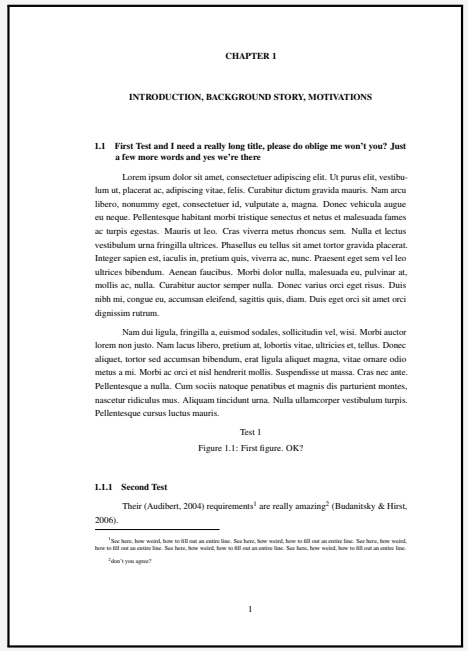


TABLE OF CONTENTS

vii



CHAPTER 1

INTRODUCTION, BACKGROUND STORY, MOTIVATIONS

1.1 First Test and I need a really long title, please do oblige me won't you? Just a few more words and yes we're there

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Test 1

Figure 1.1: First figure. OK?

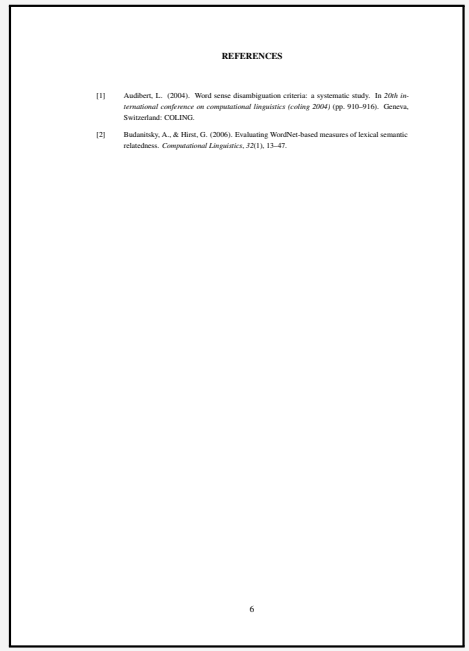
1.1.1 Second Test

Their (Audibert, 2004) requirements[1] are really amazing[2] (Budanitsky & Hirst, 2006).

[1] See here, how weird, how to fill out an entire line. See here, how weird, how to fill out an entire line. See here, how weird, how to fill out an entire line. See here, how weird, how to fill out an entire line. See here, how weird, how to fill out an entire line. See here, how weird, how to fill out an entire line.

[2] don't you agree?

1



REFERENCES

[1] Audibert, L. (2004). Word sense disambiguation criteria: a systematic study. In *20th international conference on computational linguistics (coling 2004)* (pp. 910–916). Geneva, Switzerland: COLING.

[2] Budanitsky, A., & Hirst, G. (2006). Evaluating WordNet-based measures of lexical semantic relatedness. *Computational Linguistics*, *32*(1), 13–47.

6

University Theses (cont'd)

Universiti Malaya `\documentclass{umalayathesis}`

THE UMALAYATHESIS LaTeX DOCUMENT CLASS

LIM LIAN TZE

THESIS SUBMITTED IN FULFILMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF POSTGRADUATE STUDIES
UNIVERSITY OF MALAYA
KUALA LUMPUR

2010

TABLE OF CONTENTS

v

CHAPTER 1

INTRODUCTION, BACKGROUND STORY, MOTIVATIONS

1.1 First Test and I need a really long title, please do oblige me won't you? Just a few more words and yes we're there

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Figure 1.1: First figure. OK?

1

REFERENCES

Audibert, L. (2004). Word sense disambiguation criteria: a systematic study. In *20th International Conference on Computational Linguistics (COLING 2004)* (pp. 910-916). Geneva, Switzerland: COLING.

