

# 目 录

目 录 .....	i
图形列表 .....	iii
表格列表 .....	v
符号列表 .....	vii
<b>第一部分 L<sup>A</sup>T<sub>E</sub>X .....</b>	<b>1</b>
<b>第一章 引言 .....</b>	<b>3</b>
1.1 系统要求 .....	3
1.2 问题反馈 .....	4
<b>第二章 使用简介 .....</b>	<b>5</b>
2.1 先试试效果 .....	5
2.2 各文档及目录简介 .....	5
2.3 数学公式、图片插入、参考文献等功能 .....	7
2.4 常见使用问题 .....	12
<b>第二部分 增强学习 .....</b>	<b>17</b>
<b>第三章 理论简介 .....</b>	<b>19</b>
<b>第四章 开放模型下的策略迭代 .....</b>	<b>23</b>
4.1 价值函数的近似 .....	23
4.2 近似价值函数的设计 .....	23
4.3 采样复用 .....	23
4.4 主动学习 .....	23
4.5 稳健迭代 .....	23
<b>第五章 开放模型下的策略搜寻 .....</b>	<b>25</b>
5.1 梯度上升算法 .....	25
5.2 期望最大化算法 .....	25
5.3 先验策略 .....	25

---

<b>第六章 限定模型的增强学习 .....</b>	<b>27</b>
6.1 转移模型估计 .....	27
6.2 维度压缩 .....	27

## 图形列表

2.1	Q 判据等值面图 .....	8
2.2	Shock diffraction. ....	9
2.3	总声压级。(a) <i>A</i> , (b) <i>B</i> , (c) <i>C</i> , (d) <i>D</i> .....	10



表格列表

2.1 This is sample table..... 7



## 符号列表

### Characters

Symbol	Description	Unit
$R$	the gas constant	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
$C_v$	specific heat capacity at constant volume	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
$C_p$	specific heat capacity at constant pressure	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
$E$	specific total energy	$\text{m}^2 \cdot \text{s}^{-2}$
$e$	specific internal energy	$\text{m}^2 \cdot \text{s}^{-2}$
$h_T$	specific total enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
$h$	specific enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
$k$	thermal conductivity	$\text{kg} \cdot \text{m} \cdot \text{s}^{-3} \cdot \text{K}^{-1}$
$T$	temperature	K
$t$	time	s
$p$	thermodynamic pressure	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
$\hat{p}$	hydrostatic pressure	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
$\mathbf{f}_b$	body force	$\text{kg} \cdot \text{m}^{-2} \cdot \text{s}^{-2}$
$S$	boundary surface	$\text{m}^2$
$V$	volume	$\text{m}^3$
$\mathbf{V}$	velocity vector	$\text{m} \cdot \text{s}^{-1}$
$u$	x component of velocity	$\text{m} \cdot \text{s}^{-1}$
$v$	y component of velocity	$\text{m} \cdot \text{s}^{-1}$
$w$	z component of velocity	$\text{m} \cdot \text{s}^{-1}$
$c$	speed of sound	$\text{m} \cdot \text{s}^{-1}$
$\mathbf{r}$	position vector	m
$\mathbf{n}$	unit normal vector	1
$\hat{\mathbf{t}}$	unit tangent vector	1
$\tilde{\mathbf{t}}$	unit bitangent vector	1
$C_R$	coefficient of restitution	1
$Re$	Reynolds number	1
$Pr$	Prandtl number	1

$Ma$	Mach number	1
$\alpha$	thermal diffusivity	$\text{m}^2 \cdot \text{s}^{-1}$
$\mu$	dynamic viscosity	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-1}$
$\nu$	kinematic viscosity	$\text{m}^2 \cdot \text{s}^{-1}$
$\gamma$	heat capacity ratio	1
$\rho$	density	$\text{kg} \cdot \text{m}^{-3}$
$\sigma_{ij}$	stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
$S_{ij}$	deviatoric stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
$\tau_{ij}$	viscous stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
$\delta_{ij}$	Kronecker tensor	1
$I_{ij}$	identity tensor	1

## Operators

Symbol	Description
$\Delta$	difference
$\nabla$	gradient operator
$\delta^\pm$	upwind-biased interpolation scheme

## Abbreviations

Acronym	Description
ANFO	Ammonium Nitrate Fuel Oil
CFD	Computational Fluid Dynamics
CFL	Courant-Friedrichs-Lewy
CJ	Chapman-Jouguet
EOS	Equation of State
JWL	Jones-Wilkins-Lee
TVD	Total Variation Diminishing
WENO	Weighted Essentially Non-oscillatory
ZND	Zel'dovich-von Neumann-Doering



