Sample Book Title

Sample Book Title

First Author's Name
Affiliation for the First Author

Second Author's Name
Affiliation for the Second Author

SYNTHESIS LECTURES ON XYZ #13

MORGAN & CLAYPOOL PUBLISHERS

ABSTRACT

The abstract goes here.

The Abstract and the keywords have to fit in this page.

KEYWORDS

xxx, yyyy, zz

Contents

Preface

This is the Preface.

First Author's Name and Second Author's Name
December 2015

Acknowledgments

This the Acknowledgments' page.

First Author's Name and Second Author's Name
December 2015

CHAPTER 1

Information Theory and Rate Distortion Theory

Please see below examples of equations using \begin{equation} ...\end{equation} and align.

The M&C macros support various environments like: constraint, construction, convention, conventions, corollary, definition, dictionary, example, lemma, note, notation, observation, property, proposition, remark, rules, theorem, etc.

What if you need an environment that the macros do not provide, like "guesswork?" You create it with newMCtheorem:

newMCtheorem}{guesswork}{GuessWork}

and use then use it:

GuessWork 1.1 This is pure guess work.

1.1 INTRODUCTION

Thus far in the book, the term *information* has been used sparingly and when it has been used, we have purposely been imprecise as to its meaning.

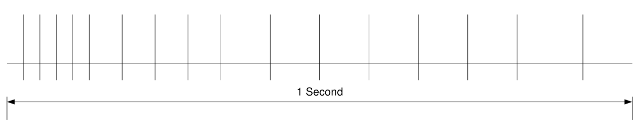

Figure 1.1: Communication system block diagram.

Examples of side-by-side figures.

1.2 ENTROPY AND AVERAGE MUTUAL INFORMATION

Consider a discrete random variable U that takes on the values $\{u_1, u_2, \ldots, u_M\}$, where the set of possible values of U is often called the *alphabet* and the elements of

Figure 1.2: Communication system block diagram.

(a) Annotated visualization of the structure of a biological neuron, reconstructed from electron microscope images of 30nm-thick slices of a mouse brain Tesnière [1959].

(b) The shape of an action potential. A small external voltage stimulus (blue) triggers a cascade of charge build-up inside a neuron (red) via voltage-gated ion channels. The activation threshold is shown as a dotted line. Simulated using a Hodgkin-Huxley model of a neuron Weber [1997].

Figure 1.3: The structure and operation of biological neurons.

the set are called *letters* of the alphabet. Let $P_U(u)$ denote the probability assignment over the alphabet, then we can define the *self-information* of the event $u = u_j$ by

$$I_U\left(u_j\right) = \log \frac{1}{P_U(u_j)} = -\log P_U\left(u_j\right). \tag{1.1}$$

Example 1.2 Given a random variable U with four equally likely letters in its alphabet, we wish to find $H(U)$. Clearly, $M = 4$ and $P_U(u_i) = \frac{1}{4}$ for $i = 1, 2, 3, 4$.

$$\begin{aligned} I_{W;X}\left(w_j; x_k\right) &= \log \frac{P_{WX}\left(w_j, x_k\right)}{P_W\left(w_j\right) P_X\left(x_k\right)} \\ &= \log \frac{P_{X|W}\left(x_k | w_j\right)}{P_X\left(x_k\right)} = I_{X;W}\left(x_k; w_j\right). \end{aligned} \tag{1.2}$$

Property 1.3 Let U be a random variable with possible values $\{u_1, u_2, \ldots, u_M\}$.

Example 1.2 illustrates Property 1.3.

Property 1.4 Let W and X be jointly distributed random variables.

Example 1.5 Here we wish to calculate the mutual information and the average mutual information for the probability assignments (with $M = 2$ and $N = 2$)

$$P_W(w_1) = P_W(w_2) = \tfrac{1}{2} \tag{1.3}$$

Example 1.6 **Boeringer and Werner [2004]** The source output is a ternary-valued random variable that takes on the values $\{u_1, u_2, u_3\}$ with probabilities $P(u_1) = 0.7, P(u_2) = 0.15 = P(u_3)$.

Theorem 1.7 **(Source Coding Theorem)..** *For a DMS with entropy $H(U)$, the minimum average codeword length per source letter $(\bar{n})$ for any code is lower bounded by $H(U)$, that is, $\bar{n} \geq H(U)$, and further, $\bar{n}$ can be made as close to $H(U)$ as desired for some suitably chosen code.*

Theorem 1.8 **(Channel Coding Theorem Liu [2005])..** *Given a DMS with entropy H bits/source letter and a DMC with capacity C bits/source letter, if $H \leq C$, the source output can be encoded for transmission over the channel with an arbitrarily small bit error probability. Further, if $H > C$, the bit error probability is bounded away from 0.*

(a) Points in $\mathbb{R}^2$, subdivided by a single linear classifier. One simple way of understanding linear classifiers is as a line (or hyper-plane, in higher dimensions) which splits space into two regions. In this example, points above the line are mapped to class 1 (red); those below, to class 0 (blue).

