1

The cool title for an interesting book about a wonderful topic

Giacomo Marciani

August 26, 2016

[1] Department of Civil Engineering and Computer Science Engineering
University of Rome Tor Vergata
Rome, Italy

Abstract

In quantum computing, where algorithms exist that can solve computational problems more efficiently than any known classical algorithms, the elimination of errors that result from external disturbances or from imperfect gates has become the “holy grail,” and a worldwide quest for a large scale fault-tolerant and computationally superior quantum computer is currently taking place. Optimists rely on the premise that, under a certain threshold of errors, an arbitrary long fault-tolerant quantum computation can be achieved with only moderate (i.e., at most polynomial) overhead in computational cost. Pessimists, on the other hand, object that there are in principle (as opposed to merely technological) reasons why such machines are still inexistent, and that no matter what gadgets are used, large scale quantum computers will never be computationally superior to classical ones. Lacking a complete empirical characterization of quantum noise, the debate on the physical possibility of such machines invites philosophical scrutiny. Making this debate more precise by suggesting a novel statistical mechanical perspective thereof is the goal of this project.

Keywords

computational complexity, decoherence, error-correction, fault-tolerance, Landauer’s Principle, Maxwell’s Demon, quantum computing, statistical mechanics, thermodynamics

Contents

Preface

Which labor market institutions worked better in containing job losses during the Great Recession of 2008–2009? Is it good for employment to increase the progressiveness of taxation? Does it make sense to contrast "active" and "passive" labor market policies? Who actually gains and who loses from employment protection legislation? Why are minimum wages generally diversified by age? Is it better to have decentralized or centralized bargaining systems in monetary unions? Should migrants have access to welfare benefits? Should governments regulate working hours? And can equal opportunity legislation reduce discrimination against women or minority groups in the labor market?

Current labor economics textbooks neglect these relevant policy issues. In spite of significant progress in analyzing the costs and benefits of labor market institutions, these textbooks have a setup that relegates institutions to the last paragraph of chapters or to a final institutional chapter. Typically a book begins by characterizing labor supply (including human capital theory), labor demand, and the competitive equilibrium at the intersection of the two curves; it subsequently addresses such topics as wage formation and unions, compensating wage differentials, and unemployment without a proper institutional framework. There is little information concerning labor market institutions and labor market policies. Usually labor market policies are mentioned only every now and then, and labor market institutions are often not treated in a systematic way. When attention is given to these institutions, reference is generally made to the U.S. institutional landscape and to competitive labor markets in which, by definition, any type of policy measure is distortionary.

Your Name
March 2014

1 Classification and Prediction

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. Although, everyone has an intuitive feeling for what information is, it is difficult to attach a meaningful quantitative definition to the term. In the context of communication systems, Claude Shannon was able to do exactly this, and as a result, opened up an entirely new view of communication systems analysis and design [**?**, p. 123]. The principal contribution of Shannon's information theory to date has been to allow communication theorists to establish absolute bounds on communication systems performance that cannot be exceeded no matter how ingeniously designed or complex our communication systems are. Fundamental physical limitations on communication systems performance is another topic that has been largely ignored in the preceding chapters, but it is a subject of exceptional practical importance. For example, for any of the numerous communication systems developed thus far in the book, we could decide to design a new system that would outperform the accepted standard for a particular application. The first question that we should ask is: how close is the present system to achieving theoretically optimum performance? If the existing communication system operates at or near the fundamental physical limit on performance, our task may be difficult or impossible. However, if the existing system is far away from the absolute performance bound, this might be an area for fruitful work.

1.1 What is Classification? What is Prediction?

Of course, in specifying the particular communication system under investigation, we must know the important physical parameters, such as transmitted power, bandwidth, type(s) of noise present, and so on, and information theory allows these constraints to be incorporated. However, information theory does not provide a way for communication system complexity to be explicitly included. Although, this is something of a drawback, information theory itself provides a way around this difficulty, since it is generally true that as we approach the fundamental limit on the performance of a communication system, the system complexity increases, sometimes quite drastically. Therefore, for a simple communication system operating far from its performance bound, we may be able to improve the performance with a relatively modest increase in complexity. On the other hand, if we have a rather complicated communication system operating near its fundamental limit, any performance improvement may be possible only with an extremely complicated system.

