TITLE WITH ALL CAPS

By

Your Full Name

Bachelor of Science, The University of Montana, Missoula, MT, 2012

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Science
in Computer Science

The University of Montana
Missoula, MT

Summer 2015

Approved by:

J.B. Alexander S. Ross Ph.D., Dean
Graduate School

John Doe Ph.D., Chair
Computer Science

Another Person Ph.D.
Computer Science

Amazing Scientist Ph.D.
Biological Sciences

Last, First M., M.S., August 2015 Computer Science

Standard Caps Title

Chairperson: John Doe

This is the short summary of why this thesis is awesome and *YOU* really want to read it!

ACKNOWLEDGMENTS

Thanks guys ...

TABLE OF CONTENTS

CODE LISTINGS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

Before starting, I *highly* recommend you take a good look at the *macros.tex* file. It contains a lot of new commands that I added to this thesis template and is extremely useful. It is also the clearing house for author-specific information that you will need to address before attempting to build this document.

Be sure to import the *dvi2pdf.tco* build configuration into TeXniCenter (if you're using that) to get the proper build sequence. It is in the root directory of this template. You can also refer to the *isuthesis.pdf* file which came from the Iowa State University thesis template upon which this template is built.

1.1 Template Issues

Many of the issues found in the previous version have been fixed, but there is no guarantee everything will work smoothly. *google is your best friend!*

I've written a DOS/Windows based build system that calls on a perl script to establish bounding boxes for the figures and then compile your thesis into a pdf. LaTeXneeds these bounding boxes in order to size the images correctly. I'm sure that any Linux user will have no problem hacking make.bat and clean.bat into bash commands in order to accommplish the same thing...

The previous version of this template used a MATLAB file to deal with figures,

which you may or may not prefer.

`saveFig.m` will save the currently active figure to *.fig*, *.jpg*, and *.eps* formats. It assumes you already have the folder hierarchy that comes with this thesis template and is mainly geared to archiving your graphics so that you don't have to completely redo them later. You also have the ability to later switch from *.eps* to *.jpg* driven graphics later via the *addfigure* macro.

If you move it from its current directory, you'll need to update the directories inside the function.

1.2 Motivation

Check out this minipage with a shadowbox around it:

> *"all algorithms that search for an extremum of a cost function will eventually choke like Mama Cass on a ham sandwich."*

and the same minipage with no shadowbox:

> *"all algorithms that search for an extremum of a cost function will eventually choke like Mama Cass on a ham sandwich."*

1.3 Goal

The goal of this thesis is ...

1.4 Benefits

...

1.5 Thesis Organization

The rest of this thesis is organized as follows:

- **Chapter 2** Literature review and overview of the framework
- **Chapter 3** Data acquisition and computational methods
- **Chapter 4** Presentation of case study results
- **Chapter 5** Discussion of results, conclusions, and future directions

CHAPTER 2 LITERATURE REVIEW

This can be a hard section to write because it will force you to learn about the field that your research is a part of. It's also a really good place to start on your thesis because the stuff you learn here will likely drive the rest of your research.

2.1 Citations

This chapter should have lots of citations and possibly several sections [1] [2] [3]. I would recommend using a tool like *JabRef* to manage your citation database in **`references.bib`**. JabRef will allow you to lookup and import citations from Search - Web Search (F5), which will open a panel on the left where you can pick a citation database like google scholar to search for journals by title.

You can check out the *natbib* package for more options to the above macros, such as including text in the reference itself.

Check out the macro.tex file that has all the macros I hacked up for this thesis. I created macros for inserting figures as well as referencing figures, tables, and equations. Use them if you wish or strike out on your own.

2.2 Equations

An unnumbered equation array:

$$\begin{aligned} O_{1,max} &= 8.0752 \quad \text{(Output 1's global maximum)} \\ O_{1,min} &= -6.5466 \quad \text{(Output 1's global minimum)} \end{aligned}$$

An in-line math environment:

$$-6.5466 \leq O_1 \leq 8.0752$$

An equation. Note that it is split across several lines of text.

$$\begin{aligned} z =& 3(1-x)^2 e^{-(x^2)-(y+1)^2} - \dots \\ & 10(\frac{x}{5} - x^3 - y^5) e^{-x^2-y^2} - \dots \\ & \frac{1}{3} e^{-(x+1)^2-y^2} \end{aligned} \tag{2.1}$$

CHAPTER 3 METHODS

The methods described below are roughly organized in the order they were performed during the case studies. Some methods are independent of the rest, so these methods are arranged at the end.

