Quantum Information and Computation, Vol. 2, No. 6 (2002) 515–516

BOOK REVIEW

on

A First Course in Information Theory

by Raymond W. Yeung

Kluwer Academic/Plenum Publishers, 2002

Hardcover (including CD-ROM) $125.00 (412 pages) ISBN: 0-306-46791-7

Raymond W. Yeung's *First Course in Information Theory* is an up–to–date account of the subject of modern information theory as seen by one of its most eminent researchers. More than fifty years after its conception by Claude Shannon, information theory is a vast and still highly active area which is probably beyond being captured comprehensively in one textbook. Therefore every author who ventures on the enterprise of writing such a textbook may be granted the license to select some topics according to personal taste and expertise. While early presentations of the subject by the M.I.T. school, for example Fano's *Transmission of Information* (1961) and Jelinek's *Probabilistic Information Theory* (1968), could still aim at an overview that left out little, the defining textbook of the 1980s, Csiszár and Kőrner's *Information Theory* (1980) concentrated on a narrower area of research. The emphasis in the book by Csiszár and Kőrner's was on combinatorial techniques, most famously the *method of types* and on the then thriving theory of multi–terminal coding problems, and all theory was set in the domain of memoryless systems. When Cover and Thomas in 1991 wrote their classic *Elements of Information Theory*, they must have felt that one can no longer go to the extremes of technical sophistication (the treatise of Csiszár and Kőrner can be difficult to read, even today). Instead the breadth of the field and the need to expose fundamental principles demanded an approach that would leave out some of the subtler points available in the original literature. As an introduction which both covers the broad range of applications of information science and the fundamental concepts, "Cover–Thomas" is, in my view, still unchallenged, as is "Csiszár–Kőrner" for mathematical sophistication.

Yet, there is a new textbook to be added to the ranks of these veterans and champions, and the reader of this review is likely to ask why one should read this particular one. I have two answers to this question which are linked to each other. My first observation is that the book really is less a first course than a "Menu" composed by an accomplished chef: it offers the reader a highly personal tour of current information theory. Indeed, the title may be misleading, as the first half of the book (chapters 1 through 10) are what would be in an introductory course at best. This part covers all the time–honored "musts" of the subject, starting with an abstract discussion of entropy and mutual information, proceeding through source coding and channel capacity (including feedback and joint source–channel coding) to rate–distortion theory. It is to be noted that the *discrete* point of view is embraced throughout (like in the book of Csiszár and Kőrner), excluding the treatment of continuous systems like the Gaussian channel. As a refinement of the almost–equipartition property (AEP) approach of Cover and Thomas, yet not going as far as Csiszár and Kőrner in its development, the notion of "strong typicality" is introduced, and discussed by re–interpreting the basic entropy inequalities as combinatorial counting inequalities.

Even in this first part the chapters 6 and 7 (on so–called "I–measures") are already preparing the ground for the theory that is to take up the second part: chapters 12 through 16 expand in detail on the theory developed by the author and his collaborators on generic inequalities which entropies of a family of random variables have to satisfy. These are divided into "Shannon–type" inequalities, which are all those that can be derived from the positivity of conditional entropy and conditional mutual information, i.e. $H(X|Y) \geq 0$ and $I(X;Y|Z) \geq 0$ for random variables X, Y, Z, and the remaining "non–Shannon-type" inequalities, which were first exhibited by Yeung *et al.* For example, Theorem 14.7 in the book states that for random variables U, V, X, Y, we have $2I(X;Y) \leq I(U;V) + I(U;XY) + 3I(X;Y|U) + I(X;Y|V)$. Astonishingly (Theorem 14.11), this inequality does not follow from the basic Shannon–type inequalities. Chapter 15 shows that the theory of these newly developed inequalities is not the idle play of researchers enamored with the entropy functional. This chapter is on the theory of multi–source network coding which provides a nontrivial application of the new inequalities.

This, in brief, is the content of this finely written book. My first answer to the question above "why read?" is a recommendation, as I find Prof. Yeung's selection very tasteful and interesting, apart from being very readable. In addition, the first half of the book also is a thorough and consistent introduction to the fundamentals of information theory, while the later chapters feature several advanced and more recent topics. My second answer is a recommendation especially aimed at the quantum information theorist. Shannon entropy is far from being understood mathematically: the latest advances of Yeung and his followers should be viewed not as the final resolution of a problem but as the opening up of new vistas. Our mathematical understanding of the quantum version of the Shannon entropy, the von Neumann entropy, is even less extensive, yet its properties play a crucial role in the emerging theory of quantum information. A most characteristic question, to my mind, is that of the additivity of the entanglement of formation and of the classical capacity of a quantum channel. Note that the von Neumann entropy (thanks to strong subadditivity) satisfies all but one of the fundamental inequalities, the exception being that conditional entropies can be negative. Which of the non–Shannon–type inequalities remain valid in the quantum domain is unknown. It is to be expected that further, not yet known, inequalities for the von Neumann entropy will greatly advance quantum information science, and Yeung's theory might well serve as a template for their discovery. At least, it can be hoped that it may inspire the reader of *A First Course in Information Theory* to her or his own reflections on entropy.

Andreas Winter (Email: `winter@cs.bris.ac.uk`)
Department of Computer Science
University of Bristol
Merchant Venturers Building, Woodland Road
Bristol BS8 1UB, United Kingdom