In this chapter we are concerned with the rather general block diagram shown in Figure 1.1. Most of the early work by Shannon and others ignored the source encoder/decoder blocks and concentrated on bounding the performance of the channel encoder/decoder pair. Subsequently, the source encoder/decoder blocks have attracted much research attention. In this chapter we consider both topics and expose the reader to the nomenclature used in the information theory literature. Quantitative definitions of information are presented in Sec. 1.3 that lay the foundation for the remaining sections. In Secs. 1.3 and 1.3 we present the fundamental source and channel coding theorems, give some examples, and state the implications of these theorems. Section 1.3 contains a brief development of rate distortion theory, which is the mathematical basis for data compression. A few applications of the theory in this chapter are presented in Sec. 1.3, and a technique for variable-length source coding is given in Sec. 1.3.

$$x = a + b - \sqrt{q} \tag{1.1}$$

Of course, in specifying the particular communication system under investigation, we must know the important physical parameters, such as transmitted power, bandwidth, type(s) of noise present, and so on, and information theory allows these constraints to be incorporated. However, information theory does not provide a way for communication system complexity to be explicitly included. Although, this is something of a drawback, information theory itself provides a way around this difficulty, since it is generally true that as we approach the fundamental limit on the performance of a communication system, the system complexity increases, sometimes quite drastically. Therefore, for a simple communication system operating far from its performance bound, we may be able to improve the performance with a relatively modest increase in complexity. On the other hand, if we have a rather complicated communication system operating near its fundamental limit, any performance improvement may be possible only with an extremely complicated system.

1.1.1 And Yet More of the Same

In this chapter we are concerned with the rather general block diagram shown in Figure 1.1. Most of the early work by Shannon and others ignored the source encoder/decoder blocks and concentrated on bounding the performance of the channel encoder/decoder pair. Subsequently, the source encoder/decoder blocks have attracted much research attention. In this chapter we consider both topics and expose the reader to the nomenclature used in the information theory literature. Quantitative definitions of information are presented in Sec. 1.3 that lay the foundation for the remaining sections. In Secs. 1.3 and 1.3 we present the fundamental source and channel coding theorems, give some examples, and state the implications of these theorems. Section 1.3 contains a brief development of rate distortion theory, which is the mathematical basis for data compression. A few applications of the theory in this chapter are presented in Sec. 1.3, and a technique for variable-length source coding is given in Sec. 1.3.

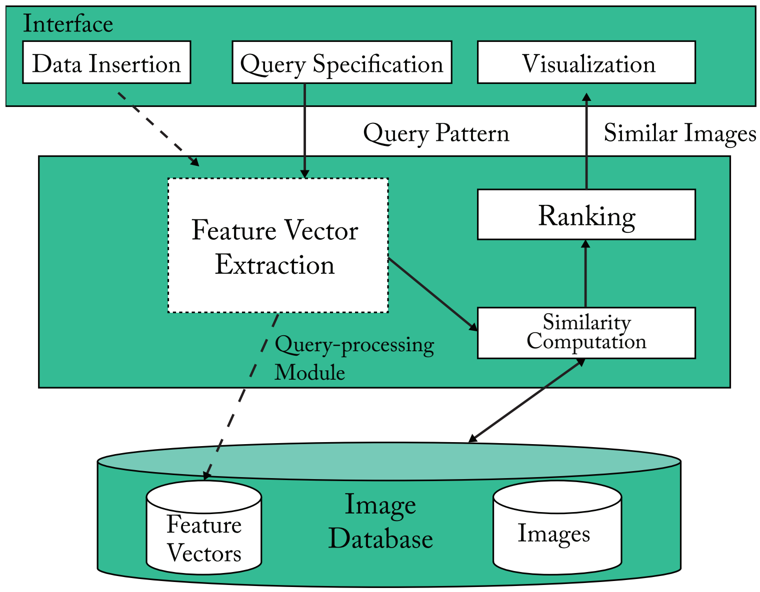

Figure 1.1 This is a caption for this figure. It is fairly long so we can make sure it looks good when occupying more than one line. Here is one more sentence to make it longer.

- This is a bullet list with a short item.
- And another item that is much longer so that we can make sure it is formatted correctly and so forth and so on.
- And a final short item.

