Exam II

- Suffix Trees
- Suffix Arrays
- Genome Scale Alignment
- Phylogenetic Trees
  - Parsimony
  - Distance-based
    - Maximum likelihood
- Genome Rearrangement
  - Signed
  - Unsigned
- Protein Structure Alignment
  - DP + Superposition
  - Double DP
  - Geometric Hashing
    - Distance Matrix
- Structure Prediction
  - HP model
Lattice Model: \( Z_2 \) for 2D lattice

\( Z_3 \) for 3D lattice

Hydrophobic core

Energy (conformation) = \(-\text{Hydrophobic contact}\)

Self-Avoiding walk

Self-avoiding walk

Energy = -2

Hydrophobic interface

Non consecutive in sequence
1) **Simulations (Monte Carlo)**

Set of neighbor move

a) End move

b) Corner move: rotation/flip in the same plane (XY, YZ, XZ plane in 3D)

c) Crank-shaft move: rotate 90° (3D)

2D (180°)

d) Pivot move: choose a pivot point i

   Divide the configuration into 2 sides

   Rotate one side in any of the 2d-1 possible neighbors

   \( d = 2 \implies 5 \) neighboring cells

make sure self-avoiding
2) Deterministic approach

\[ 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \]

\[ H \quad P \quad P \quad H \quad H \quad P \quad P \quad H \quad P \]

\[ \Sigma = 2 \cdot 2 \cdot 2 \cdot 2 \]

\[ \text{optimum} = \sqrt{9} \]

\[ \text{lower bound} = \frac{9}{2} \]

\[ = \frac{-5 \cdot (\# \text{internal } H + \# \text{terminal } H)}{2} \]

**Theorem:** In an integer lattice, only possible \( H-H \) contacts are between even & odd points in cells.

\[ \text{Proof:} \]

\[ p \quad 23 \]

\[ 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \]

\[ H \quad P \quad P \quad H \quad H \quad P \quad P \quad H \quad P \]

\[ -1 \]
\[ P_{34} : \text{same as } P_{23} \]

\[ P_{45} : \]

\[ P_{52} : \]

\[ \text{best } : -1 \]

\[ \text{approx. ratio: } 4 \quad (\text{asymptotically}) \]

\[ \lim_{n \to \infty} \]

\[ \text{max} \left\{ \begin{array}{l}
\# a, \text{even } H \\
\# 3, \text{even } H
\end{array} \right\} \]

\[ \begin{array}{ccccccc}
| & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 |
\end{array} \]

\[ \begin{array}{ccccccc}
P & 1 & H & P & H & P & H & P & H
\end{array} \]

\[ \begin{array}{ccccccc}
11 & 12 & 13 & 14 & 15 & 16 & 17 & 18
\end{array} \]

\[ \begin{array}{ccccccc}
& 1 & H & P & H & P & H & P & H
\end{array} \]

\[ \begin{array}{ccccccc}
\text{odd} & 14 & 15 & 16 & 17 & 18 & 19 & 20
\end{array} \]

\[ \begin{array}{ccccccc}
\text{even} & 21 & 22 & 23 & 24 & 25 & 26 & 27
\end{array} \]

\[ N_e = 2 \]

\[ N_o = 3 \]

\[ N_e = 2 \]

\[ N_o = 2 \]
1) Define the even and odd block.

2) Try the prior points that are between blocks.

3) Balance the # of odds with the # of evens on either side.

\[
\text{Score}(d, R) = \max \left\{ \min (N_L^E, N_0^R), \min (N_0^L, N_R^E) \right\}
\]

Choose the max score \((d, R)\)