

Programma del Corso di Analisi di Sequenze di Dati
A.A.Laurea Specialistica in Matematica per le Applicazioni
A.A. 2010 - 2011

E. Caglioti

4 giugno 2012

Nota. Per quanto riguarda la Teoria dell'Informazione gli argomenti trattati sono perlopiù contenuti in [1]. In [2] vengono descritti gli algoritmi di Viterbi e di allineamento ed i metodi di ricostruzione filogenetica. I restanti argomenti, compresi alcuni approfondimenti, si trovano su [3].

1 Teoria dell'Informazione

1.1 Entropia, Entropia Relativa ed Informazione Mutua

Entropia ([1] 13-15) - Informazione Mutua ([1] 15-21) - Chain Rule per l'entropia ([1] 21 -22) - Diseguaglianza di Jensen e sue conseguenze ([1] 23-28) - Diseguaglianza Log-Sum e sue applicazioni ([1] 29-31) - La Seconda Legge della Termodinamica ([1] 33-36) - Catene di Markov: convergenza all'equilibrio ([3]) - Diseguaglianza di Csiszar - Kullback ([1] 299 - 301)

1.2 Teorema di Shannon-McMillan

Teorema di Shannon-McMillan per sequenze di variabili indipendenti ([1] 50-53)- Compressione Dati ([1] 53-55)

1.3 Codifica e Tasso di Entropia

Catene di Markov[1] 60-62 -Tasso di Entropia [1] 63-66 - Modelli di Markov Nascosti ([1] 69 - 71)

1.4 Compressione Dati

Esempi di Codici ([1] 78 - 82) - Diseguaglianza di Kraft ([1] 82 - 84) - Codici Ottimali e diseguaglianze([1] 82 -92) - Codice Huffman ed esempi ([1] 92 - 97) - Ottimalità del Codice Huffman ([1] 97 - 101) - Codice di Shannon - Fano - Elias ([1] 101 - 103) - Codifica Aritmetica ([1] 104 - 107) - Generazione di una distribuzione di probabilità discreta a partire da una moneta onesta ([1] 110 - 116) - Entropia dell' Inglese ([1] - [2] 13 - 14) - ^S Codifica di Lempel e Ziv ([1] 319 - 326) - Codice di Slepian-Wolf ([1] 407 - 413)

1.5 Complessità di Kolmogorov

([1] 144 - 153) - Complessità di Kolmogorov ed Entropia ([1] 153 - 155)

2 Segmentazione ed Allineamento di Sequenze

Algoritmo di Viterbi ([2] 21 -26) - Metodi di Allineamento ([2] 26 - 31)

3 Algoritmi di Filogenesi

3.1 Introduzione

Generalità sugli alberi ([2]) - Metodi di Parsimonia e metodo di Fitch ([2] 35 - 37)

3.2 Neighbor Joining

Distanze Ultrametriche ed UPGMA([2] 33 - 34) -Algoritmo di Neighbor Joining ([2] 34 - 35) - NJ ricostruisce correttamente l'albero da una matrice additiva ([3])

3.3 Formula di Pauplin

Formula di Pauplin ([3])

Riferimenti bibliografici

[1] T. M. Cover, J. A. Thomas *Elements of Information Theory*, Wiley & Sons

[2] E. Caglioti, *Fondamenti di Teoria dell'Informazione ed Analisi di Sequenze di Dati: Dispense*,

<http://sites.google.com/site/ecaglioti/didattica>

[3] E. Caglioti, *Analisi di Sequenze di Dati: complementi*,

<http://sites.google.com/site/ecaglioti/didattica>