In this chapter we are concerned with the rather general block diagram shown in Figure 1.1. Most of the early work by Shannon and others ignored the source encoder/decoder blocks and concentrated on bounding the performance of the channel encoder/decoder pair. Subsequently, the source encoder/decoder blocks have attracted much research attention. In this chapter we consider both topics and expose the reader to the nomenclature used in the information theory literature.

$$a = b + c \tag{1.2}$$
$$x = \frac{1}{2}a \tag{1.3}$$

Quantitative definitions of information are presented in Sec. 1.3 that lay the foundation for the remaining sections. In Secs. 1.3 and 1.3 we present the fundamental source and channel coding theorems, give some examples, and state the implications of these theorems. Section 1.1 contains a brief development of rate distortion theory, which is the mathematical

Table 1.1 This is a little table here.

Foo	Bar	Zoo	Snork	Quux
0	0	1	2	4
0	0	3	2	4
3	3	1	0	1
3	1	4	2	1
3	4	4	2	2

basis for data compression. A few applications of the theory in this chapter are presented in Sec. 1.3, and a technique for variable-length source coding is given in Sec. 1.3.

Only the binary Huffman procedure has been described here, but nonbinary codes can be designed using the Huffman method. The details are somewhat more complicated and nonbinary codes are less commonly encountered than binary ones, so further discussion is left to the problems and the literature.

1.2 Case Study

In this section, we exemplify how the 5S extensions for content-based image retrieval can be explored to define an image search service in the context of the CTRnet project. The Crisis, Tragedy, and Recovery Network (CTRnet) [**?**] objectives include to develop better approaches toward making technology useful for archiving information about such events, and to support analysis of rescue, relief, and recovery, from a digital library perspective. CTRnet has several modules, including crawling, filtering, a Facebook application, visualization, metadata search, and Content-Based Image Retrieval (CBIR).

The CBIR module builds upon the EVA tool for evaluating image descriptors for content-based image retrieval [**?**]. Eva integrates the most common stages of an image retrieval process and provides functionalities to facilitate the comparison of image descriptors in the context of content-based image retrieval.

In this case study, we consider the scenario in which a user is interested in finding images in the CTRnet collection that are similar to a particular photo provided as example. The objective is to identify images that could be used in a report on damages caused by an earthquake. In this example, the query specification q would be a tuple $q = (H_q, Contents_q, P_q)$, where q is an image (see Figure 1.1) Thus, $q = ((V_q, E_q), L_q, F_q), Contents_q, P_q)$, where $V_q = v_1$; $E_q = \emptyset$; $L_q = 'Cathedral\_P\_A\_P.jpg'$; $F_q : V_q \cup E_q \rightarrow L_q$, $Contents_q$ is the stream of the query image; and $P_q : V_q \rightarrow Contents_q$.

In this chapter we are concerned with the rather general block diagram shown in Figure 1.1. Most of the early work by Shannon and others ignored the source encoder/decoder blocks and concentrated on bounding the performance of the channel encoder/decoder pair. Subsequently, the source encoder/decoder blocks have attracted much research attention. In this chapter we consider both topics and expose the reader to the nomenclature used in the information theory literature.

1.3 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 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

$$I_U(u_j) = \log \frac{1}{P_U(u_j)} = -\log P_U(u_j) \ . \tag{1.4}$$

The quantity $I_U(u_j)$ is a measure of the information contained in the event $u = u_j$. Note that the base of the logarithm in Eq. (1.4) is unspecified. It is common to use base e, in which case $I_U(\cdot)$ is in natural units (nats), or base 2, in which case $I_U(\cdot)$ is in binary units (bits). Either base is acceptable since the difference in the two bases is just a scaling operation. We will use base 2 in all of our work, and hence $I_U(\cdot)$ and related quantities will be in bits. The average or expected value of the self-information is called the *entropy,* also discrete entropy or absolute entropy, and is given by

$$H(U) = -\sum_{j=1}^{M} P_U(u_j) \log P_U(u_j) \ . \tag{1.5}$$

The following example illustrates the calculation of entropy and how it is affected by probability assignments.

