Testi del Syllabus

Resp. Did.	**BONONI Alberto**	Matricola:	**004854**

Anno offerta:	**2015/2016**
Insegnamento:	**1005249 - INFORMATION THEORY**
Corso di studio:	**5052 - COMMUNICATION ENGINEERING - INGEGNERIA DELLE TELECOMUNICAZIONI**
Anno regolamento:	**2015**
CFU:	**6**
Settore:	**ING-INF/03**
Tipo Attività:	**B - Caratterizzante**
Anno corso:	**1**
Periodo:	**Primo Semestre**
Sede:	**PARMA**

Testi in italiano

Lingua insegnamento

Inglese

Contenuti

1.1 Definition of basic information theory quantities: entropy, mutual information
1.2 Basic theorems, relations and inequalities
1.3 Sufficient statistics
2.1 Asymptotic equipartition property (AEP) and typical set
2.2 Entropy rate
2.3 Discrete-time Markov chains (DTMC)
2.4 AEP theorem for stationary ergodic sources
3.1 Data Compression: definitions, examples,
3.2 Kraft inequality, Noiseless coding theorem
3.3 Huffman and Lempel-Ziv codes
4.1 Channel capacity: definitions, examples
4.2 Typical-sequence dcoding. Jointly typical set and its properties.
4.2 Proof Channel Coding Theorem
4.3 Joint source-channel coding theorem
5.1 Differential entropy: definition, examples
6.1 Mutual information for discrete X and continuous Y: examples
6.2 Capacity of additive Gaussian channel, Shannon Capacity formula
6.3 Parallel Gaussian channels
6.4 Capacity of discrete-time additive Gaussian channel with memory.
6.5 Capacity of continuous-time additive Gaussian channel with memory. Water pouring.

Testi di riferimento

[1] T. M. Cover, J. A. Thomas, "Elements of Information Theory". John Wiley and Sons, 1991.

Testi Secondari:

[2] R. Blahut, "Principles and Practice of Information Theory". Addison-Wesley, 1988.
[3] J. Cioffi, "Ch. 8: Fundamental Limits of Coding and Sequences", http://www.stanford.edu/~cioffi

Obiettivi formativi

L’obiettivo del corso è fornire allo studente la capacità di comprendere ed applicare le regole di base della teoria dell'informazione, e in particolare:
- il significato "fisico" delle principali grandezze entropia e informazione mutua e le loro inter-relazioni
- il concetto di sequenze tipiche e il loro ruolo nella compressione dei dati
- il concetto di capacità di canale

Le capacità di applicare le conoscenze sopra elencate risultano essere in particolare:
- progettare ed analizzare codici di compressione dei dati
- calcolare la capacità di un canale di trasmissione

Prerequisiti

Vedi testo inglese

Metodi didattici

Didattica frontale 42 ore. Esercitazioni 6 ore. Esercizi assegnati per casa.

Altre informazioni

Vedi testo inglese

Modalità di verifica dell'apprendimento

Esami solo orali. Dettagli nel testo inglese.

Programma esteso

Vedi testo inglese.

Testi in inglese

Lingua insegnamento

Course Language

English

Contenuti

1.1 Definition of basic information theory quantities: entropy, mutual information
1.2 Basic theorems, relations and inequalities
1.3 Sufficient statistics
2.1 Asymptotic equipartition property (AEP) and typical set
2.2 Entropy rate
2.3 Discrete-time Markov chains (DTMC)
2.4 AEP theorem for stationary ergodic sources
3.1 Data Compression: definitions, examples,
3.2 Kraft inequality, Noiseless coding theorem
3.3 Huffman and Lempel-Ziv codes
4.1 Channel capacity: definitions, examples
4.2 Typical-sequence dcoding. Jointly typical set and its properties.
4.2 Proof Channel Coding Theorem
4.3 Joint source-channel coding theorem
5.1 Differential entropy: definition, examples
6.1 Mutual information for discrete X and continuous Y: examples
6.2 Capacity of additive Gaussian channel, Shannon Capacity formula
6.3 Parallel Gaussian channels
6.4 Capacity of discrete-time additive Gaussian channel with memory.
6.5 Capacity of continuous-time additive Gaussian channel with memory. Water pouring.

