CMSC 858E
Lecture 27: Last lecture!!!!
12/12/06
- Z algorithm
- Boyer-Moore: bad character rule, good suffix rule, proof of linearity
- KMP
- Aho-Corasick: many patterns, don’t care symbols
- Suffix trees: Ukkonen’s construction, applications
- Inexact alignment: dynamic programming, speed-up/memory reduction heuristics
- Multiple alignments: exact solution, progressive alignment, sum of pairs score, star score, Steiner score, phylogenetic alignment, consensus score
- Phylogenetic trees: scoring, distance methods (neighbor joining, UPGMA), maximum likelihood
- RNA folding: dynamic programming solution, covariance models
- Protein folding: folding on a lattice, threading
Other cool bioinformatics topics

• Gene finding and genome assembly – CMSC 828N
• Gene association studies
• Synthetic biology
• Systems biology
  – protein-protein interaction networks
  – regulatory networks
• Gene expression – microarrays, SAGE, etc.
• Proteomics – protein arrays, mass spectrometry, etc.
• High-throughput experimentation
• Metagenomics – sequencing bugs from the environment
• Databases (good ones) and data mining
• Modeling