EDIT DISTANCE

INPUT: Two strings $x[1 \ldots m]$ $y[1 \ldots n]$.

GOAL: Compute the smallest number of edits to go from $x \longrightarrow y$

- insert a character
- delete
- substitution

Example

$x =$ FAST
$y =$ CATS

F	A	S	T	_	_	_	_	cost = 8
_	_	_	_	C	A	T	S	

cost = 4

F	_	A	_	S	T
_	C	A	T	S	_
d	i	.	i	.	d

Cost = 3

$\boxed{F}$	A	S	T
$\boxed{C}$	A	T	S
s	.	s	s

Insertion	_ / A
Deletion	A / _
Substitution	F / C

$$
\begin{array}{cccccc}
\multicolumn{6}{c}{\begin{pmatrix} F & \_ & A & \_ & S & T \\ \_ & C & A & T & S & \_ \end{pmatrix}} \\
d & i & . & i & . & d
\end{array}
$$

"Optimal Solution"

_	F	A	. . . .	T
C	_	A		_

4

optimal way of editing

$x[1 \ldots n] \to y[1 \ldots m-1]$

$x[1 \ldots n-1] \to y[1 \ldots m]$

$x[1 \ldots n-1] \to y[1 \ldots n-1]$

Subproblem: $E[i,j] =$ edit distance between $x[1..i]$ and $y[1..j]$

$$\{ E[i,j] : i = 1..m \quad j = 1..n \}$$

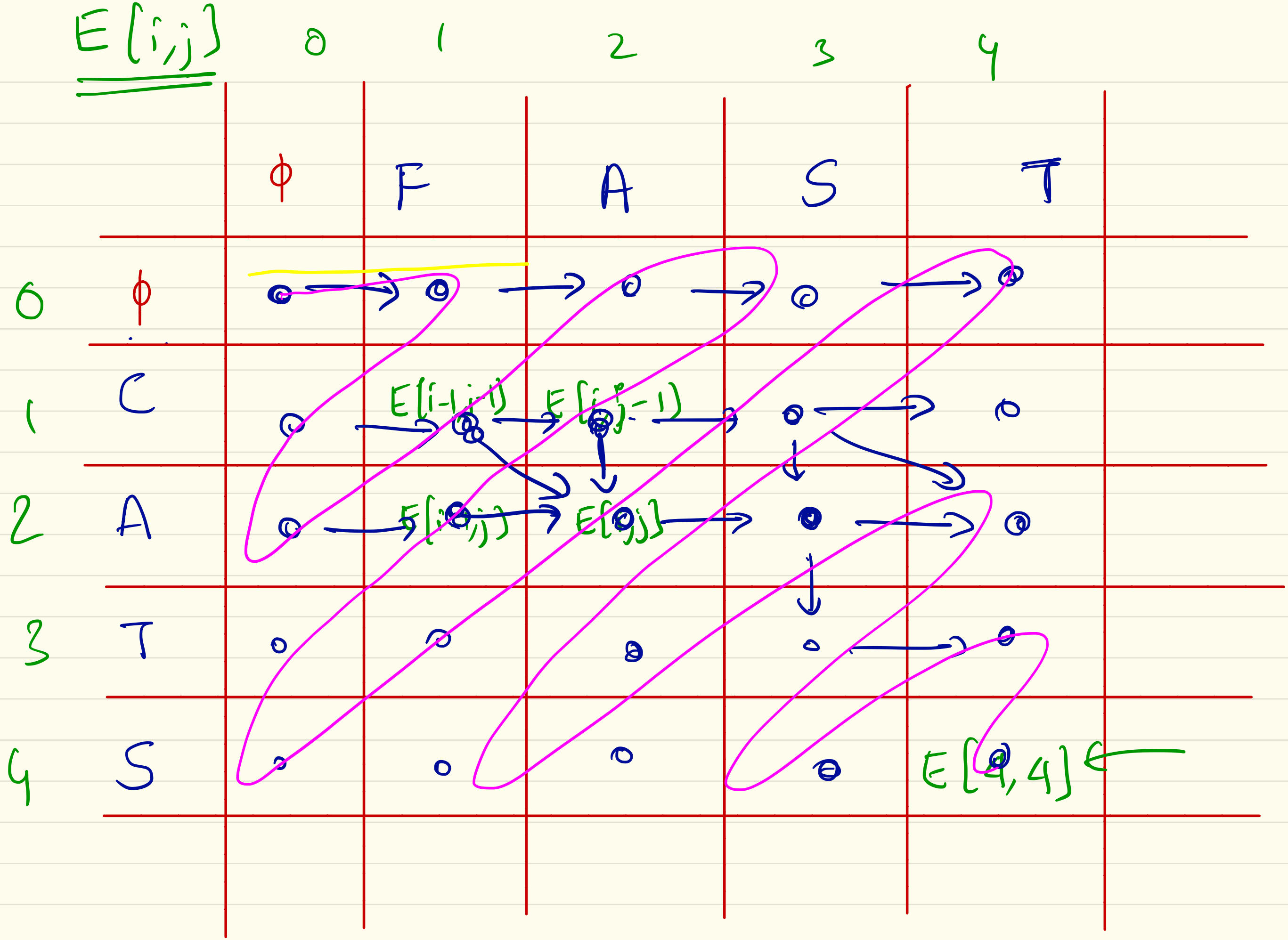
~~Re~~ $E[m,n] =$ edit distance between x & y.