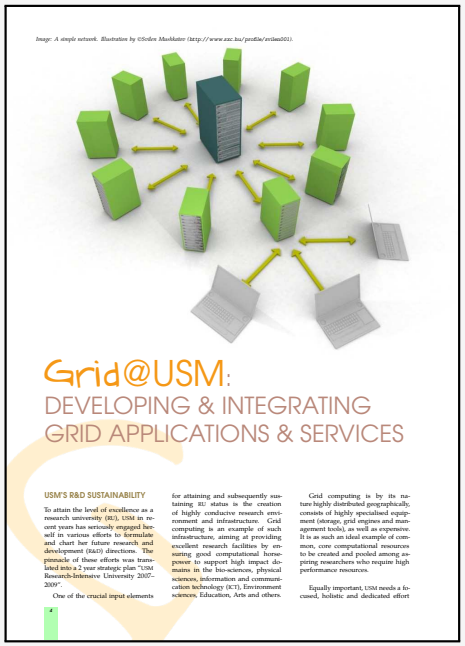
Budanitsky, A., & Hirst, G. (2006). Evaluating WordNet-based measures of lexical semantic relatedness. *Computational Linguistics*, *32*(1), 13–47.

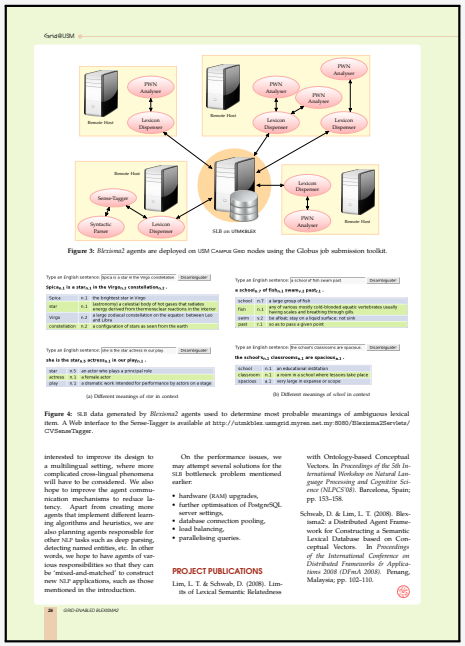
9

Highly Configurable Documents

memoir and KOMA-Script Classes

- Sectional headings
- Running headers and footers
- Good font, colour and illustration choices
- http://latex-my.blogspot.com/search/label/bookdesign

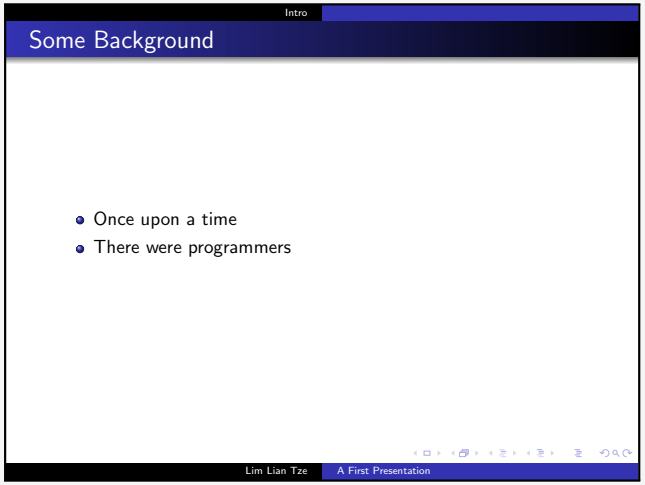




Presentation Slides

- This presentation was made with LaTeX!
- Many possible classes: powerdot, beamer