Testi di riferimento

Textbook:

[1] T. M. Cover, J. A. Thomas, "Elements of Information Theory". John Wiley and Sons, 1991.

	Complementary Reading [2] R. Blahut, "Principles and Practice of Information Theory". Addison-Wesley, 1988. [3] J. Cioffi, "Ch. 8: Fundamental Limits of Coding and Sequences", http://www.stanford.edu/~cioffi
Obiettivi formativi	Course Objectives Objective of the course is to provide the student with the ability to understand and apply the basic rules of information theory, and in particular: - the physical meaning of the main information quantities, namely, entropy and mutual information and their inter-relationships - the concept of typical sequences and their role in data compression - the concept of channel capacity The abilities in applying the above-mentioned knowledge are in particular in the: - design and analysis of data compression codes - calculation of the capacity of a given transmission channel
Prerequisiti	Pre-requisistes A first course in basic probability theory
Metodi didattici	Teaching Methodology Classroom teaching, 42 hours. In-class problem solving, 6 hours. Homeworks assigned weekly.
Altre informazioni	Office Hours Monday 11:30-13:30 (Scientific Complex, Building 2, floor 2, Room 2/19T).
Modalità di verifica dell'apprendimento	Exams Oral only, to be scheduled on an individual basis. When ready, please contact the instructor by email alberto.bononi[AT]unipr.it by specifying the requested date. The exam consists of solving some proposed exercises and explaining theoretical details connected with them, for a total time of about 1 hour. You can bring your summary of important formulas in an A4 sheet to consult if you so wish. To get userid and password, please send an email to the instructor from your account nome@studenti.unipr.it.
Programma esteso	Syllabus (every class = 2 hours) CLASS 1: Intro: Course organization, objectives, textbooks, exam details. Sneaky preview of the course, motivations, applications. Assigned Reading of Ch.1 of textbook. Physical justification and definition of entropy. Examples of entropy calculation. Up to sec. 2.1. CLASS 2: Definition of joint and conditional entropy, eample 2.2.1. Relative entropy, mutual information and their relation. Chain rules for PMFs and entropy. CLASS 3: Relative conditional entropy, conditional mutual information, chain rules for D and I. Inequalities for D and I. max and min of H, H(X|Y)<=H(X) and generalizations. Convex functions. Jensen's inequality, examples. CLASS 4: first hour: logsum inequality, convexity of D, concavity of H. concavity of I in p(x) and convexity in p(y|x). Exercise: mixing increases entropy. second hour: Definition of Markov chain and first properties for 3 random variables (RV) X,Y,Z. Data processesng inequality. Counterexample. CLASS 5: first hour: sufficient statistics: definition in terms of mutual information.

Examples: number of successes in repeated trials; sample mean in estimation of common mean in a vector of independent gaussian RVs. Sufficient statistics and hypothesis testing: factorization theorem. Second hour: Fano inequality. Exercise 2.32.

CLASS 6:
Exercises 2.5, 2.4, 2.27, 2.30 (after brief introduction to the method of Lagrange multipliers), 2.21.

CLASS 7:
Ch 3 asymptotic equipartition property (AEP): introduction. Probability theory refresher: i.p. convergence, Chebychev inequality, Weak law of large numbers, AEP. Typical set and properties. Example with binary sequences.

CLASS 8:
first hour: relation among Typical set and high-probability sets. Theorem 3.3.1. second hour: Problem solving: Exercises 3.8, 3.9. Ch 4: Entropy rates: introduction. Definition of discrete-time stochastic process and stationarity.

CLASS 9:
first hour: introduction to discrete-time Markov chains (DTMC): transition matrix, updatge law (Chapman-Kolmogorov), stationary distribution. Two-state example: state diagram, evolution towards limit distribution. Evaluation with flux balancing. second hour: Entropy rates H and H'. H=H' for stationary processes. Statement of AEP theorem for stationary ergodic sources (Shannon/Breiman/McMillan). Explicit evaluation of H for DTMC. Examples.