Recurrence Relation:

$$E[i,j] = \min \begin{cases} \overset{\text{insertion}}{\boxed{\begin{matrix} - \\ a \end{matrix}}} & E[i, j-1] + 1/2 \leftarrow \text{insertion} \\ & E[x[1..i], y[1..j-1]] + 1 \\ \boxed{\begin{matrix} a \\ - \end{matrix}} & E[i-1, j] + 1/7 \leftarrow \text{deletion} \\ x[i] = y[j] \quad \boxed{\begin{matrix} a \\ a \end{matrix}} & E[i-1, j-1] \leftarrow \text{carry over.} \\ x[i] \neq y[j] \quad \boxed{\begin{matrix} a \\ b \end{matrix}} & E[i-1, j-1] + 1/17 \leftarrow \text{substitution} \end{cases}$$

$$\text{''}$$

$$E[x[1..i], y[1..j]]$$


E[i,j]
0 1 2 3 4
ϕ F A S T
0 ϕ
1 C
2 A
3 T
4 S
E[i-1,j-1]
E[i,j-1]
E[i-1,j]
E[i,j]
E[4,4]

int E[m, n]

for i = 0 to m

$\quad$ E[i,0] = i $\leftarrow$ deletion $\left\{ x[1..i] \xrightarrow{i} \phi_{\text{empty}} \right\}$

for j = 0 to n

$\quad$ E[0,j] = j $\leftarrow$ insertion $\left\{ \phi \longrightarrow y[1..j] \right\}$

for i = 1 to m

$\quad$ for j = 1 to n

$\quad\quad$ E[i,j] = $\left\{ \text{reccurence relation.} \right\}$

$E[i,j] =$ minimum

prev$[i,j]$

		F	A	S	T
	0	1	2	3	4
C	1	1	2		
A	2				
T	3				
S	4				

Case 1

$\boxed{\begin{array}{c} - \\ F \end{array}}$ $E[\underbrace{i, j-1}_{\text{insert}}] + 1$

Case 2

$\boxed{\begin{array}{c} F \\ - \end{array}}$ $E[\underbrace{i-1}_{\text{deletion}}, j] + 1$

Case 3

$\boxed{\begin{array}{c} A \\ B \end{array}}$ $E[i-1, j-1] + 1$

$\| \; E[i-1, j-1] + \text{diff}(x[i], y[j])$

Case 4

$\boxed{\begin{array}{c} A \\ A \end{array}}$ $E[i-1, j-1] + 0$

$$E[1,3] = \min \begin{cases} \underline{E[0,3]} + 1 \\ \underline{E[1,2]} + 1 \\ \underline{E[0,2]} + \mathrm{diff}(x[1], y[3]) \end{cases}$$

$$\mathrm{diff}(x[1], y[3]) = \begin{cases} 1 & \text{if } x[i] \neq y[j] \\ 0 & \text{otherwise} \end{cases}$$

GAMBLING STRATEGY

- Play exactly $n = 200$ gambles in a casino

- 2 games of choice

Game 1: → w. prob 1/2 win 2 \$
→ w. prob 1/2 lose 2 \$

Game 2 → w. prob 2/3 win 7 \$
→ w. prob 1/3 lose 7 \$

- Win a prime number of \$ at end

win 170 \$ at end

GOAL: Find the optimal strategy that succeeds with maximum prob.