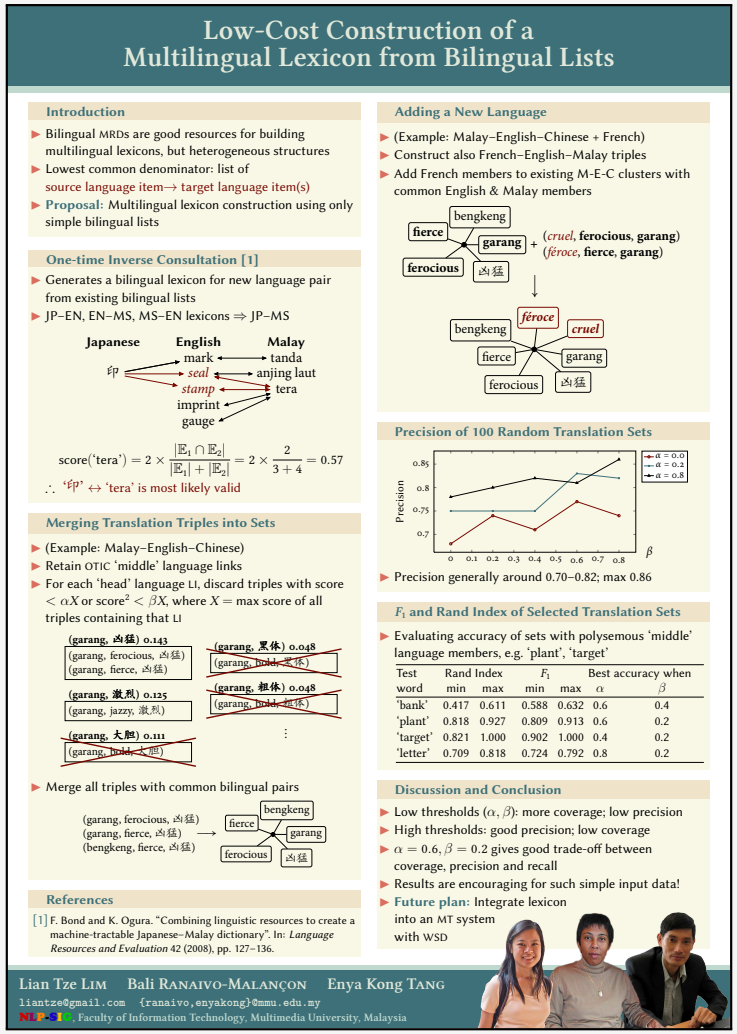
```
\documentclass{beamer}
\usetheme{Warsaw}

\author ...

\begin{document}
\titleframe

\section{Intro}

\begin{frame}
\frametitle{Some Background}
...
\end{frame}
\end{document}
```



Oversized Posters

- Many possible solutions: `sciposter`, `flowfram`, `beamerposter`

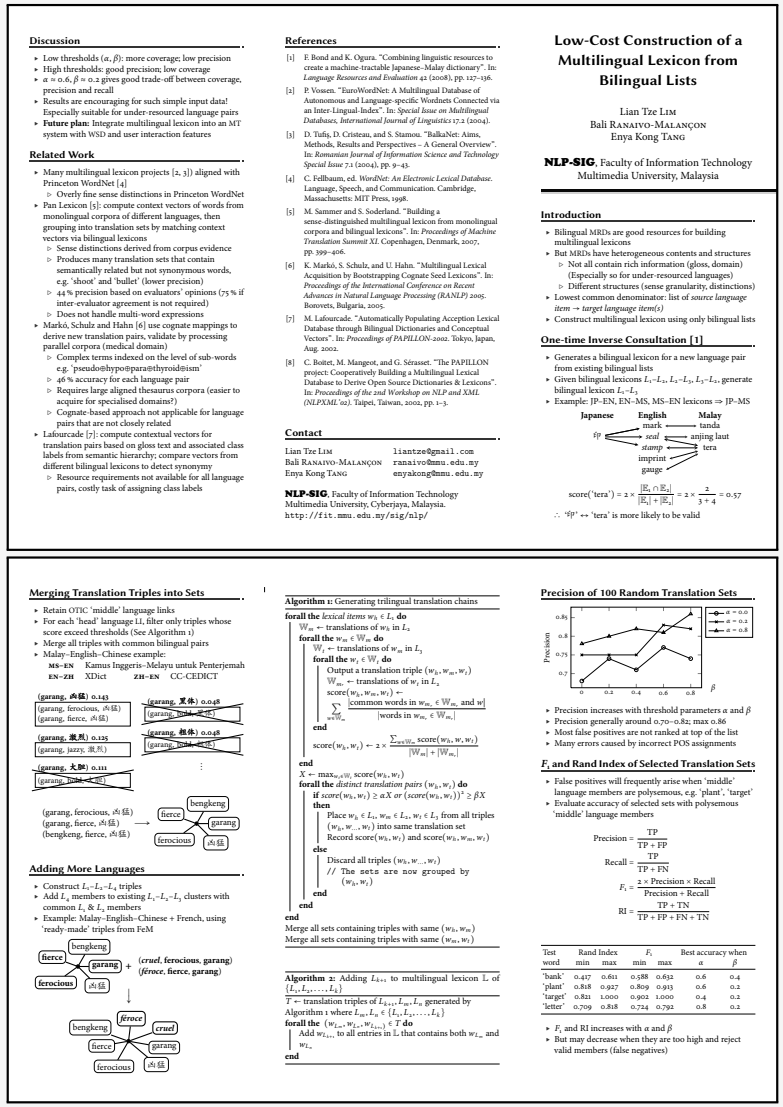
```
\documentclass{beamer}
\usepackage[orientation=portrait,
↪ size=a0]{beamerposter}
\usetheme{...}
\author ... % Meta—information

\begin{document}
\begin{frame}
... % Poster contents goes here
\end{frame}
\end{document}
```



