Dynamic Programming

Sequence Alignment Problem

A DNA strand consists of a string of molecules called bases!

4 Types of bases:

- Adenine A
- Cytosine C
- Guanine G
- Thymine T

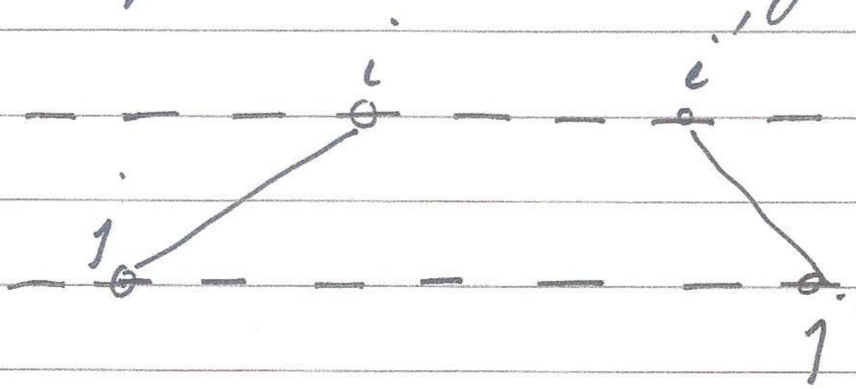
Suppose we have 2 strings X & Y.

$X = \{x_1 x_2 \ldots \ldots x_m\}$

$Y = \{y_1 y_2 \ldots \ldots y_n\}$

Def. A matching is a set of ordered pairs with property that each item occurs at most once.

__Def.__ A matching is an __alignment__ if there are no crossing pairs.



$(i,j),\ (i',j') \in M$

& $i < i' \Rightarrow j < j'$

For an alignment M between X & Y

1- We incur a "gap penalty" of $\underline{\delta}$ for each gap.

2- For each mismatch (of letters p & q) we incur a mismatch cost α_{pq}

Def. **Similarity** between strings X & Y is the minimum cost of an alignment between X & Y.

$X = \{x_1, \ldots, x_m\}$

$Y = \{y_1, \ldots, y_n\}$

Say M is an opt. solution.

either $(x_m, y_n) \in M$ or $(x_m, y_n) \notin M$

Define OPT(

In an optimal alignment M, at least one of the following is true:

1) - $(X_m, Y_n) \in M \Rightarrow OPT(m,n) = OPT(m-1,n-1) + \alpha_{X_m Y_n}$

2) - X_m is not matched $\Rightarrow OPT(m,n) = OPT(m-1,n) + \delta$

3) - Y_n is not matched $\Rightarrow OPT(m,n) = OPT(m,n-1) + \delta$

Matrix Chain Multiplication

Shortest Path Problem

Dynamic Programming

Bellman-Ford Alg.

Shortest-path (G, s, t)

n = no. of nodes in G

define $M[0,t]=0$, $M[0,v]=\infty$

for $i=1$ to $n-1$

for $v \in V$ in any order

$$M[i,v] = \min\left(M[i-1,v], \min_{w \in Adj(v)}\left(M[i-1,w] + c_{vw}\right)\right)$$

end for

end for

Discussion 7

1. When their respective sport is not in season, USC's student-athletes are very involved in their community, helping people and spreading goodwill for the school. Unfortunately, NCAA regulations limit each student-athlete to at most one community service project per semester, so the athletic department is not always able to help every deserving charity. For the upcoming semester, we have S student-athletes who want to volunteer their time, and B buses to help get them between campus and the location of their volunteering. There are F projects under consideration; project i requires s_i student-athletes and b_i buses to accomplish, and will generate $g_i > 0$ units of goodwill for the university. Our goal is to maximize the goodwill generated for the university subject to these constraints. Note that each project must be undertaken entirely or not done at all -- we cannot choose, for example, to do half of project i to get half of g_i goodwill.

2. Suppose you are organizing a company party. The corporation has a hierarchical ranking structure; that is, the CEO is the root node of the hierarchy tree, and the CEO's immediate subordinates are the children of the root node, and so on in this fashion. To keep the party fun for all involved, you will not invite any employee whose immediate superior is invited. Each employee j has a value v_j (a positive integer), representing how enjoyable their presence would be at the party. Our goal is to determine which employees to invite, subject to these constraints, to maximize the total value of invitees.

3. You are given a set of n types of rectangular 3-D boxes, where the i^{th} box has height h(i), width w(i) and depth d(i) (all real numbers). You want to create a stack of boxes which is as tall as possible, but you can only stack a box on top of another box if the dimensions of the 2-D base of the lower box are each strictly larger than those of the 2-D base of the higher box. Of course, you can rotate a box so that any side functions as its base. It is also allowable to use multiple instances of the same type of box.