

## B363: Bioinformatics algorithms

Review sheet for Final Exam (**Dec. 15 Thursday, 8:30-10:00a**)

<http://darwin.informatics.indiana.edu/col/courses/B363-16>

Basic skills: understanding pseudocodes, desired output; run time analysis, big-O notation; designing algorithms for bioinformatics problems related to those learned in the lectures from Chapter 5-8.

Chapter 5: Manhattan tourist problem. The longest path in a directed graph problem; The Dynamic programming algorithm. The change problem and DP algorithm; DP for sequence alignment; Exercising DP in the Alignment graph; Scoring matrix for sequence alignment; Global alignment and local alignment; DP algorithm for sequencing alignment using affine gap penalty; Divide-and-Conquer algorithm for sequence alignment using linear space.

Chapter 6: Random breakage model; Genome rearrangements; Sorting by reversals; Reversal distances; Greedy algorithm for sorting reversals; Breakpoint theorem; translocations, fusion and fissions; 2-break distance problem; Breakpoint graph; 2-break distance theorem; Fragile breakage model; Synteny block construction;

Chapter 7: Evolutionary tree; Distance-based phylogeny construction; additive matrix; four-point condition; Distance between leaves problem; Limb-length theorem; Additive phylogeny; Least-square phylogeny; Ultrametric trees; UPGMA; Neighbor-joining algorithm; Character-based phylogeny; Small-parsimony problem; DP algorithm for solving small parsimony problem; Large-parsimony problem; nearest-neighbor interchange.

Chapter 8: Diauxic shift; Gene expression matrix; Clustering as an optimization problem; k-means clustering problem; Lloyd algorithm; coin flipping problem; EM algorithm; soft k-means clustering; Distance-based clustering problem; Hierarchical clustering;