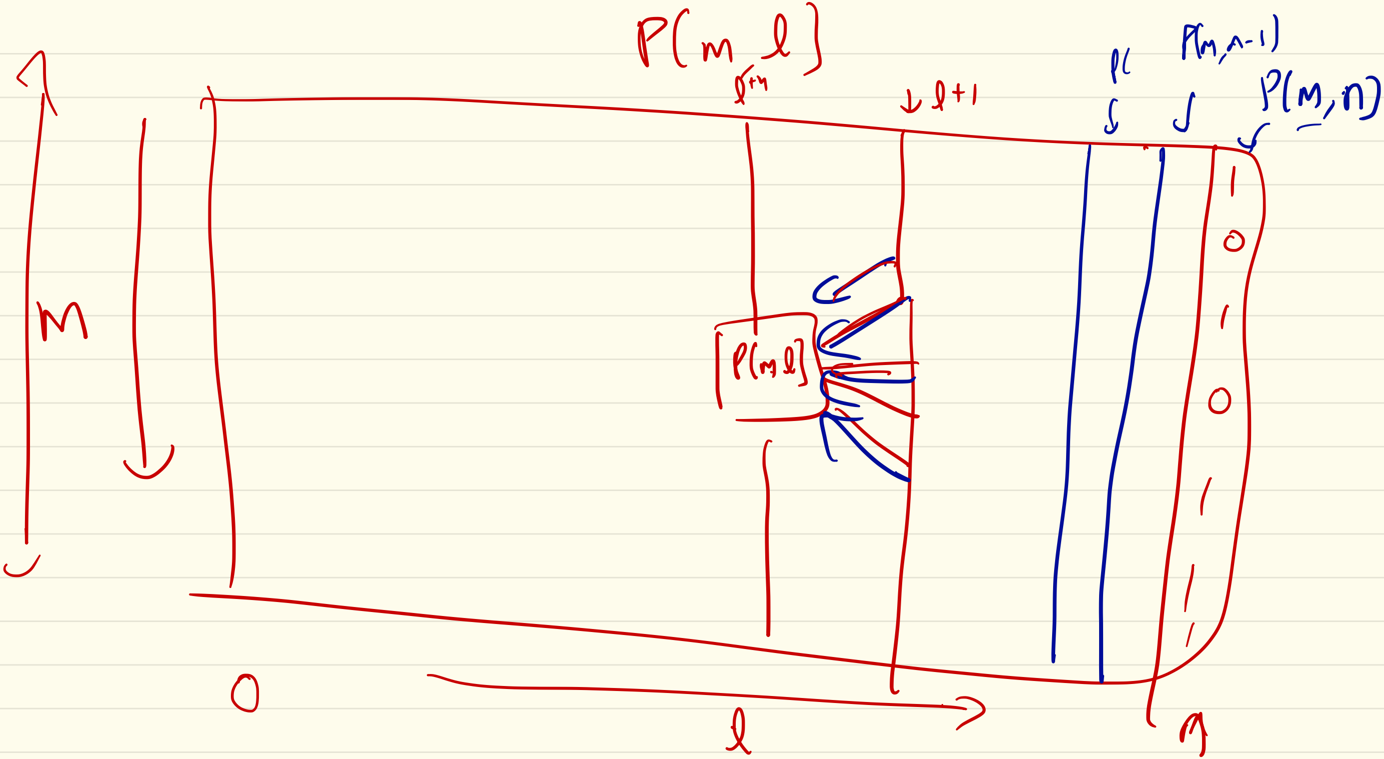
Strategy: $A \left[\begin{matrix} \text{won} \\ m\$ \end{matrix} \;,\; \begin{matrix} \text{end of} \\ \ell \text{ games} \end{matrix} \right]$ → 1 ?? → 2 ??

Subproblem: $P[m, \ell] =$ probability of "success" of optimal strategy given that $m\$$ after ℓ games.

Reccurrence Relation:

Base Case: $P[m, n] = \begin{cases} 1 & \text{if } m \text{ is a prime} = 170 \\ 0 & \text{otherwise} \end{cases}$ o.w.

$$P[m, \ell] = \max \begin{cases} \text{Game 1: } & \frac{1}{2} P[m+2, \ell+1] + \frac{1}{2} P[m-2, \ell+1] \\ \text{Game 2: } & \frac{2}{3} P[m+7, \ell+1] + \frac{1}{3} P[m-7, \ell+1]. \end{cases}$$



"decreasing order of l"