1.4 Exercises and Projects

1. How might CBIR be applied so teachers with a computer and connected camera can be reminded of the names of students in their class?
2. Consider the two colourful images (Image A and Image B) showed below, represented in the RGB color space. Suppose that the intensity values of each pixel in all bands (R, G, and B) are the same. Furthermore, each (R, G, B) triplet is represented by a single intensity value. For example the triplet $(R, G, B) = (2, 2, 2)$ is represented by the intensity value 2.

   Suppose also that the colour space was quantized in five colors with intensity values 0, 1, 2, 3, and 4.
3. Compute the L_1 between the *Color Histograms* (5 bins) of the two images. The L_1 distance between two color histograms H_A and H_B is computed as follows: $L_1(H_A, H_B) = \sum_{i=1}^{K} |H_A[i] - H_B[i]|$, where K is the size of both histograms (5, in the case).

4. By considering both the feature vector extraction function and the distance function defined of the descriptor *Color Coherence Vector – CCV* [**?**], compute the distance $\delta_{CCV}(A, B)$ between the two images.
5. Consider the existence of two classes (*class 1* and *class 2*) composed of five images each. Consider the existence of three different descriptors (*descriptor 1*, *descriptor 2*, and *descriptor 3*), whose feature vector extraction functions extract vectors belonging to the R^2 space. Table 1.1 shows the coordinate of each image of each class, considering the three descriptors.

A This Is an Appendix

Of course, in specifying the particular communication system under investigation, we must know the important physical parameters, such as transmitted power, bandwidth, type(s) of noise present, and so on, and information theory allows these constraints to be incorporated. However, information theory does not provide a way for communication system complexity to be explicitly included. Although, this is something of a drawback, information theory itself provides a way around this difficulty, since it is generally true that as we approach the fundamental limit on the performance of a communication system, the system complexity increases, sometimes quite drastically. Therefore, for a simple communication system operating far from its performance bound, we may be able to improve the performance with a relatively modest increase in complexity. On the other hand, if we have a rather complicated communication system operating near its fundamental limit, any performance improvement may be possible only with an extremely complicated system.

In this chapter we are concerned with the rather general block diagram shown in Figure 1.1. Most of the early work by Shannon and others ignored the source encoder/decoder blocks and concentrated on bounding the performance of the channel encoder/decoder pair. Subsequently, the source encoder/decoder blocks have attracted much research attention. In this chapter we consider both topics and expose the reader to the nomenclature used in the information theory literature. Quantitative definitions of information are presented in Sec. 1.3 that lay the foundation for the remaining sections. In Secs. 1.3 and 1.1 we present the fundamental source and channel coding theorems, give some examples, and state the implications of these theorems. Section 1.3 contains a brief development of rate distortion theory, which is the mathematical basis for data compression. A few applications of the theory in this chapter are presented in Sec. 1.3, and a technique for variable-length source coding is given in Sec. 1.3.

In this chapter we are concerned with the rather general block diagram shown in Figure 1.1. Most of the early work by Shannon and others ignored the source encoder/decoder blocks and concentrated on bounding the performance of the channel encoder/decoder pair. Subsequently, the source encoder/decoder blocks have attracted much research attention. In this chapter we consider both topics and expose the reader to the nomenclature used in the information theory literature. Quantitative definitions of information are presented in Sec. 1.3 that lay the foundation for the remaining sections. In Secs. 1.3 and 1.1 we present the fundamental source and channel coding theorems, give some examples, and state the implications of these theorems.

A.1 What is Classification? What is Prediction?

Of course, in specifying the particular communication system under investigation, we must know the important physical parameters, such as transmitted power, bandwidth, type(s) of noise present, and so on, and information theory allows these constraints to be incorporated. However, information theory does not provide a way for communication system complexity to be explicitly included. Although, this is something of a drawback, information theory itself provides a way around this difficulty, since it is generally true that as we approach the fundamental limit on the performance of a communication system, the system complexity increases, sometimes quite drastically. Therefore, for a simple communication system operating far from its performance bound, we may be able to improve the performance with a relatively modest increase in complexity. On the other hand, if we have a rather complicated communication system operating near its fundamental limit, any performance improvement may be possible only with an extremely complicated system.

1. This is a number list with a short item.
2. And another item that is much longer so that we can make sure it is formatted correctly and so forth and so on.
3. And a final short item.

Bibliography

Author’s Biography

Giacomo Marciani

Giacomo Marciani began life as a small child ...