Leaflets

- `leaflet`: arrange contents into 6 pages on a foldable double-sided sheet

```
\documentclass[foldmark,a4paper]
{leaflet}
\author ...  % Meta—information

\begin{document}
\maketitle
\section ...
... % Leaflet contents
\end{document}
```



Discussion

- Low thresholds (α, β): more coverage; low precision
- High thresholds: good precision; low coverage
- $\alpha = 0.6$, $\beta = 0.2$ gives good trade-off between coverage, precision and recall
- Results are encouraging for such simple input data! Especially suitable for under-resourced language pairs
- **Future plan**: Integrate multilingual lexicon into an MT system with WSD and user interaction features

Related Work

- Many multilingual lexicon projects [2, 3]) aligned with Princeton WordNet [4]
  - Overly fine sense distinctions in Princeton WordNet
- Pan Lexicon [5]: compute context vectors of words from monolingual corpora of different languages, then grouping into translation sets by matching context vectors via bilingual lexicons
  - Sense distinctions derived from corpus evidence
  - Produces many translation sets that contain semantically related but not synonymous words, e.g. 'shoot' and 'bullet' (lower precision)
  - 44 % precision based on evaluators' opinions (75 % if inter-evaluator agreement is not required)
  - Does not handle multi-word expressions
- Markó, Schulz and Hahn [6] use cognate mappings to derive new translation pairs, validate by processing parallel corpora (medical domain)
  - Complex terms indexed on the level of sub-words e.g. 'pseudo@hypo@para@thyroid@ism'
  - 46 % accuracy for each language pair
  - Requires large aligned thesaurus corpora (easier to acquire for specialised domains?)
  - Cognate-based approach not applicable for language pairs that are not closely related
- Lafourcade [7]: compute contextual vectors for translation pairs based on gloss text and associated class labels from semantic hierarchy; compare vectors from different bilingual lexicons to detect synonymy
  - Resource requirements not available for all language pairs, costly task of assigning class labels

References

[1] F. Bond and K. Ogura. "Combining linguistic resources to create a machine-tractable Japanese–Malay dictionary". In: *Language Resources and Evaluation* 42 (2008), pp. 127–136.

[2] P. Vossen. "EuroWordNet: A Multilingual Database of Autonomous and Language-specific Wordnets Connected via an Inter-Lingual-Index". In: *Special Issue on Multilingual Databases, International Journal of Linguistics* 17.2 (2004).

[3] D. Tufiş, D. Cristeau, and S. Stamou. "BalkaNet: Aims, Methods, Results and Perspectives – A General Overview". In: *Romanian Journal of Information Science and Technology Special Issue* 7.1 (2004), pp. 9–43.

[4] C. Fellbaum, ed. *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. Cambridge, Massachusetts: MIT Press, 1998.

[5] M. Sammer and S. Soderland. "Building a sense-distinguished multilingual lexicon from monolingual corpora and bilingual lexicons". In: *Proceedings of Machine Translation Summit XI*. Copenhagen, Denmark, 2007, pp. 399–406.

[6] K. Markó, S. Schulz, and U. Hahn. "Multilingual Lexical Acquisition by Bootstrapping Cognate Seed Lexicons". In: *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP) 2005*. Borovets, Bulgaria, 2005.

[7] M. Lafourcade. "Automatically Populating Acception Lexical Database through Bilingual Dictionaries and Conceptual Vectors". In: *Proceedings of PAPILLON-2002*. Tokyo, Japan, Aug. 2002.

[8] C. Bottet, M. Mangeot, and G. Sérasset. "The PAPILLON project: Cooperatively Building a Multilingual Lexical Database to Derive Open Source Dictionaries & Lexicons". In: *Proceedings of the 2nd Workshop on NLP and XML (NLPXML '02)*. Taipei, Taiwan, 2002, pp. 1–3.

