

# Info based on questions I've received

- Based on the scoring matrix we've discussed, can we assume that the edge cost for any horizontal or vertical move is -1?
  - In a linear gap penalty framework, yes. Every horizontal or vertical move is a gap, because it only makes progress in one of the sequences, not both – hence the other sequence is a gap.
  - This is not true in an affine gap penalty framework, as we will see today

# Info based on questions I've received

Additional parameters:

**Genetic code:**

- ☐ 11 (Bacteria, Archaea)
- ☒ 4 (Mycoplasma/Spiroplasma)

**Topology:**

- ☐ circular
- ☒ linear

Run GLIMMER v3.02

- Should we use the linear or circular settings on NCBI Glimmer?
  - Linear. We are not dealing with organisms with circular DNA.
- Regarding the genetic coding, should we use “Bacteria, Archaea” or “Mycoplasma, Spiroplasma”?
  - Mycoplasma. I seem to be getting better results with this, even though neither groups is any closer to viruses than the other.

# A general overview on deciphering contigs



## A jigsaw puzzle on contigs

- The contigs you get are not assembleable: assume they are the result of some assembly process/software
- Genes from the virus are in the contigs, but the sequencing process does not guarantee that the genes are not spread across a few contigs
- Thus, gene predictions made with Glimmer and Genemark can be inaccurate for a number of reasons:
  - The gene might not be wholly contained on the contig, and thus it's incompleteness might throw off Glimmer
  - Glimmer and Genemark are not 100% accurate, so they could simply be wrong (note that they are not trained on viruses either)

# Making gene predictions on this data

1. Use Nano to assemble putative contigs
  - You know the contigs fit together somehow, so there are only so many combinations. Try putting a few contigs together and submitting the concatenated string to see how that changes Glimmer/Genemark's predictions
2. Use Blast to search for genes predicted by Glimmer and Genemark
  - Blast can help you find the gene you are looking for, and it's total length – use the length and the length of the contig, to guess if you need another contig added on, and to also guess which contigs connect to which.

You should be able to find all genes and which contigs they are on through this kind of deduction.

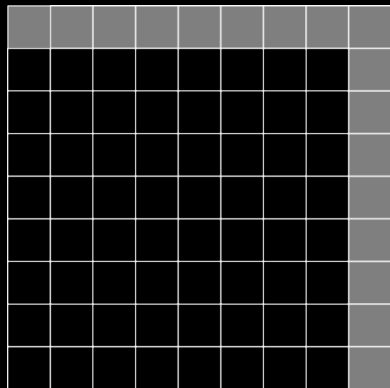
# Dynamic Programming with Affine Gap Penalties Part 1

# A biological rationale for gap penalties

- The purpose of gap scoring:
  - To bias automated alignment algorithms towards realistic explanations of insertion and deletion
- The grid of alignment possibilities outlines all possible alignments – mostly implausible:

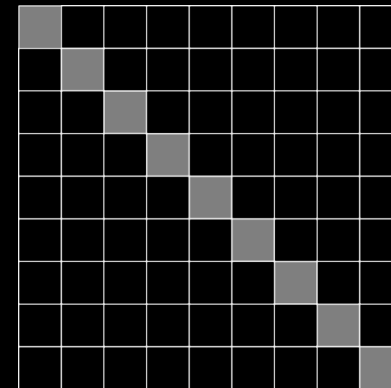
AAAGAATCA-----  
-----AAAGAATCA

Implausible alignment



AAAGAATCA  
AAAGAATCA

Plausible alignment



# Linear gap penalties are implausible

- We score implausible alignments negatively
  - The more implausible the alignment is, the more negatively it SHOULD score
    - Otherwise our scoring system is bad.
- But large insertions are not much more implausible than small insertions.
  - The same mechanics cause them:
    - Crossovers, viral insertions, etc etc
  - The probability of small insertions is not so different from large insertions, relative to the probability of no insertion at all.
- Affine penalties penalize less for longer deletions or insertions

# An example of linear gap failure

AAAG~~GA~~TTCA

A-A-A-T-CA

An implausible alignment

AAAG~~GA~~TTCA

AAA-----TCA

An plausible alignment

- We score implausible alignments negatively
- Large insertions are not much more implausible than small insertions.
  - The same mechanics cause them:
    - Crossovers, viral insertions, etc
- The probability of small insertions is not so different from large insertions
- Affine penalties penalize less for longer deletions or insertions



# What is an Affine Gap Scoring Penalty?

- The penalty has two parts:
  - S, the Gap Opening Penalty: -10
    - This is the penalty for opening a new gap
  - E, the Gap Extension Penalty: -.5
    - This is the penalty for making a gap bigger
- Gap Scoring Penalty =  $S + d * E$ 
  - For gaps of size d.

# Affine Gaps and Dynamic Programming

		G	G	A	T	C
C  G  C	0	-1	-2	-3	-4	-5
	-1	-2	-3	-4	-5	-3
	-2	0	-1	-2	-3	-4
	-3	-1	-2	-3	-4	-2

	G	G	A	T	C	
	0	-1	-2	-3	-4	-5
C	-1	-2	-3	-4	-5	-3
G	-2	0	-1	-2	-3	-4
C	-3	-1	-2	-3	-4	-2

- Using Dynamic Programming, we could fill out a grid and backtrack to find the optimal alignment
  - Every grid square had an optimal score
  - Adding a new nucleotide to the end of the sequence requires calculation only of the new squares
- With Affine Gaps, the score depends on how long the gap was. But we don't know that from a single grid.