# 第一部分

L<sup>A</sup>T<sub>E</sub>X



# 第一章 引言

考虑到大多数用户并无  $\text{\LaTeX}$  使用经验，本模板将  $\text{\LaTeX}$  的复杂性尽可能地进行了封装，开放出简单的接口，以便于使用者可以轻易地使用。同时，对使用  $\text{\LaTeX}$  撰写论文所遇到的一些主要难题，如插入图片、文献索引等，进行了详细的说明，并提供了相应的代码样本，理解了上述问题后，对于初学者而言，使用此模板撰写其学位论文将不存在实质性的困难。所以，如果您是初学者，请不要直接放弃，因为同样作为初学者的我，十分明白让  $\text{\LaTeX}$  变得简单易用的重要性，而这正是本模板所体现的。

此中国科学院大学学位论文模板 `ucasthesis` 基于吴凌云的 `CASthesis` 模板发展而来，`ucasthesis` 文档类的基础架构为 `ctexbook` 文档类。当前 `ucasthesis` 模板满足最新的中国科学院大学学位论文撰写要求和封面设定。模板兼顾不同操作系统 (Windows, Linux, Mac OS) 并兼容 `pdflatex` 和 `xelatex` 编译方式，完美地支持中文书签、中文渲染、中文粗体显示、拷贝 pdf 中的文本到其他文本编辑器等特性，此外，对模板的文档结构进行了精心设计，撰写了编译脚本提高模板的易用性和使用效率。

宏包的目的是简化学位论文的撰写，模板文档的默认设定是十分规范的，从而论文作者可以将精力集中到论文的内容上，而不需要在版面设置上花费精力。同时，在编写模板的  $\text{\LaTeX}$  文档代码过程中，作者对各结构和命令进行了十分详细的注解，并提供了整洁一致的代码结构，对文档的仔细阅读可以为初学的您提供一个学习  $\text{\LaTeX}$  的窗口。除此之外，整个模板的架构十分注重通用性，事实上，本模板不仅是中国科学院大学学位论文模板，同时，也是使用  $\text{\LaTeX}$  撰写中英文 `article` 或 `book` 的通用模板，并为使用者的个性化设定提供了接口和相应的代码。

## 1.1 系统要求

`ucasthesis` 宏包可以在目前大多数的  $\text{\TeX}$  编译系统中使用，例如 `CTEX`、`MiKTeX`、`TeXLive`。推荐的  $\text{\TeX}$  编译系统 + 文本编辑器为

- Linux: `TeXLive` + `vim` or `Texmaker`
- MacOS: `TeXLive` or `MacTeX` + `Macvim` or `Texmaker`
- Windows: `TeXLive` or `MikTeX` + `Texmaker`

$\text{\TeX}$  编译系统 (如 `MiKTeX`、`TeXLive`) 用于提供编译环境，文本编辑器 (如 `Texmaker`、`vim`) 用于编辑  $\text{\TeX}$  源文件。

## 1.2 问题反馈

莫晃锐 (mohuangrui)    mohuangrui@gmail.com

模版下载地址: <https://github.com/mohuangrui/ucasthesis>

欢迎大家反馈模板不足之处，一起不断改进模板。希望大家向同事积极推广 $\text{\LaTeX}$ ，一起更高效地做科研。

## 第二章 使用简介

为方便使用及更好地展示 L<sup>A</sup>T<sub>E</sub>X 排版的优秀特性，本人对模板的框架和文件体系进行了细致地处理，尽可能地对各个功能和板块进行了模块化和封装，对于初学者来说，众多的文件目录也许会觉得有些无所适从，但阅读完下面的使用说明后，您会发现原来使用思路是简单而清晰的，而且，当对 L<sup>A</sup>T<sub>E</sub>X 有一定的认识和了解后，会发现其相对 Word 类排版系统的极具吸引力的优秀特性。所以，如果您是初学者，请不要退缩，请稍加尝试和坚持，让自己领略到 L<sup>A</sup>T<sub>E</sub>X 的非凡魅力，并可以通过阅读相关资料如 Wikibook<sup>[?]</sup> 来完善自己的使用知识。

### 2.1 先试试效果

ucasthesis 模板不仅只是提供了相应的类文件，同时也提供了包括参考文献等在内的完成学位论文的一切要素，所以，下载时，推荐下载整个 ucasthesis 文件夹，而不是单独的文档类。

下载 ucasthesis 文件夹并解压后，请在文件夹内找到 Compile.bat，双击运行，即可获得本说明文档，而这，也完成了学习使用此模板撰写论文的一半进程，什么？这就学成一半了，这么简单???，是的，就这么简单！

编译完成后，可以进入各个子目录逛逛，熟悉下模板框架。

### 2.2 各文档及目录简介

#### 2.2.1 Thesis.tex 文档

Thesis.tex 文档为主文档，其设计和规划了论文的整体框架，通过对其的阅读可以让用户了解整个论文框架的搭建。

#### 2.2.2 Compile.bat

Compile.bat 为编译此模板的 Dos 脚本，通过双击运行此脚本即可获得编译后的 PDF 文档，编译生成的文档及临时文件皆位于 Tmp 文件夹内。在此脚本中可以设定编译器为 pdf<sub>l</sub>atex or xelatex(默认设定，推荐)。

#### 2.2.3 Tmp 文件夹

运行编译脚本 Compile.bat 后，编译所生成的文档皆存于 Tmp 文件夹内，包括编译得到的 pdf 文档，其存在是为了保持工作空间的整洁，因为好的心情是很

重要的.