(b) Points in $\mathbb{R}^2$, subdivided by a combination of four linear classifiers. Each classifier maps *all* points to class 0 or 1, and an additional linear classifier is used to combine the four. This hierarchical model is strictly more expressive than any linear classifier by itself.

Figure 1.4: Simple elements can be combined to express more complex relationships. This is one basic tenet of deep neural networks.

Proof. This result can be proved in several ways, including calculus of variations Wang et al [2005] inequality; however, an alternative method is used here. □

Table 1.1: Timer0 Compare Output Mode, non-PWM Mode

COM0x1-0	Description
00	Normal port operation
01	Toggle on Compare Match
10	Clear on Compare Match
11	Set on Compare Match

COM0x1-0	Description
00	Normal port operation
01	Toggle on Compare Match
10	Clear on Compare Match
11	Set on Compare Match

SUMMARY

In this chapter we have discussed very briefly some of the salient results from information theory and rate distortion theory and have indicated how these results can be used to bound communication system performance.

PROBLEMS

1.1. A random variable U has a sample space consisting of the set of all possible binary sequences of length N, denoted $\{u_j, j = 1, 2, \ldots, 2^N\}$.

1.2. Given a random variable U with the alphabet $\{u_1, u_2, u_3, u_4\}$ and probability assignments $P(u_1) = 0.8$, $P(u_2) = 0.1$, $P(u_3) = 0.05$, $P(u_4) = 0.05$, calculate the entropy of U. Compare your result to a random variable with equally likely values.

CHAPTER 2

Jordan Canonical Form

2.1 THE DIAGONALIZABLE CASE

Although, for simplicity, most of our examples will be over the real numbers (and indeed over the rational numbers), we will consider that *all of our vectors and matrices are defined over the complex numbers* $\mathbb{C}$. It is only with this assumption that the theory of Jordan Canonical Form (JCF) works completely. See Remark 2.3 for the key reason why.

Definition 2.1 If $v \neq 0$ is a vector such that, for some λ,

$$Av = \lambda v,$$

then v is an *eigenvector* of A associated to the *eigenvalue* λ.

Example 2.2 Let A be the matrix $A =$ Then, as you can check, if $v_1 =$ then $Av_1 = 3v_1$, so v_1 is an eigenvector of A with associated eigenvalue 3, and if $v_2 =$ then $Av_2 = -2v_2$, so v_2 is an eigenvector of A with associated eigenvalue -2.

Remark 2.3 This is the customary definition of the characteristic polynomial. But note that, if A is an n-by-n matrix, then the matrix $\lambda I - A$ is obtained from the matrix $A - \lambda I$ by multiplying each of its n rows by -1, and hence In practice, it is most convenient to work with $A - \lambda I$ in finding eigenvectors—this minimizes arithmetic—and when we come to find chains of generalized eigenvectors in Section 1.2, it is (almost) essential to use $A - \lambda I$, as using $\lambda I - A$ would introduce lots of spurious minus signs.

CHAPTER 3

An Algorithm

The navigation oriented heuristic ($h3$) considers the site topology. Accesses to cached pages are not recorded in the Web log due to the browser or proxy cache. Therefore, references to those pages are missed. The missing references in the log file can be found using a set of assumptions. The referrer field of the Web log or the Web site structure can be used to infer cached pages. If a requested Web page p_i is not reachable from previously visited pages in a session, then a new session is constructed starting with page p_i.

Algorithm 3.1 Construction of user sessions from Web server logs using $h1$ heuristic

Input : Web server logs
Output : set of user sessions $\mathcal{S} = \{s_1, ..., s_M\}$

$\mathcal{S} = \{\emptyset\}$
Order all Web logs by user IDs (u_k) and time increasingly
for all user ID u_k **do**
  Create a new user session in $\mathcal{S}$ for user u_k
  for i=1 to the number of records of this u_k **do**
    if $t_{i+1} - t_i < \triangle t$ **then**
      insert this record into user session
    else
      Create a new user session in $\mathcal{S}$ for user u_k
    end if
  end for
end for

CHAPTER 4

Shaded Areas

M&C provides a macro called `mcframe` that allows us to create a frame around some text which automatically breaks its contents across pages.

```
\begin{mcframe}[width](text-width)<background-color>

...
\end{mcframe}
```

The optional arguments are:

- `width` defaults to the width of the page
- `text-width` defaults to an indent of 10pt in each margin
- `background-color` defaults to the `bluetwenty` color of the M&C books.

