Book Review

A First Course in Information Theory R. W. Yeung (New York: Kluwer/Plenum, 2002. 412 + xxi pp.) *Reviewed by Imre Csiszár, Fellow, IEEE*

The title of this book is rather misleading, for less than half of it is devoted to an introductory treatment of basic concepts of information theory. The rest is advanced material, not surprisingly pertaining to the author's research interests. These topics, here appearing first in a book, represent a very valuable contribution to the literature. Compared to them, the introductory part appears of lesser value (this review does not follow the author's classification of Chapters 1–13 as the basic and Chapters 14–16 as the advanced topics). The book comes with a CD containing the Information Theoretic Inequality Prover (ITIP) software, developed by the author and Y. O. Yan.

The introductory material in Chapters 1–5 and 8–10 covers information measures, lossless source coding, weak and strong typicality, capacity of discrete memoryless channels (DMCs) including those with feedback, rate-distortion theory, and the Blahut–Arimoto algorithm. Multiuser theory is not entered here, except for one problem in Chapter 9 related to the Slepian–Wolf theorem. Efforts are made at a mathematically careful treatment of subtle points, for example, the role of strict positivity is properly emphasized when dealing with conditional independence. The problem of zeros is overlooked, however, in Definition 5.1 of a strongly typical set, where the exclusion of sequences containing some letter x with $p(x) = 0$ is missing. Types and type classes are not treated, though some messy proofs such as of Theorems 5.2, 5.3, 5.9 would be much more transparent using types; the concept would come handy also in Section 16.3 presenting a major result of the book. A commendable feature is the visualization of information-theoretic inequalities via properties of typical sets (Section 5.4). An attempt is made at a definition of a (DMC) that covers both the nonfeedback and feedback cases; unfortunately, Definition 8.2 appears obscure, and the reliance upon "dependency graphs" in Sections 8.3 and 8.6 suffers from the lack of a formal definition of how such graphs give rise to probability distributions.

The novel features of the book include the treatment of I-measures and information diagrams, in Chapters 6 and 7. The I-measure corresponding to an n-tuple of random variables (RVs) is a signed measure on a set of size $2^n - 1$, such that information-theoretic identities correspond to set-theoretic ones for the I-measure, often admitting graphical visualization. Part of the theory of I-measures is elementary and may well belong to a first course, but another part is highly advanced. The latter includes a characterization of Markov random fields (Theorem 7.25), with the important consequence that the I-measure corresponding to a Markov chain is always a true measure, that is, nonnegative.

Chapters 12, 13, 14, and 16 are centered around a recent but conceptually fundamental problem about information measures. Associate with each n-tuple of RVs $X_1, \ldots, X_n$ a $2^n - 1$-dimensional vector h with components indexed by the nonempty sets $\alpha \subseteq \{1, \ldots, n\}$, whose component h_α is the joint entropy of the RVs X_i, $i \in \alpha$. The problem is to determine the set of all such vectors h, denoted by Γ_n^*, or at least its closure $\overline{\Gamma_n^*}$. The mentioned chapters contain results pertinent to this—apparently very hard—problem. Determining $\overline{\Gamma_n^*}$ would amount to determining all information-theoretic inequalities that can be reduced to the nonnegativity of some linear combination of joint entropies as above. All such inequalities that have been known until recently are "Shannon type," that is, consequences of the nonnegativity of Shannon's information measures. Shannon-type inequalities are discussed in Chapter 13, in particular, they can be machine-proved by the ITIP software. Non-Shannon-type inequalities are treated in Chapter 14. Most remarkably, in Section 16.3, an equivalent characterization of the set $\overline{\Gamma_n^*}$ is found in terms of algebraic concepts, viz. finite groups and their subgroups.

Finally, two chapters (11 and 15) are devoted to "network coding," a new direction in multiuser information theory. Chapter 11 addresses lossless transmission of a source to several destinations via a network of noiseless channels with given capacities, when coding operations can be performed at each node. The maximum achievable source rate is shown to be as expected, but the proof is quite hard. The generalization of the problem to several sources, treated in Chapter 15, has no complete solution yet. Still, similarly looking inner and outer bounds are given to the region of achievable source rate tuples, when the zero error requirement is relaxed to small probability of error. These bounds, though not computable, are of major conceptual interest for they involve Γ_n^* and $\overline{\Gamma_n^*}$ as above, with suitable n.

The book is well written, although some proofs are (perhaps unavoidably) tedious. There are many examples, and problems at the end of each chapter. I did not find many errors, but systematically erroneous reference numbers to the bibliography are a nuisance. It is awkward that the examples are set in italic.

In summary, the introductory material in this book is suitable for a first course in information theory, but not preferable, e.g., to the Cover-Thomas book. The advanced material is much more valuable, and the reviewer unequivocally recommends reading it. Learning about these new research topics should be beneficial to all information theorists, as well as to computer scientists. This material could also be used for an advanced course for students with a mathematical orientation.

Manuscript received March 10, 2003.

The reviewer is with A. Renyi Institute of Mathematics, The Hungarian Academy of Sciences, H-1364 Budapest, Hungary.

Communicated by S. Verdú, Book Reviews Editor.
Digital Object Identifier 10.1109/TIT.2003.813510

0018-9448/03$17.00 © 2003 IEEE