#### 2.2.4 Style 文件夹

Style 文件夹内包含 ucasthesis 文档类的定义文件和配置文件, 对于有特殊需求的用户, 通过对它们的修改可以实现特定的类设定。用户若需更新模板, 一般只需用新的样式文件替换旧的即可。

1. ucasthesis.cls: 文档类定义文件, 论文的最核心的格式即通过它来定义的。
2. ucasthesis.cfg: 文档类配置文件, 通过它设定论文的某些项目的显示内容, 如 abstract 显示为摘要, table of content 显示为目 录而不是目录等 (如果愿意, 你也可以改过来)。
3. commons.sty: 常用宏包的加载及文档的设定, 如参考文献样式, 文献引用样式, 页眉页脚设定等。模板为这些功能提供了开关选项, 从而只需在 Thesis.tex 中的 `\usepackage[options]{commons}` 中进行启用即可, 而一般无需修改 commons.sty 本身。
4. custom.sty: 用来实现一些个性化设定, 用户自定义命令以及添加宏包的推荐放置位置。

#### 2.2.5 Tex 文件夹

Tex 文件夹内为论文的所有实体内容, 正常情况下, 这也是你使用此模板撰写学术论文时, 主要关注和修改的一个位置, 注: 所有文件都必须采用 UTF-8 编码, 否则编译后将出现乱码文本, 详细分类介绍如下:

- Frontpage.tex: 为论文封面内容及中英文摘要。
- Main\_Content.tex: 对需要出现的 Chapter 进行索引, 开始写论文时, 可以只索引当前章节, 以便快速编译和查看, 当最终所有章节完成后, 再对所有章节进行索引即可。
- Chap\_XXXXX.tex: 为论文主体的各个章节, 用户可根据需要添加和撰写, 最终需要包含在论文中的章节, 须在 Main\_Content.tex 中进行索引。
- Appendix.tex: 为附录内容
- Backmatter.tex: 为发表文章信息, 致谢部分等。

### 2.2.6 Img 文件夹

Img 文件夹用于放置论文中所需要的图类文件，支持格式有：.jpg, .png, .pdf。其中，ucas.pdf 为国科大校徽。不建议再为各个章节的图片建立子目录，即使图片众多，若命名规则合理，各个案例的图片仍将有序的聚集在一起，查询亦是十分方便。若坚持引入子目录以增加额外约束条件，则需在 commons.sty 文件的 291 行附近对增加的子目录进行索引：

```
\graphicspath{{Img/}{Img/subdir1}{Img/subdir2/}{Img/subdirn/}}
```

### 2.2.7 Biblio 文件夹

Biblio 文件夹用于放置参考文献的索引信息文件：ref.bib，此文件包含需要引用的参考文献信息。文件夹内包含符合国标的参考文献样式文件（从 zepinglee/gbt-7714-2015 <https://github.com/zepinglee/gbt-7714-2015> 引入，建议用户追踪其更新）。

## 2.3 数学公式、图片插入、参考文献等功能

### 2.3.1 数学公式

Navier-Stokes equations:

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) = 0 \\ \frac{\partial (\rho \mathbf{V})}{\partial t} + \nabla \cdot (\rho \mathbf{V} \mathbf{V}) = \nabla \cdot \boldsymbol{\sigma} \\ \frac{\partial (\rho E)}{\partial t} + \nabla \cdot (\rho E \mathbf{V}) = \nabla \cdot (k \nabla T) + \nabla \cdot (\boldsymbol{\sigma} \cdot \mathbf{V}) \end{cases} \quad (2.1)$$

Row number	This is a multicolumn						
Row 1	1	2	4	5	6	7	8
Row 2	1	2	4	5	6	7	8
Row 3	1	2	4	5	6	7	8
Row 4	1	2	4	5	6	7	8

表 2.1: This is sample table

常用数学公式的命令代码模板，请见 WiKibook：<https://en.wikibooks.org/wiki/LaTeX/Mathematics>。custom.sty 中定义了一系列数学命令，使用它们可以提高数学代码对不同样式的适应性。