Here are a few examples using some text from the book *General Game Playing* by M. Genesereth and M. Thielscher that was published by M&C in March, 2014:

`\begin{mcframe}`

Games of strategy, such as Chess, couple intellectual activity with competition. We can exercise and improve our intellectual skills by playing such games. The competition adds excitement and allows us to compare our skills to those of others. The same motivation accounts for interest in Computer Game Playing as a testbed for Artificial Intelligence. Programs that think better should be able to win more games, and so we can use competitions as an evaluation technique for intelligent systems.

Unfortunately, building programs to play specific games has limited value in AI. (1) To begin with, specialized game players are very narrow. They can be good at one game but not another. Deep Blue may have beaten the world Chess champion, but it has no clue how to play checkers. (2) A second problem with specialized game playing systems is that they do only part of the work. Most of the interesting analysis and design is done in advance by their programmers. The systems themselves might as well be tele-operated.

All is not lost. The idea of game playing can be used to good effect to inspire and evaluate good work in Artificial Intelligence, but it requires moving more of the design work to the computer itself. This can be done by focussing attention on General Game Playing.

```
\end{mcframe}
```

```
\begin{mcframe}[.7\textwidth](.6\textwidth)
```

Games of strategy, such as Chess, couple intellectual activity with competition. We can exercise and improve our intellectual skills by playing such games. The competition adds excitement and allows us to compare our skills to those of others. The same motivation accounts for interest in Computer Game Playing as a testbed for Artificial Intelligence. Programs that think better should be able to win more games, and so we can use competitions as an evaluation technique for intelligent systems.

Unfortunately, building programs to play specific games has limited value in AI. (1) To begin with, specialized game players are very narrow. They can be good at one game but not another. Deep Blue may have beaten the world Chess champion, but it has no clue how to play checkers. (2) A second problem with specialized game playing systems is that they do only part of the work. Most of the interesting analysis and design is done in advance by their programmers. The systems themselves might as well be tele-operated.

```
\end{mcframe}
```

```
\begin{mcframe}[.8\textwidth](.7\textwidth)<green>
```

Games of strategy, such as Chess, couple intellectual activity with competition. We can exercise and improve our intellectual skills by playing such games. The competition adds excitement and allows us to compare our skills to those of others. The same motivation accounts for interest in Computer Game Playing as a testbed for Artificial Intelligence. Programs that think better should be able to win more games, and so we can use competitions as an evaluation technique for intelligent systems.

Unfortunately, building programs to play specific games has limited value in AI. (1) To begin with, specialized game players are very narrow. They can be good at one game but not another. Deep Blue may have beaten the world Chess champion, but it has no clue how to play checkers. (2) A second problem with specialized game playing systems is that they do only part of the work. Most of the interesting analysis and design is done in advance by their programmers. The systems themselves might as well be tele-operated.

```
\end{mcframe}
```

CHAPTER 5

Math

5.1 BOLD MATH

Sometimes you have a math symbol in bold to indicate, for example, a vector or special functions. Use `\mathbf{x}` for letters and `\bm{\Phi}` for Greek letters.
...$\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ are the distribution functions governing the likelihood of $\mathcal{X}$ and $\mathbf{Z}$ ...

Some notes:

- use `\bm{\Phi}` to produce a bold Greek symbol Phi.
- use `\mathcal{X}` to display the caligraphic letter X
- use `\mathbf{Z}` to dispaly bold non-Greek symbols

If you plan to use Greek symbols to identify, for example, vectors, you could create a macro, like

```
\newcommand{\vect}[1]{\bm{#1}}
```

and then use

```
\ldots the vectors are $\vect{\Phi}$ and $\vect{\Theta}$ \ldots
```

to produce
...the vectors are $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ ...

The idea is to have macro names for different constructs. If you chose to use a different representation for vectors, you change the macro and rerun your manuscript.

5.2 MATH VARIABLES

When we have a math variable with more than one character, like *abcd*, LaTeX treats each character as a separate variable. With Caslon fonts, that we use to produce your book, that space is noticeable.

We eliminat the extra space using `\mathit`:

```
\mathit{abcd} = 3.
```

$$abcd = 3.$$

If we use the variable *abcd* often, as in

```
\mathit{abcd} = \mathit{abcd}^2 - \sqrt{\mathit{abcd}} + \mathit{abcd},
```

$$\mathit{abcd} = \mathit{abcd}^2 - \sqrt{\mathit{abcd}} + \mathit{abcd},$$

we can create a separate macro for the variable:

```
\newcommand{\abcd}{\mathit{abcd}}
```

Then we can use `\abcd` in the expressions:

```
\abcd = \abcd^2 - \sqrt{\abcd} + \abcd,
```

$$\mathit{abcd} = \mathit{abcd}^2 - \sqrt{\mathit{abcd}} + \mathit{abcd},$$

If there is already a predefined command with the same name, we use instead something like `Abcd`:

```
\newcommand{\Abcd}{\mathit{abcd}}
```

Then we can use `\Abcd` in the expressions:

```
\Abcd = \Abcd^2 - \sqrt{\Abcd} + \Abcd.
```

$$\mathit{abcd} = \mathit{abcd}^2 - \sqrt{\mathit{abcd}} + \mathit{abcd}.$$

CHAPTER 6

Citations with Author-Year

An alternative to numerical LaTeX citations makes reference to published works by citing the author's name and year of publication. The entries in the bibliographic listing are not numbered. The citation itself may be either parenthetical like [Carroll et al., 1998] or textual as shown by Carroll et al. [1998].