3.1 Methods Syntax And Conventions

Some of the special syntax conventions used in this document are as follows.

- *Scientific Names* - capitalized and italicized.
- **`Scripts`** - bold and Courier font.
- *Variables* - italicized
- **Functions** - bold
- *System Commands* - italicized and underlined

3.2 Code With Syntax Highlighting

The tsv file must also have lineage strings available, which can be created with the bash command in Code Listing 3.1.

Code Listing 3.1: Convert Biom Matrix To TSV

```
biom convert -i infile -o outfile --to-tsv --header-key taxonomy
```

This process is necessary because QIIME's OTU matrix output is in the biom format [1]. In order to convert a biom format matrix into a tsv matrix, a UNIX or Linux style operating system (OS) is recommended. Ubuntu 15.2 run from VirtualBox [4] was used to test the *biom* command successfully at the time of this writing. From a Linux based OS the *biom* command can be installed in a terminal window as seen in Code Listing 3.2. In order for Code Listing 3.2 to work, the *pip* command needs to be available. Pip is the PyPA recommended tool for installing python packages, which makes it a good way to install many useful and important bioinformatics tools [5].

Code Listing 3.2: Install PIP

```
pip install numpy
pip install biom-format
```

Code Listing 3.3: Find N-Dimensional Pareto Frontier

```
#this function will return the rows of the matrix
#that make up the pareto frontier
findParetoBoundary = function(m){
  rows = nrow(m)
  cols = ncol(m)
  pareto_frontier = c()
  for(point in 1:rows){
    domination = apply(m[-point,], 1, function(d)
                   paretoDomination(m[point,], d))
    if(sum(domination) == 0){ #no point dominates this point
      pareto_frontier = c(pareto_frontier, point)
    }
  }
  o = order(m[pareto_frontier,1])
  return(pareto_frontier[o])
}
```

Code Listing 3.4: Check For Pareto Domination

```
paretoDomination = function(point, dominator){
  numdims = length(point)
  if(numdims != length(dominator)){
    print("points of varying lengths can't exist in the same
    vectorspace")
    return(NULL)
  }
  greater = FALSE
  peer    = TRUE
  for(d in 1:numdims){
    if(dominator[d] > point[d]){ greater = TRUE; }
    if(dominator[d] < point[d]){ peer = FALSE; }
  }
  return(greater && peer)
}
```

CHAPTER 4 RESULTS

these results prove ...

4.1 Figures

The addfigure macro, used to typset Figure 4.1. Note that the caption in the list of tables is different than the caption below the figure.

```
\addfigure{ML_LDA_4Chamber_NoTitle}
            {1}{Mouse LDA Plot With Four Chambers}
            {long caption.}
            {fig:mlda4c}
```