Contact

Lian Tze Lim — liantze@gmail.com
Bali Ranaivo-Malançon — ranaivo@mmu.edu.my
Enya Kong Tang — enyakong@mmu.edu.my

NLP-SIG, Faculty of Information Technology
Multimedia University, Cyberjaya, Malaysia.
http://fit.mmu.edu.my/nlp/

Low-Cost Construction of a Multilingual Lexicon from Bilingual Lists

Lian Tze Lim
Bali Ranaivo-Malançon
Enya Kong Tang

NLP-SIG, Faculty of Information Technology
Multimedia University, Malaysia

Introduction

- Bilingual MRDs are good resources for building multilingual lexicons
- But MRDs have heterogeneous contents and structures
  - Not all contain rich information (gloss, domain) (Especially so for under-resourced languages)
  - Different structures (sense granularity, distinctions)
- Lowest common denominator: list of *source language item* → *target language item(s)*
- Construct multilingual lexicon using only bilingual lists

One-time Inverse Consultation [1]

- Generates a bilingual lexicon for a new language pair from existing bilingual lists
- Given bilingual lexicons L_1–L_2, L_2–L_3, L_3–L_4, generate bilingual lexicon L_1–L_4
- Example: JP–EN, EN–MS, MS–EN lexicons ⇒ JP–MS

$$\text{score}(\text{`tera'}) = 2 \times \frac{|\mathbb{E}_s \cap \mathbb{E}_p|}{|\mathbb{E}_s| + |\mathbb{E}_p|} = 2 \times \frac{2}{3+4} = 0.57$$

$\therefore$ 'とら' ↔ 'tera' is more likely to be valid

Merging Translation Triples into Sets

- Retain OTIC 'middle' language links
- For each 'head' language $L1$, filter only triples whose score exceed thresholds (See Algorithm 1)
- Merge all triples with common bilingual pairs
- Malay–English–Chinese example:
  - **MS–EN** Kamus Inggeris–Melayu untuk Penterjemah
  - **EN–ZH** XDict
  - **ZH–EN** CC-CEDICT

Adding More Languages

- Construct L_3–L_4–L_4 triples
- Add L_4 members to existing L_1–L_2–L_3 clusters with common L_3 & L_4 members
- Example: Malay–English–Chinese + French, using 'ready-made' triples from FeM

Algorithm 1: Generating trilingual translation chains

forall the *lexical items* $w_3 \in L_3$ **do**
  $\mathbb{W}_4 \leftarrow$ translations of w_3 in L_4
  forall the $w_4 \in \mathbb{W}_4$ **do**
    $\mathbb{W}_1 \leftarrow$ translations of w_4 in L_1
    forall the $w_1 \in \mathbb{W}_1$ **do**
      Output a translation triple (w_3, w_4, w_1)
      $\mathbb{W}_{w_1} \leftarrow$ translations of w_1 in L_3
      $\text{score}(w_3, w_4, w_1) \leftarrow \sum_{w \in \mathbb{W}_{w_1}} \frac{|\text{common words in } w_4 \in \mathbb{W}_{w_1} \text{ and } w|}{|\text{words in } w_4 \in \mathbb{W}_{w_1}|}$
    end
    $\text{score}(w_3, w_1) \leftarrow 2 \times \frac{\sum_{w \in \mathbb{W}_{w_1}} \text{score}(w_3, w, w_1)}{|\mathbb{W}_{w_1}| + |\mathbb{W}_{w_1}|}$
  end
  $X \leftarrow \max_{w_1 \in \mathbb{W}} \text{score}(w_3, w_1)$
  forall the *distinct translation pairs* (w_3, w_1) **do**
    if $\text{score}(w_3, w_1) \geq \alpha X$ *or* $(\text{score}(w_3, w_1))^2 \geq \beta X$ **then**
      Place $w_3 \in L_3$, $w_4 \in L_4$, $w_1 \in L_1$ from all triples (w_3, w_-, w_1) into same translation set
      Record $\text{score}(w_3, w_1)$ and $\text{score}(w_3, w_4, w_1)$
    else
      Discard all triples (w_3, w_-, w_1)
      // The sets are now grouped by (w_3, w_1)
    end
  end
end
Merge all sets containing triples with same (w_3, w_4)
Merge all sets containing triples with same (w_4, w_1)

Algorithm 2: Adding L_{4+1} to multilingual lexicon $\mathbb{L}$ of $\{L_1, L_2, \ldots, L_4\}$

$\mathbb{T} \leftarrow$ translation triples of L_{4+1}–L_m–L_n generated by Algorithm 1 where $L_m, L_n \in \{L_1, L_2, \ldots, L_4\}$
forall the $(w_{1-}, w_{2-}, w_{3+1-}) \in \mathbb{T}$ **do**
  Add w_{3+1} to all entries in $\mathbb{L}$ that contains both w_{1-} and w_{2-}

Precision of 100 Random Translation Sets

- Precision increases with threshold parameters α and β
- Precision generally around 0.70–0.82; max 0.86
- Most false positives are not ranked at top of the list
- Many errors caused by incorrect POS assignments

