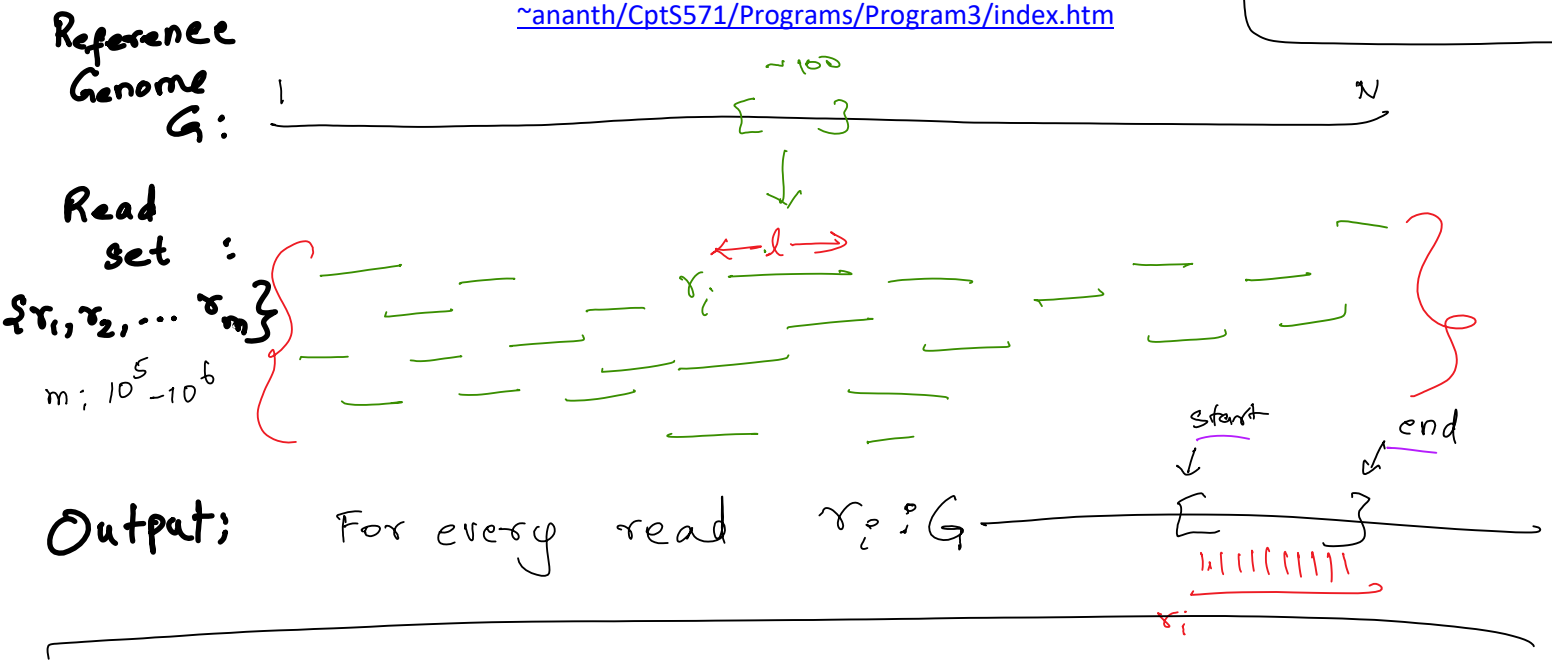
Read Mapping: Project 3

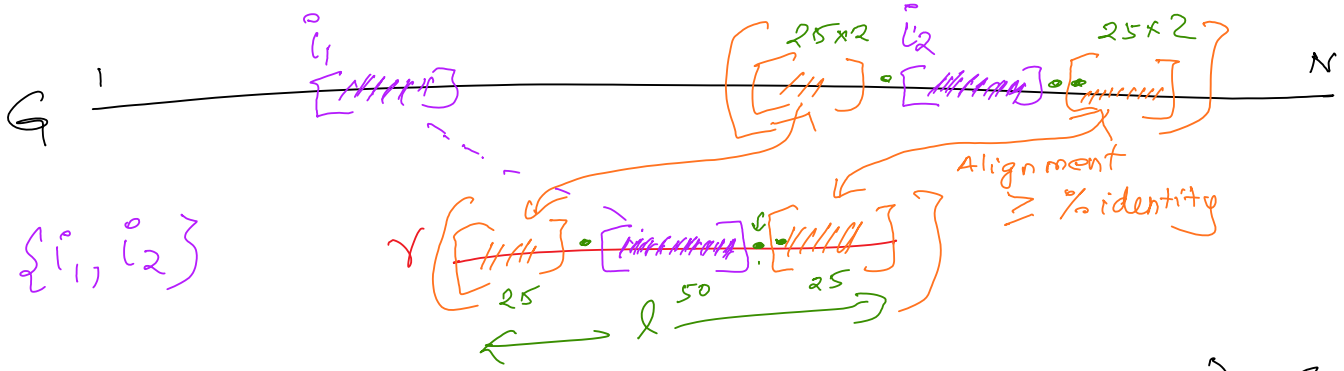
Tuesday, April 2, 2019 11:45 AM

471, 571 → Proj 3 + hw + Exam

571 → Survey

https://eecs.wsu.edu/~ananth/CptS571/Programs/Program3/index.htm





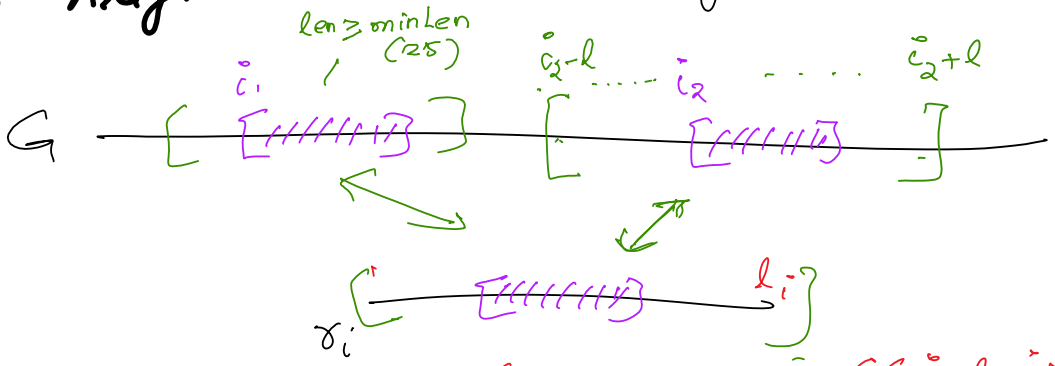
Read Mapping ($G[1..N]$, $R = \{r_1, r_2, \ldots r_m\}$) {

1. "Preprocess" (G)