### 2.3.2 图片插入

论文中图片的插入通常分为单图和多图，下面分别加以介绍：

单图插入：假设插入名为 ITC\_Q\_Criteria（后缀可以为.jpg、.png、.pdf，下同）的图片，其效果如图2.1，其命令可为：

```
\begin{figure}[!htbp]
  \centering
  \includegraphics[width=0.45\textwidth]{ITC_Q_Criteria}
  \caption{Q判据等值面图}
  \label{fig:ITC_Q_Criteria}
\end{figure}
```

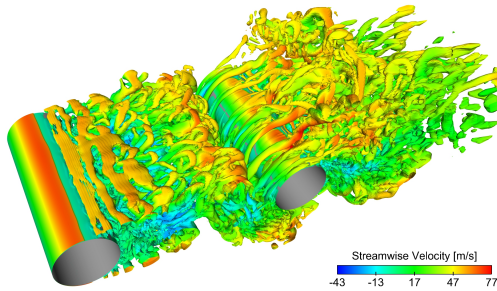


图 2.1: Q 判据等值面图

如果插图的空白区域过大,希望减少插入图片后的留白,以图片 Y 为例(图2.2),可以使用如下代码模板：

```
\begin{figure}[!htbp]
  \centering
  %trim option's parameter order: left bottom right top
  \includegraphics[trim = 30mm 0mm 30mm 0mm, clip, width=0.45\textwidth]{Y}
  \caption{Shock diffraction}
  \label{fig:Y}
\end{figure}
```

多图的插入如图2.3，其代码如下。

```
\begin{figure}[!htbp]
  \centering
  \begin{subfigure}[b]{0.45\textwidth}
    \includegraphics[width=\textwidth]{HC_OASPL_A}
```



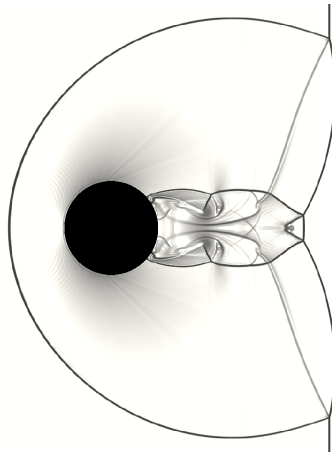


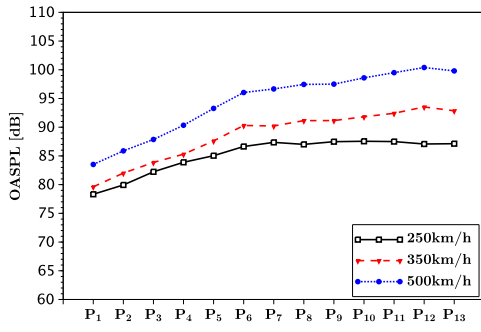
图 2.2: Shock diffraction.

```

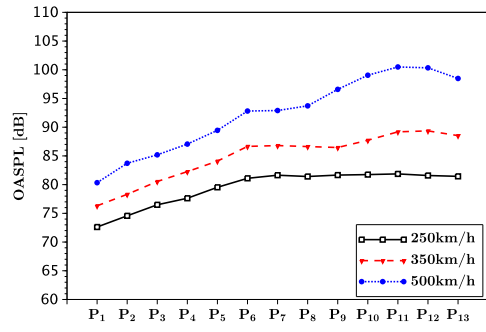
\caption{}
\label{fig:HC_OASPL_A}
\end{subfigure}%
~%add desired spacing
\begin{subfigure}[b]{0.45\textwidth}
\includegraphics[width=\textwidth]{HC_OASPL_B}
\caption{}
\label{fig:HC_OASPL_B}
\end{subfigure}
\begin{subfigure}[b]{0.45\textwidth}
\includegraphics[width=\textwidth]{HC_OASPL_C}
\caption{}
\label{fig:HC_OASPL_C}
\end{subfigure}%
~%add desired spacing
\begin{subfigure}[b]{0.45\textwidth}
\includegraphics[width=\textwidth]{HC_OASPL_D}
\caption{}
\label{fig:HC_OASPL_D}
\end{subfigure}
\caption{总声压级。 (a)  $A$ , (b)  $B$ , (c)  $C$ , (d)  $D$ }
\label{fig:HC_OASPL}
\end{figure}

```

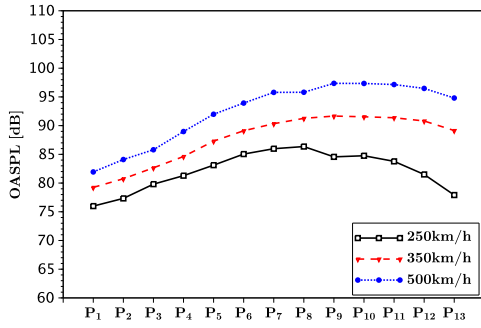
撰写论文中，插图和制表常用到的命令，已在 **Useful Commands.txt** 这个