F_1 and Rand Index of Selected Translation Sets

- False positives will frequently arise when 'middle' language members are polysemous, e.g. 'plant', 'target'
- Evaluate accuracy of selected sets with polysemous 'middle' language members

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{RI} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

Test word	Rand Index min	Rand Index max	F_1 min	F_1 max	Best accuracy when α	Best accuracy when β
'bank'	0.417	0.611	0.588	0.632	0.6	0.4
'plant'	0.818	0.927	0.809	0.913	0.6	0.2
'target'	0.821	1.000	0.902	1.000	0.4	0.2
'letter'	0.709	0.818	0.724	0.792	0.8	0.2

- F_1 and RI increases with α and β
- But may decrease when they are too high and reject valid members (false negatives)

Flash Cards

```
\documentclass[avery5388,frame]
{flashcards}
\cardfrontstyle{headings}
\cardfrontfoot{Linux}

\begin{document}
\begin{flashcard}[Security]
{Certificate}
...
\end{flashcard}

\begin{flashcard}[Security]
{MAC ...}
...
\end{flashcard}
\end{document}
```

Examination Questions

```
\documentclass{exam}
...
\begin{questions}\printanswers
\question[5]
What is Paul McCartney's middle name?
\begin{oneparchoices}
\choice John \CorrectChoice Paul
\choice Ringo \choice James
\end{oneparchoices}

\question[10] What was the Beatles' first
↪single in 1962?
\begin{solution}Love Me Do\end{solution}

\question
\begin{parts}
\part[5] What was George's inspiration for
↪`While My Guitar Gently Weeps'?
\begin{solution}
He opened a random book and saw the words
↪``gently weep''.
\end{solution}
...
\end{questions}
```

1. What is Paul McCartney's middle name? (5)

   A. John **B. Paul** C. Ringo D. James

2. What was the Beatles' first single in 1962? (10)

   Solution: Love Me Do

3. (a) What was George's inspiration for 'While My Guitar Gently Weeps'? (5)

   Solution: He opened a random book and saw the words "gently weep".

   (b) Who guest-performed for the song and why? (5)

   Solution: Eric Clapton; he wanted a spiffy guitar solo.

Mathematics

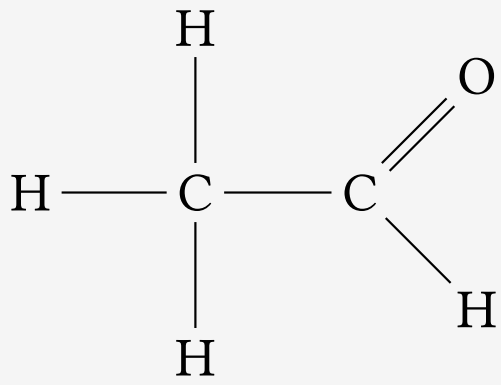
(1) relates the golden ratio and the Fibonacci series. Recall that the golden ratio, $\phi = \frac{1}{2}(1 + \sqrt{5})$.

$$\phi = 1 + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{F_n F_{n+1}} \tag{1}$$

```
\eqref{eq:gratio} relates the golden ratio and the Fibonacci series.
Recall that the golden ratio, $\phi = \frac{1}{2} (1 + \sqrt{5})$.

\begin{equation}\label{eq:gratio}
\phi = 1 + \sum^{\infty} _{n=1}
                \frac{ (-1)^{n+1} }{ F_n F_{n+1} }
\end{equation}
```

Chemical Equations and Molecules

$$\mathrm{Zn^{2+}} \underset{+2\,\mathrm{H^+}}{\overset{+2\,\mathrm{OH^-}}{\rightleftharpoons}} \underset{\text{amphoteres Hydroxid}}{\mathrm{Zn(OH)_2}\downarrow} \underset{+2\,\mathrm{H^+}}{\overset{+2\,\mathrm{OH^-}}{\rightleftharpoons}} \underset{\text{Hydroxozikat}}{\mathrm{[Zn(OH)_4]^{2-}}}$$