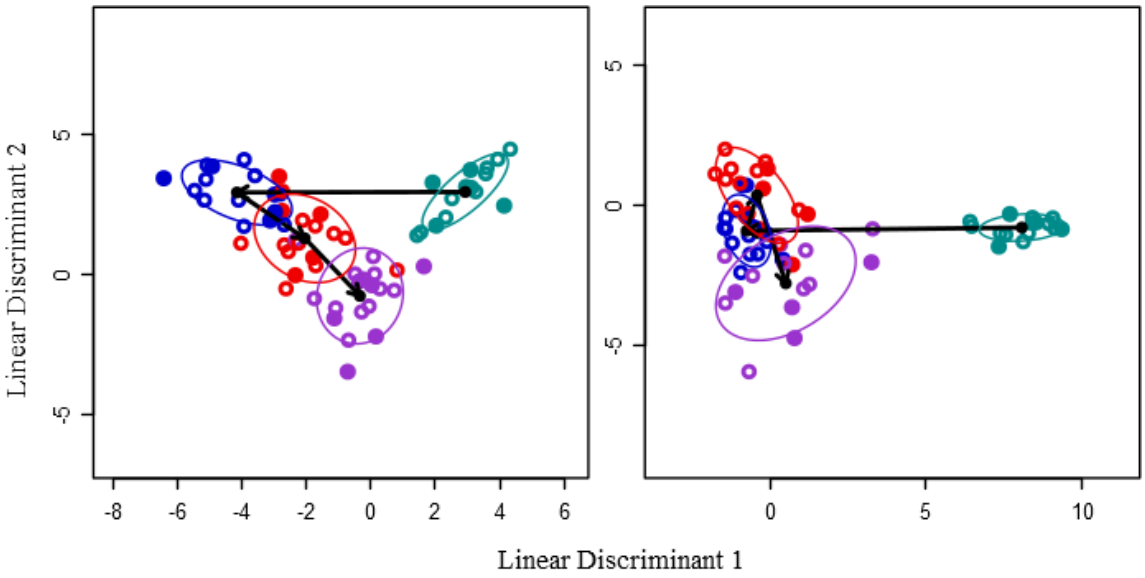

Figure 4.1: Linear Discriminant Analysis (LDA) of the four main compartments sampled from C57B1/6 strain mice (left panel) and CD-1 strain mice (right panel). Filled circles and open circles represent cohorts 1 and 2, respectively. Black dots represent the centroid for each cluster and ellipses indicate 1 standard deviation. The arrows show the flow of digesta between chambers. The plots were made using vote-determined genera shown in Tables 4.1 and 4.2. The accuracies were 78.79% (62.12%)(left panel) and 63.93% (65.57%)(right panel). The first accuracies listed used the vote-determined genera, while the right side accuracies were for genera identified using 'floating search within each fold'.

4.2 Tables

Tables can be built directly in your document, but you may find it useful to import data like a csv from an external source.

Tables 4.1 and 4.2 represent the genera identified using the voting process for the four chamber LDA plot visualized in Figure 4.1.

Table 4.1: Strain B - 14 genera identified using the voting process

Genera	Rank
Oscillibacter	6.75
Lactobacillus	6.38
Robinsoniella	6.24
Ruminococcus	5.93
Barnesiella	5.65
Dorea	5.63
Coprobacillus	4.97
Coprococcus	4.76
Butyricimonas	4.64
Blautia	4.33
Turicibacter	4.28
Mucispirillum	4.03
Anaerotruncus	3.57
Parabacteroides	3.46

Table 4.2: Strain C - 15 genera identified using the voting process

Genera	Rank
Lactobacillus	6.80
Dorea	6.54
Turicibacter	6.49
Oscillibacter	6.31
Sporacetigenium	6.23
Robinsoniella	6.01
Akkermansia	5.80
Marvinbryantia	5.63
Asaccharobacter	5.45
Anaerotruncus	5.35
Bacteroides	5.17
Butyricicoccus	5.09
Coprobacillus	4.89
Papillibacter	3.98
Sporobacter	3.60

Example of a table directly in your document.

(see Table 4.3) with multirow, multicolumn, and cline all in use.

Table 4.3: Let the water hold me down

Sample	Point Values			ΔCF	$\overline{\Delta CF}$
	x	y	z		
1	0.8191	-0.2945	0.0000	–	–
	-2.8080	-0.7454	-0.0390	-0.0390	-0.0390
2	-2.8080	-0.7454	-0.0390	–	–
	-1.7669	0.0308	-2.1083	-2.0692	-1.0541
	0.2503	-1.6039	-6.5415	-0.0055	-2.7268

CHAPTER 5 DISCUSSION

It would be a good idea to refer to every figure or table presented in chapter 4 as you discuss what it all means.

5.1 Conclusions

It is clear from the analysis of the experimental data that ...

5.2 Future Directions

Certainly a valid avenue of future research would be ...

BIBLIOGRAPHY

[1] D. McDonald, J. C. Clemente, J. Kuczynski, J. R. Rideout, J. Stombaugh, D. Wendel, A. Wilke, S. Huse, J. Hufnagle, F. Meyer *et al.*, "The biological observation matrix (biom) format or: how i learned to stop worrying and love the ome-ome," *GigaScience*, vol. 1, no. 1, p. 7, 2012.

[2] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 7, pp. 179–188, 1936.

[3] W. R. Pearson and D. J. Lipman, "Improved tools for biological sequence comparison," *Proceedings of the National Academy of Sciences*, vol. 85, no. 8, pp. 2444–2448, 1988.

[4] J. Watson, "Virtualbox: bits and bytes masquerading as machines," *Linux Journal*, vol. 2008, no. 166, p. 1, 2008.

[5] J. L. C. M. Donald Stufft, Marcus Smith *et al.* (2008, october) The pypa recommended tool for installing python packages. [Online]. Available: https://pypi.python.org/pypi/pip