2. For (i = 1 to m) // every read r_i
   {
   2.a) $\{l_1, l_2, \ldots l_k\} \leftarrow$ **FindLoc** (r_i , ST(G)) = ?
   // candidate loci which are starting indices for lcs (r_i , G)

Read Mapping

Tuesday, April 2, 2019 12:26 PM

2.b) "Align" candidate loci against the read r_i.



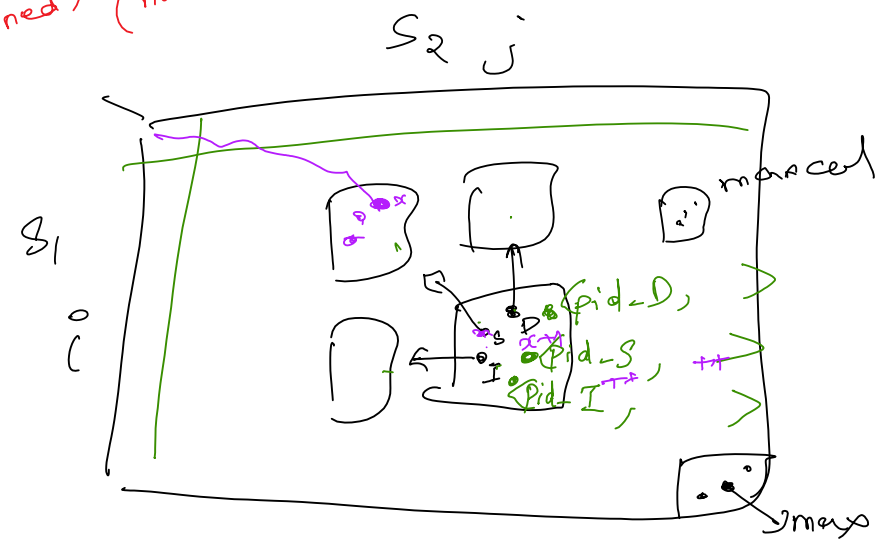
Use local Align ($r_i[1..l_i]$ vs. $G[i_1-l_{..}i_1+l]$)