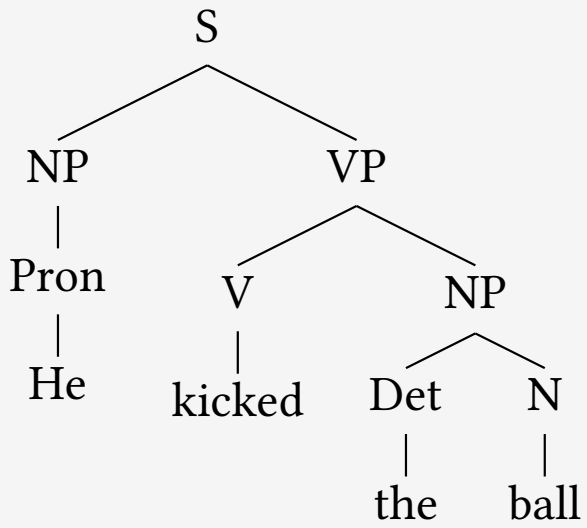



```
\usepackage[version=3]{mhchem}      % sufficient for chemical equations
\usepackage{chemfig}      % for 2-D molecule drawings
...
\ce{Zn^2+ <=>[\ce{+ 2OH-}][\ce{+ 2H+}]
$\underset{\text{amphoteres Hydroxid}}{\ce{Zn(OH)2 v}}$
<=> C[+2OH-][{+ 2H+}]
$\underset{\text{Hydroxozikat}}{\cf{[Zn(OH)4]^2-}}$ }

\chemfig{H-C(-[2]H)(-[6]H)-C(-[7]H)=[1]O}
```

Linguistics

(1) %*Wen liebt seine Mutter?
Whom loves his mother
'Who does his mother love?'

(2) [[$_{NP}$ He] [$_{VP}$ kicked [$_{NP}$ the ball]]]$_{S}$



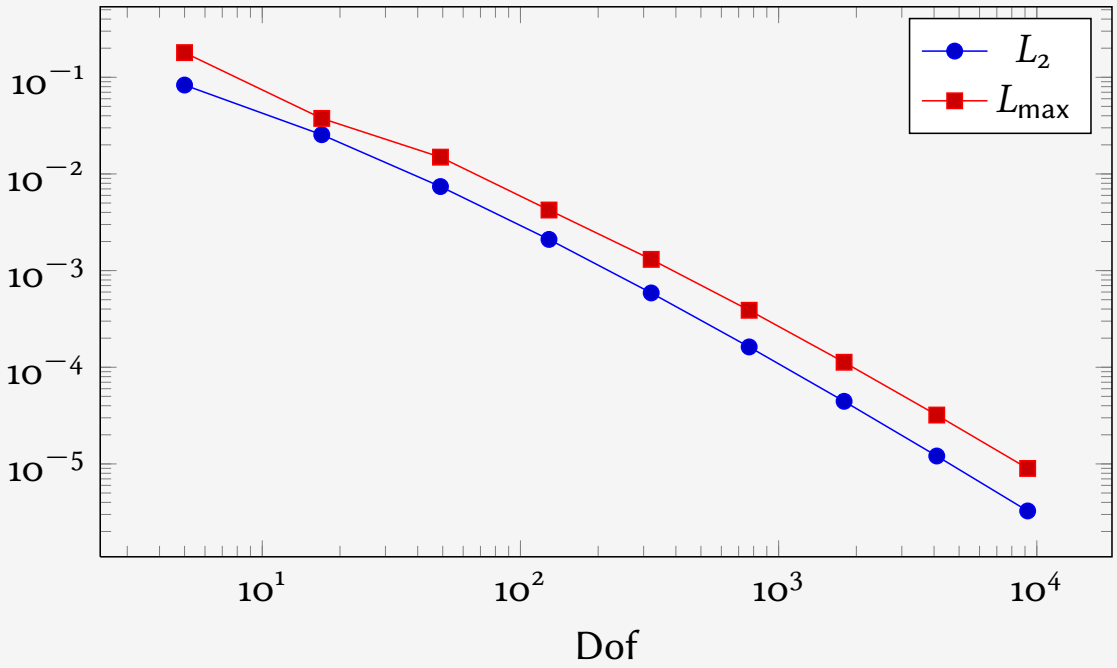

```
\usepackage{linguex,qtree}
...
\exg. \%*Wen liebt seine Mutter?\\
Whom loves his mother\\
`Who does his mother love?'

\exi. [[NP He ] [VP kicked [NP the ball ]]]S

\Tree [ .S [.NP [.Pron He ] ] [.VP [.V kicked ] [.NP [.Det the ] [.N ball ]
↪ ] ] ]
```