CLASS 10:
Doubly-stochastic matrices and uniform steady-state distribution. Connections with entropy as defined in statistical thermodynamics: DTMC on microstates with doubly-stochastic transition matrix. Entropy increases towards steady-state distribution entropy. Example 4 (eq 4.50-4.52). Hidden Markov models (HMM): entropy rate.

CLASS 11:
Problem solving: Ex. 4.1 mixing increases entropy. Conditions for observable Y in a HMM sian una DTMC. examples where Y is not a DTMC. Point a. of Ex. 4.18 on Entropy Rate of stationary but not ergodic process.

CLASS 12:
Problem solving: first hour: points b, c of Ex. 4.18. second hour: Es 4.10 entropy rate of a second order markov process: study of hidden markov chain. Ex. 4.6.

CLASS 13:
Ch 5: Data compression. Examples of codes. Kraft inequality. Search of optimal codes with Lagrange multipliers method. Noiseless coding theorem.

CLASS 14:
Comments on first Shannon Theorem: when p not dyadic. Quasi-optimal Shannon Codes. Shannon super-codes are asymptotically optimal. Extra cost on minimal code length when using a PMF that differs from true PMF. McMillan Theorem: every uniquely decodable theorem satisfies Kraft inequality. Introduction to Huffman codes: examples 1, 2.

CLASS 15:
Huffman codes: example 3 (dummy symbols), Exercise 5.32, example 5.73 (set of different optimal codelengths). Competitive optimality of Shannon code. Proof of optimality of Huffman code.

CLASS 16:
first hour: Optimal compression of Markov sources. Description of Lempel-Ziv algorithm for universal compression. second hour: Channel capacity: introduction, definition of discrete memoryless channel (DMC), examples of capacity computation: ideal channel, noisy channel with disjoint outputs, noisy typewriter.

CLASS 17:
Capacity of bynary symmetric channel (BSC), binary erasure channel (BEC). Symmetric, weakly-symmetric and Gallager-symmetric channels. Convexity of C on convex set of input PMFs. Hints to numerical techniques to evaluate max I.

CLASS 18:
Introduction to proof of II Shannon Theorem. Channel Coding, ideas on typical-sequence dcoding. Jointly typical set and its properties. Average and maximum error probability, achievable rate and operative channel capacity. Statement of II Shannon theorem.

CLASS 19:
first hour: proof of direct part of II Shannon theorem. second hour: proof of converse part of II Shannon theorem.

CLASS 20:
first hour: joint source-channel coding theorem. second hour: exercises on channel capacity: 7.8, 7.9 (Z channel) 7.3 (memory increases capacity) 7.12 (unused symbol). Ex 7.23 assigned as homework.

CLASS 21:
Differential entropy (Ch 9): definition, examples (uniform, Gaussian); AEP, properties of Typical set. 2^h=edge of Typical set. Joint and conditional diff. entropy. Ex: multivariate Gaussian. Relative Entropy and mutual information. Inequalities. Hadamard. Shift and change of scale. multivariate Gaussian maximizes entropy at given covariance matrix.

CLASS 22:
Mutual information for discrete X and continuous Y. Ex: Evaluation for PAM signal with equially likely symbols on discrete-time memoryless additive gaussian channel (DTMAGC). Capacity of DTMAGC. Sampling Theorem and Shannon Capacity formula. Gaussian additive noise is a worst case for capacity.

CLASS 23:
Parallel Gaussian channels: capacity. discrete time additive gaussian channels (DTAGC) with memory: capacity. Introduction to Toeplitz matrices and Toeplitz distribution theorem. DTAGC capacity evaluation (water-pouring).

CLASS 24:
continuous-time additive gaussian channel (CTAGC) with memory: capacity evaluation of CTAGC using equivalent input noise, Karhunen-Loeve basis and continuous-time Toeplitz distribution theorem. Examples.