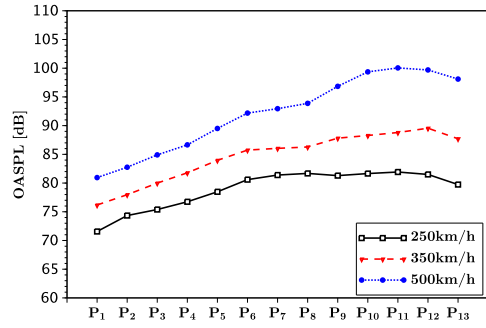
(a)



(b)



(c)



(d)

图 2.3: 总声压级。(a)  $A$ , (b)  $B$ , (c)  $C$ , (d)  $D$

文本中给出了参考代码，大家只需 copy 使用即可。

### 2.3.3 参考文献引用

参考文献引用过程以实例进行介绍，假设需要引用名为 Document Preparation System 的文献，步骤如下：

1) 使用 google scholar 搜索 Document Preparation System，在目标条目下点击 Cite，展开后选择 Import into BibTeX 打开此文章的 BibTeX 索引信息，将它们 copy 添加到 ref.bib 文件中（此文件位于 Biblio 文件夹下）。

2) 你会发现索引信息中第一行为 `@article{lampport1986document}`，其中 `lampport1986document` 即为此文献的 label（中文文献也必须使用英文 label，一般遵照：姓氏拼音 + 年份 + 标题第一字拼音的格式），想要在论文中索引此文献，有两种索引模式：

textual: `\citet{lampport1986document}`。正如此处所示<sup>[?]</sup>；

parenthetical: `\citep{lampport1986document}`。正如此处所示<sup>[?]</sup>。

**多文献索引用英文逗号隔开：**

`\citep{lampport1986document,chen2005zhulu}`。正如此处所示<sup>[? ?]</sup>

如此，即完成了文献的索引，请查看下本文档的参考文献一章，看看是不是就是这么简单呢？是的，就是这么简单！

不同文献样式和引用样式可在 Thesis.tex 中对 commons.sty 设置实现，如：

`\usepackage[numbered]{commons}` % default citation style. textual: Jones [1]; parenthetical: [1]

`\usepackage[authoryear]{commons}` % author year citation style. textual: Jones (1995); parenthetical: (Jones, 1995)

`\usepackage[alpha]{commons}` % alpha citation style. textual: not available; parenthetical: [Jon95]

若需将所有的上标改为嵌入式标注，则可在 commons.sty 174 行附近使用

`\RequirePackage[square,comma,numbers,sort&compress]{natbib}`

的设置替换

`\RequirePackage[square,comma,super,sort&compress]{natbib}`

如只希望在某些特定情形将上标改为嵌入式标注，则可使用

textual: `\citepns{lampport1986document,chen2005zhulu}`。正如此处所示<sup>[? ?]</sup>

parenthetical: `\citetns{lampport1986document,chen2005zhulu}`。正如此处所示<sup>[? ?]</sup>

参考文献索引更为详细的信息，请见 Wikibook<sup>[?]1</sup>。

## 2.4 常见使用问题

1. 模板文档的编码为 UTF-8 编码。所有文件都必须采用 UTF-8 编码，否则编译后生成的文档将出现乱码文本。若出现文本编辑器无法打开文档或打开文档乱码的问题，请检查您使用的编辑器对 UTF-8 编码的支持，如果使用 WinEdt 作为文本编辑器，应在其

options -> Preferences -> wrapping

选项卡下将两种 Wrapping Modes 中的内容：

TeX;HTML;ANSI;ASCII|DTX...

修改为：

TeX;**UTF-8|ACP**;HTML;ANSI;ASCII|DTX...

同时，取消

options -> Preferences -> Unicode

中的 Enable ANSI Format... 选项。

2. 推荐选择 xelatex 编译引擎编译。Compile.bat 的默认设定为 xelatex 编译引擎。你也可以选择不使用此脚本编译，如直接使用 T<sub>E</sub>X 文本编辑器编译。注：T<sub>E</sub>X 文本编辑器编译的默认设定为 pdf<sub>l</sub>atex 编译引擎，若选择 xelatex 编译引擎，请进入下拉菜单进行选择。为正确生成引用链接，编译步骤为：xelatex + bibtex + xelatex + xelatex。

### 3. Texmaker 使用简介

- (a) 使用 Texmaker 打开文档 Thesis.tex。
- (b) 菜单 Options -> Define Current Document As 'Master Document'
- (c) 菜单 User -> User Commands -> Edit User Commands -> Input Menu Item as 'Auto Build' -> Click 'wizard' -> add: xelatex + bibtex + xelatex + xelatex + pdf viewer -> Click 'OK'
- (d) 使用 Auto Build 编译带有未生成引用链接的源文件，可以仅使用 xelatex 编译带有已经正确生成引用链接的源文件。
- (e) 编译完成，View PDF，在 pdf 中'ctrl+click' 可链接到相对应的源文件。