Program Listings

```
\usepackage{listings,xcolor}
...
\begin{lstlisting}
[language=C,columns=fullflexible,
basicstyle=\ttfamily,
keywordstyle=\bfseries\color{red},
commentstyle=\sffamily\color{green},
stringstyle=\rmfamily\color{orange}]
#include <stdio.h>
/*
 | Prints "hello world"
 */
int main(void)
{
    printf("hello, world\n");
    return 0;
}
\end{lstlisting}
```

```c
#include <stdio.h>

/*
 | Prints "hello world"
 */
int main(void)
{
    printf("hello, world\n");
    return 0;
}
```

Graph Plots




```
\usepackage{pgfplots}
...
\begin{tikzpicture}
\begin{loglogaxis}[xlabel=Dof]
\addplot table[x=dof,y=L2]{datafile.dat}; \addlegendentry{$L_2$};
\addplot table[x=dof,y=Lmax]{datafile.dat};  \addlegendentry{$L_\text{max}$};
\end{loglogaxis}
\end{tikzpicture}
```

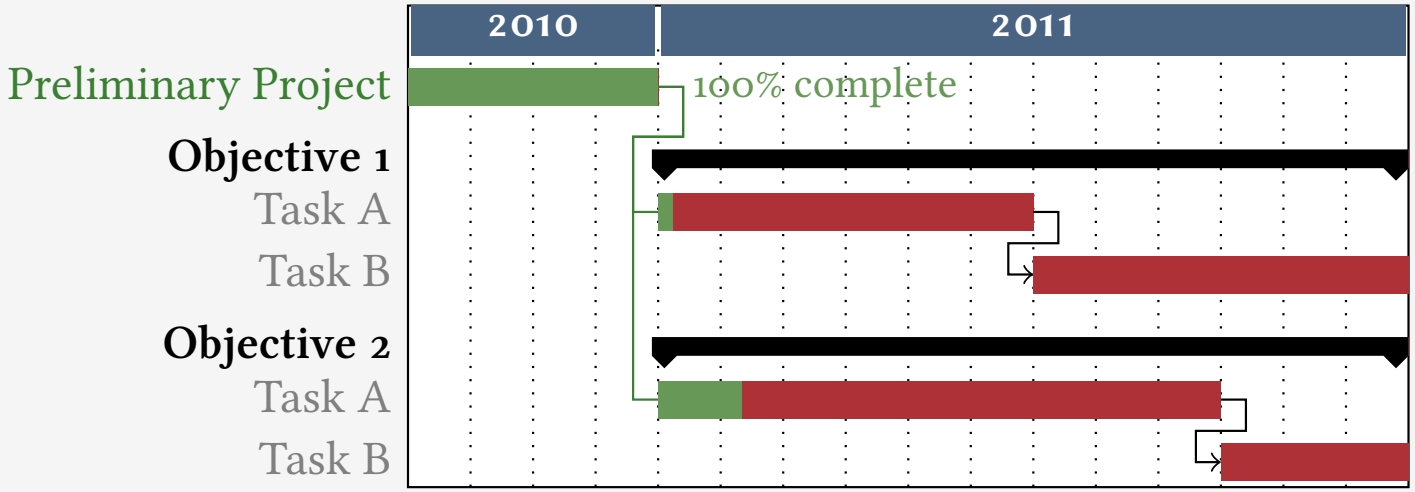
Spreadsheets

(Seriously, use a proper spreadsheet application for complex stuff.)

Year ending Mar 31	2009	2008	2007
Revenue	14580.20	11900.40	8290.30
Cost of sales	6740.20	5650.10	4524.20
Gross profit	7840.00	6250.30	3766.10

```
\STautoround*{2}
\begin{spreadtab}{{tabular}{l rrr}}
@Year ending Mar 31 & @2009 & @2008 & @2007\\ \hline
@Revenue & 14580.2 & 11900.4 & 8290.3\\
@Cost of sales & 6740.2 & 5650.1 & 4524.2\\ \cline{2-4}
@\emph{Gross profit} & \STcopy{>}{b2-b3} & &\\ \cline{2-4}
\end{spreadtab}
```

Gantt Charts




```
\usepackage{pgfgantt}
...
\begin{tikzpicture}
\begin{ganttchart}[...settings...]{16}
\ganttitle{2010}{4} \ganttitle{2011}{12} \\
\ganttbar[progress=100]{Preliminary Project}{1}{4} \\
\ganttlink[link mid=.4]{4}{2}{5}{4}  \ganttlink[link mid=.159]{4}{2}{5}{7}
\ganttgroup{Objective 1}{5}{16} \\
\ganttbar[progress=4]{Task A}{5}{10} \\
\ganttlinkedbar[progress=0]{Task B}{11}{16} \\
...
\end{ganttchart}
\end{tikzpicture}
```