The `natbib` package by Patrick W. Daly is the most universal package for Author-Year citation and it is compatible with `\bibitem` syntaxes and with the `.bst` files of author-year packages such as `apalike`, `chicago`, and `harvard`.

If you create a bibliography by hand, please see the file `biblio.tex` in the `latex` directory of `bib-AY.zip`. The `biblio.tex` files shows the syntax that `natbib.sty` requires and notice how we use that package in `book.tex` and in `ch01.tex`. You can find the details about `natbib` and Author-Year bibliographies in Kopka and Daly [2004, pages 218-221].

Note that we have bibliographic references like `carroll98`, `KD:2004`, and `weber-97`. Pick a style that works for you and stay with it, then it will be easier for yourself and others to maintain and modify your bibliography and citations.

If you use `.bib` files then you should know about `BibTeX`, `.bst` files, and you should know how to create the `.bbl` files as well. (If you use `.bib` files but you do not know the rest, you should consult a local expert before continuing with `.bib` or create your bibliography by hand.) If you know the steps we mentioned above and you wish to create an Author-Year bibliography and citations, then you are in luck: there are three `.bst` that are provided with `natbib`. They are: `plainnat`, `unsrtnat`, and `abbrvnat`.

Once you create your `.bbl` file(s), please submit them along with your manuscript files, style files, and (preferably) `.eps` figures with `BoundingBoxes` to *Morgan and Claypool Publishers*.

You can get a free copy of the `natbib` package from CTAN at

```
http://www.ctan.org/tex-archive/macros/latex/contrib/natbib/
```

and download the file `natbib.zip`.

Here are some more samples using the citations that appear in the file `biblio.tex` that we built by hand.

The next paragraph uses text citations using `\citet` (cite as text):
There are several textbooks that give a general introduction to dependency grammar but most of them in other languages than English, for example, Tarvainen [1982] and Weber [1997].

The last paragraph uses text citations with parentheses using `\citetp`:
Tesnière's seminal work was published posthumously as [Tesnière , 1959]. Other influential theories in the dependency grammar tradition include Functional Generative Description [Sgall et al., 1986].

Please see the bibliography on the next page and the file `biblio.tex`.

Bibliography

D. Boeringer and D. Werner, "Particle swarm optimization versus genetic algorithms for phased array synthesis," *IEEE Trans. Antennas Propagat.*, vol. 52, no. 3, pp. 771–779, 2004.

Carroll, John, Briscoe, Edward, and Sanfilippo, Antonio Supertagging: An approach to almost parsing, *Computational Linguistics* 25(2):237–267.

S. Ho, S. Yang, G. Ni, E. W. C. Lo, and H. C. Wong, "A particle swarm optimization-based method for multiobjective design optimizations," *IEEE Trans. Magn.*, vol. 41, no. 5, pp. 1756–1759, 2005.

Kopka, Helmut and Daly, Patrick W. *Guide to LaTeX*, fourth edition, Addison Wesley.

F. S. Levin, *An Introduction to Quantum Theory*, Cambridge: Cambridge University Press, 2002.

W-C Liu, "A design of a multiband CPW-fed monopole antenna using a particle swarm optimization approach," *IEEE Trans. Antennas Propagat.*, vol. 53, no. 10, pp. 3273–3279, 2005.

Sgall, Petr, Hajičová, Eva, and Panevová, Jarmila. *The Meaning of the Sentence in Its Pragmatic Aspects*, Reidel.

Tarvainen, Kalevi. *Einführung in die Dependenzgrammatik*, Niemeyer.

Tesnière, Lucien. *Éléments de syntaxe structurale*, Editions Klincksieck.

Weber, H. J. *Dependenzgrammatik. Ein interaktives Arbeitsbuch*, Günter Narr.

W. Wang, Y. Lu, J. S. Fu, and Y. Z. Xiong, "Particle swarm optimization and finite-element based approach for microwave filter design," *IEEE Trans. Magn.*, vol. 41, no. 5, pp. 1800–1803, 2005.

Author's Biography

FIRSTNAME LASTNAME

FirstName LastName ...