4. 若编译过程中出现无法找到某些 package 的错误, 如无法找到 xcolor.sty, mathtools.sty, ctexbook.sty, newtext.sty 等, T<sub>E</sub>X 编译程序一般可以自动下载和安装相应的文件, 否则, 请进入 L<sup>A</sup>T<sub>E</sub>X 软件的 Package Manager (Admin) 确认启用 Repository-Synchronize 状态。下次编译过程中 T<sub>E</sub>X 编译程序一般将自动下载安装 L<sup>A</sup>T<sub>E</sub>X 宏包库。
5. 模版在设计之初就尽可能地考虑了适应性。致谢, 简历及攻读学位期间发表的学术论文与科研成果等几乎所有条目都是通过最为通用的

`\chapter{item name} and \section*{item name}`

来显式实现的 (请仔细观察下 Frontpage.tex, Prematter.tex, Backmatter.tex), 从而你可以随意添加, 放置, 和修改他们, 如同一般章节。对于图表目录名称则可在 ucasthesis.cfg 中进行修改。

6. 设置正文行距: 在 custom.sty 99 行附近, 修改

`\linespread{1.3}`

设置参考文献行距: 在 custom.sty 103 行附近, 修改/注掉

`\setlength{\bibsep}{0.0pt plus 0.3ex}`

将 subsection 显示到目录当中: 在 custom.sty 107 行附近, 将 1 改为 2 就可以了

`\setcounter{tocdepth}{1}% the depth for the Table of Contents.`

如果需设置图 2.3 为图 2-3, 可将如下命令添加 custom.sty 中:

`\renewcommand{\theequation}{\arabic{chapter}-\arabic{equation}}`

`\renewcommand{\thefigure}{\arabic{chapter}-\arabic{figure}}`

`\renewcommand{\thetable}{\arabic{chapter}-\arabic{table}}`

7. 字体控制。如果对字体控制有较高需求, 请选择 xelatex 编译引擎, 并在 commons.sty 中设置需要的字体, 如启用 Times New Roman 作为英文字体, 在 commons.sty 的 105 行附近设置:

`\setmainfont{Times New Roman}`

8. 在某些情况下拷贝 pdf 文档内容到 word 时存在乱码。解决方式是选择安装 adobe 相应的字体库, 请在公共网站 (如百度云盘: <http://pan.baidu.com/share/home?uk=3188136325&view=share#category/type=0>) 搜索并下载如下四种中文字体文件:

- (a) AdobeFangsongStd-Regular.otf (adobe 仿宋)
- (b) AdobeHeitiStd-Regular.otf (adobe 黑体)
- (c) AdobeKaitiStd-Regular.otf (adobe 楷体)
- (d) AdobeSongStd-Light.otf (adobe 宋体)

下载字体文件后，双击安装相应字体。

在 Thesis.tex 中设置启用 adobe 的字体：

```
\documentclass[doublesided,fontset=adobe]{Style/ucasthesis}%
```

如果 L<sup>A</sup>T<sub>E</sub>X 软件版本比较老旧，如 Linux 用户，ctex 宏包没有更新，设置启用 adobe 的字体则为：

```
\documentclass[doublesided,adobefonts]{Style/ucasthesis}%
```

最后选择 xelatex 编译引擎编译。

因为模版的设定考虑兼顾不同操作系统 (Windows, Linux, Mac OS) 并兼顾 pdf<sub>l</sub>atex 和 xelatex，为了模版的健壮性，上述方案并未作为原始设定。

9. 页眉页脚的设定在 commons.sty 的底部。始于 323 行附近的 frontmatterstyle, mainmatterstyle, 和 backmatterstyle 分别用于定义前言，主要内容，和附录的页眉页脚样式。一般默认情况下每一章的第一页不应显示页眉页脚，若想修改此行为，请将 377-379 行附近的 plain 样式定义注空即可。即修改为

```
\fancypagestyle{plain}{%
  %\fancyhf{}% clear fields
  %\renewcommand{\headrulewidth}{0pt}% header rule
  %\renewcommand{\footrulewidth}{0pt}% footer rule
}
```

关于页眉页脚各个命令的作用和意义请参见 fancyhdr 的用户文档 <https://www.ctan.org/pkg/fancyhdr?lang=en>。如果需要在页眉页脚中添加章节字样，请使用