% identity = #matches / Align.len

s_1: a c g t - t g
s_2: a c t t a c -

3/7

% ident ≥ cutoff? Yes (aligned) No (not aligned)

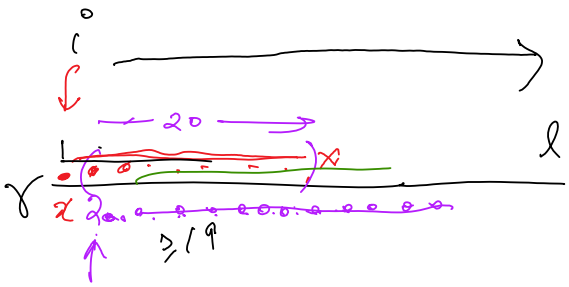


At the end, pick the best aligning region for each read.

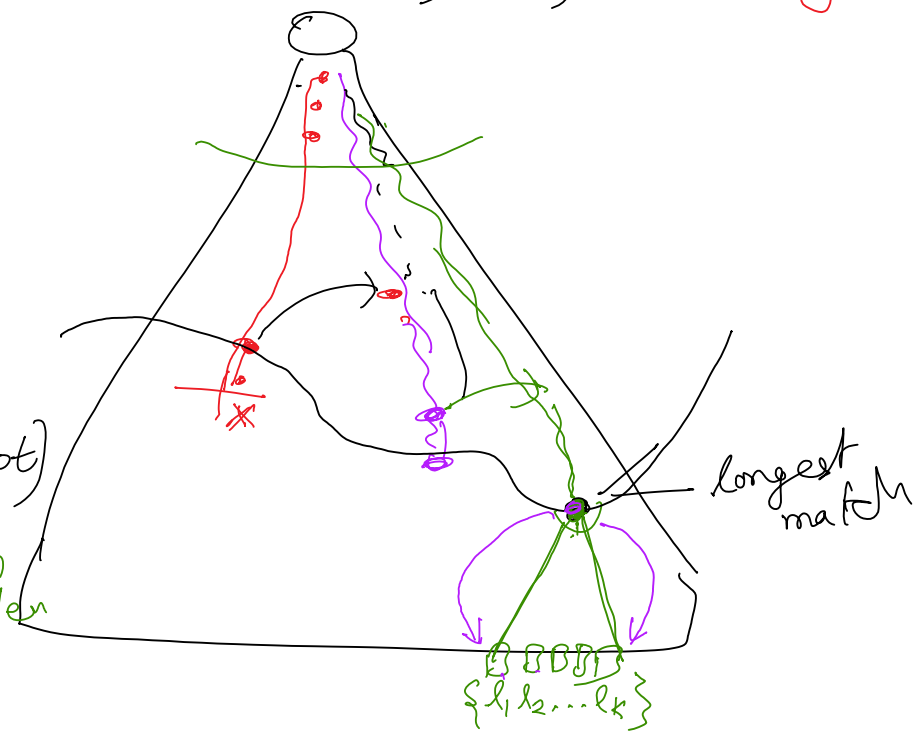
Read Mapping

Tuesday, April 2, 2019 12:50 PM

FindLoc: Find all locations that are starting indices on G, for lcs(r, G)



STG) // read only




```
For (i = 1 to l)
{
    len ← FindPath(r[i...], root)
    maxlen ← update new len > maxlen
}
```

Read Mapping

Tuesday, April 2, 2019 1:07 PM

Preprocess (G):

↳ 1. Build ST(G)

Prepare ST ⇒ DFS