- (a) \CTEXthechapter 显示：第 X 章
- (b) \CTEXthesection 显示：第 X 节

参见 ctex 宏包用户文档 <http://ctan.mirror.rafael.ca/language/chinese/ctex/ctex.pdf>

10. 一般规范下，每一章应开始于奇数页。从而若前一章结束于奇数页，则一空白页将被插入以保证上述规则。如果想修改规则以取消空白页，有如下三种方案：

- 在 `thesis.tex` 的 `documentclass` 中使用 `singlesided` 替代 `doublesided` 选项。这一命令使文档不区分奇偶页，因此章可以开始于任意页。此方案将移除所有的空白页，包括封面处的。同时，页面页脚的设定不再区分奇偶页。
- 可以在 `ucasthesis.cls` 文件中 106 行附近，将 `cleardoublepage` 命令的定义修改为：

```
\def\cleardoublepage{\clearpage}
```

这一命令使产生空白页的机制失效。这一方案将移除所有的空白页，包括封面处的。但与方案一不同的是，页面页脚的设定可以区分奇偶页。

- 在 `thesis.tex` 的 `documentclass` 中添加 `openany` 选项 (`openany` 与 `doublesided` 和 `printcopy` 都可搭配)。这一命令使章可以开始于任意页。同时，将 `custom.sty` 中 86 行和 `thesis.tex` 中 88 行附近的 `cleardoublepage` 改为 `clearpage`。此方案将移除所有的用于调整章的起始位置的空白页，而不包括封面处的。同时，页面页脚的设定可以区分奇偶页。

无论哪种方案都需要注意对页眉页脚的影响并做出合适的调整。个人的推荐是采用默认设置，尽量避免将精力花在这些无关紧要的细节上。 $\text{\LaTeX}$  的特点是标准化，而其导致的问题则是任何脱离标准的修改都将花费相当的精力。对于电子档的论文，在 `thesis.tex` 的 `documentclass` 中，若不想使用 `doublesided`，则可使用 `singlesided` 来减少空白页。而对于打印版，启用 `printcopy` 选项以替换 `doublesided/singlesided` 选项，这样可使奇偶页的排版在打印装订后更美观。

11. 若 `pdflatex` 编译出现 `pdfTeX error (font expansion): auto expansion is only possible with scalable fonts`。是因为 `MikTeX` 安装后字体配置异常，请进入软件的 `package manager` 更新你的  $\text{\LaTeX}$  宏包库，并请进入

```
C:\Program Files\MikTeX 2.9\miktex\bin\x64
```

找到并运行 `updmap.exe`。

12. 部分同学留意到一个所谓的新的模板要求，那个模板是一位学生发布的，而其英文封面的设定不满足中国科学院大学学位论文封面设定。因中国科学院大学学位论文撰写规定只是限定了封面格式但是没有要求具体的内容格式，部分所因为图方便省事就直接将其作为了模板发布到了所内。但是，这

今为止学校官网上的撰写规定和模板 (<http://onestopucas.edu.cn/home/info/abc167cb-4589-4e05-b014-052fa9291d0c/1>) 未做出任何修改，亦未发表任何官方声明表示模板需要修改。所以，当前模板将不会在官方版发布之前做出调整。



## 第二部分

### 增强学习



### 第三章 理论简介

这部分是对 Masashi Sugiyama 的《Statistical Reinforcement Learning: Modern Machine Learning Approaches》的笔记。增强学习试图解决未知环境下的决策问题。增强学习可以概括为：两个要素（**环境，玩家**），三种信号（**状态，行动，反馈**）。在一个未知的环境下，玩家基于**策略**选择行动，之后，环境更新状态，并给予玩家反馈。基于环境和玩家的互动，在没有任何指令的情况下，玩家也可以完成指定任务。通常情况下，我们通过马尔可夫决策过程来建模增强学习任务。在任一离散时间  $t$ ，玩家观察到环境状态信号  $s_t \in S$ ，给出其行动信号  $a_t \in A$ ，相应的，环境更新状态信号  $s_{t+1} \in S$ ，给出即时反馈信号  $r$ 。

$$r_t = r(s_t, a_t, s_{t+1})$$

状态信号属于状态空间，行动信号属于行动空间。 $r(s_t, a_t, s_{t+1})$  是即时反馈方程。

环境的初始状态  $s_1$  符合概率分布。如果状态空间是离散的，则初始概率分布可以通过概率质量函数计算  $P(s)$

$$0 \leq P(s) \leq 1 \quad \forall s \in S$$

$$\sum_{s \in S} P(s) = 1$$

如果状态空间是连续的，则初始概率分布可以通过概率密度函数计算

$$p(s) \geq 0 \quad \forall s \in S$$

$$\int_{s \in S} p(s) ds = 1$$

事实上，概率质量函数可以通过概率密度函数计算（基于 Dirac Delta Function），因此，之后我们只关注连续的状态空间。

$$p(s) = \sum_{s' \in S} \delta(s' - s) P(s')$$

环境的动态变化可以用条件概率密度表达成状态的转移概率分布（transition probability distribution）

$$p(s'|s, a) \geq 0 \quad \forall s, s' \in S \quad \forall a \in A$$

$$\int_{s' \in S} p(s'|s, a) ds' = 1 \quad \forall s \in S \quad \forall a \in A$$

玩家通过**策略**  $\pi$  来决定行动。当采取确定策略行动时，我们可以把策略看成状态的函数。

$$\pi(s) \in A \quad \forall s \in S$$

行动空间可以是离散的也可以是连续的。当处理比较复杂的增强学习问题时，随机策略是更好的选择。此时，在某一状态下，玩家的行动选择存在概率分布。随机策略可以表达成行动和状态的条件概率。

$$\pi(a|s) \geq 0 \quad \forall s \in S \quad \forall a \in A$$

$$\int_{a \in A} \pi(a|s) da = 1 \quad \forall s \in S$$

随机策略可以帮助玩家更好的探索环境的所有状态。玩家和环境的一系列互动([状态, 行动]) 被称为**经历**。

互动的次数可以是有限的，也可以是无限的。在此，我们只关注有限互动的增强学习任务。一次经历  $h$  可以被表达成

$$h = [s_1, a_1, \dots, s_T, a_T, s_{T+1}]$$

一次经历  $h$  的**回报**  $R$  是衰减累计反馈。 $\gamma$  是衰减因子。

$$R(h) = \sum_{t=1}^T \gamma^{t-1} r(s_t, a_t, s_{t+1})$$

增强学习的目的是促使玩家形成最佳策略  $\pi_*$ ，来获得最大的经历回报。

$$\pi_* = \arg \max_{\pi} \mathbb{E}_{p^{\pi}(h)}[R(h)]$$

$\mathbb{E}_{p^{\pi}(h)}$  是一次经历的期望， $p^{\pi}(h)$  是策略  $\pi$  下的经历  $h$  的概率密度。

$$p^{\pi}(h) = p(s_1) \prod_{t=1}^T p(s_{t+1}|s_t, a_t) \pi(a_t|s_t)$$

增强学习可以分为两类，一类基于限定模型，一类基于开放模型。这里**模型**指代上面提及到的转移概率  $p(s'|s, a)$ 。限定模型的增强学习，转移概率是先验的，转移模型是显性地用于学习行动策略。开放的增强学习，玩家不能直接使用转移概率来寻找策略。一般来说，如果存在很强关于转移概率的先验知识，限定模型的方法更合适。如果不存在很好的先验知识，开放模型的增强学习更适合。

对于限定模型的增强学习，环境的转移模型是先验的，玩家不需要花费额外的精力来理解环境变化。这种类型非常适合于数据充足的任務。

**策略迭代**是开放模型下增强学习最流行的求解算法。策略迭代使用**价值函数**。价值函数  $Q^\pi(s, a)$  是关于状态  $s$  和行动  $a$  的函数, 是某状态, 行动, 和策略的预期回报。

$$Q^\pi(s, a) = \mathbb{E}_{p^\pi(h)}[R(h)|s_1 = s, a_1 = a]$$

最优价值函数是拥有最大价值的状态和行动组合  $Q_*(s, a) = \max_\pi Q^\pi(s, a)$ 。最优策略  $\pi_*$

$$\pi_*(a|s) = \delta(a - \arg \max_{a'} Q_*(s, a'))$$

最优价值函数是未知, 我们通过策略迭代来估计策略的价值  $Q^\pi$ , 然后基于策略价值  $Q^\pi$  更新策略  $\pi$ , 通过不断的迭代, 是价值函数收敛到最优  $Q_*(s, a)$ 。一般来说, 策略迭代的效果取决于价值评估的效力。但是, 提高价值评估的效力并不意味着策略质量的提升, 而且微小的价值评估变化可能会带来不同的策略选择。此外, 当行动空间是连续的情况, 策略迭代算法计算量巨大, 非常难以训练。为了克服策略迭代的缺点, **策略搜寻**不采用价值评估, 直接学习最大化预期回报的策略函数。在策略搜寻中, 最关键的问题是如何从海量的策略空间中筛选出好的策略。

$$\pi_* = \arg \max_{\pi} \mathbb{E}_{p^\pi(h)}[R(h)]$$



## 第四章 开放模型下的策略迭代

4.1 价值函数的近似

4.2 近似价值函数的设计

4.3 采样复用

4.4 主动学习

4.5 稳健迭代





## 第五章 开放模型下的策略搜寻

5.1 梯度上升算法

5.2 期望最大化算法

5.3 先验策略



## 第六章 限定模型的增强学习

### 6.1 转移模型估计

### 6.2 维度压缩