# Example: How affine scoring is different

match = 10, mismatch = -2, gap = -7, gap\_start = -15

Alignment

Scoring

CAGT  
CA-T

Linear score:  $30 - 15 - 7 = 8$

CAGTA  
CA-T-

WRONG:  $30 - 15 - 7 - 15 - 7 = -14$

This does not correctly implement  
Affine Gap Scoring

CAGTA  
CAT--

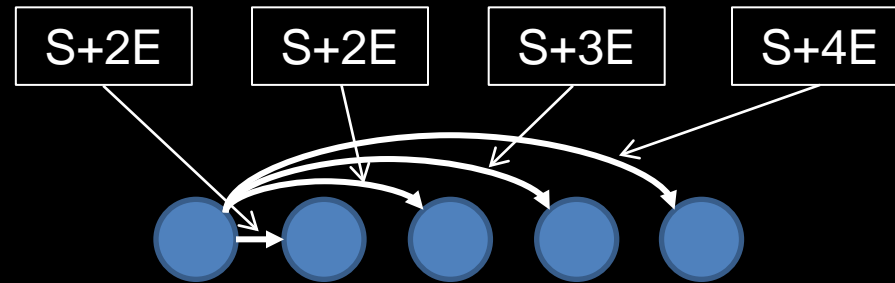
Correct:  $20 - 2 - 15 - 14 = -11$

Mismatch is outscored by the cost of a  
second gap opening penalty

- Under affine gap scoring, adding a new nucleotide changes the entire score

# A false start on affine gap penalty scoring

- You could add more gap edges in your graph

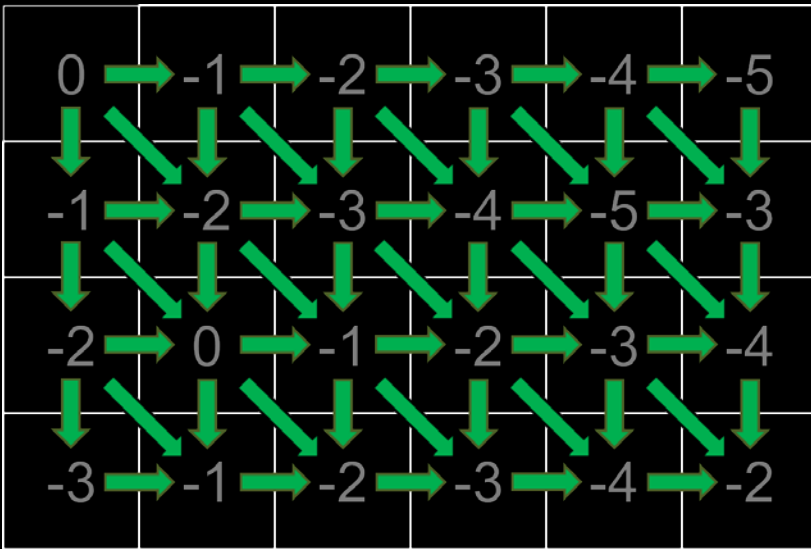


$S$  = gap opening cost    $E$  = gap extension cost

- This would correctly represent the gap opening and extension costs (if you did it in both dimensions, for every node)
- Unfortunately, you would have to build many more edges:  $\sim mn^2 + nm^2$  rather than  $\sim mn$
- It would take much longer to update and score such a graph.

# Topology of the Affine Gap Scoring Graph

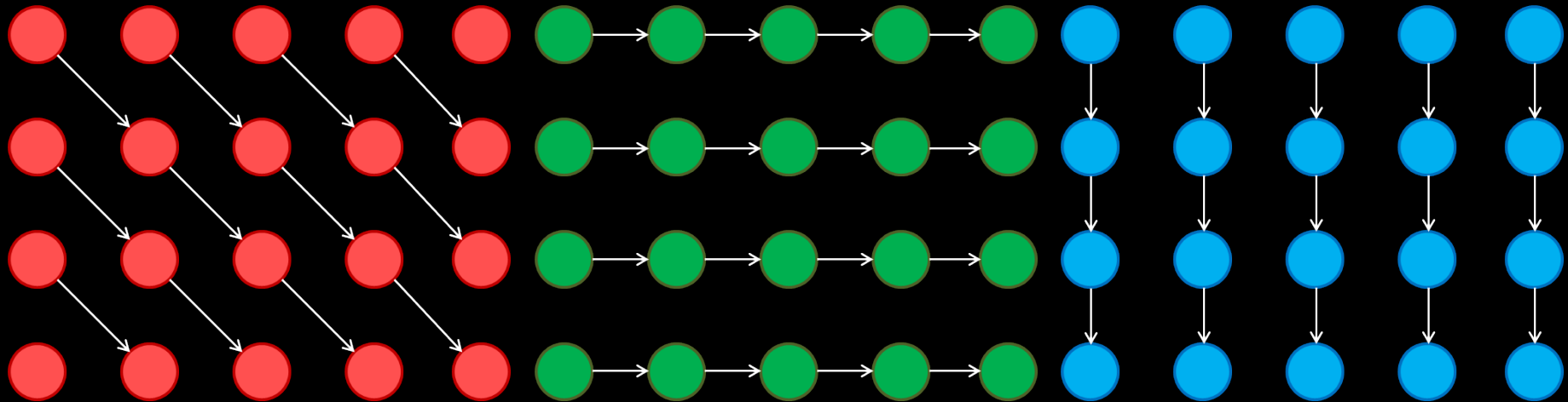
The Linear Gap scoring graph



The Affine Gap Scoring Graph

(Below)

- This graph is just a deconstructed version of the linear graph, with 3 subgraphs: **M**, **X**, **Y**.
- **M** scores matches, mismatches
- **X** and **Y** score gaps
- Edges between them score Gap Extension penalties (not shown)

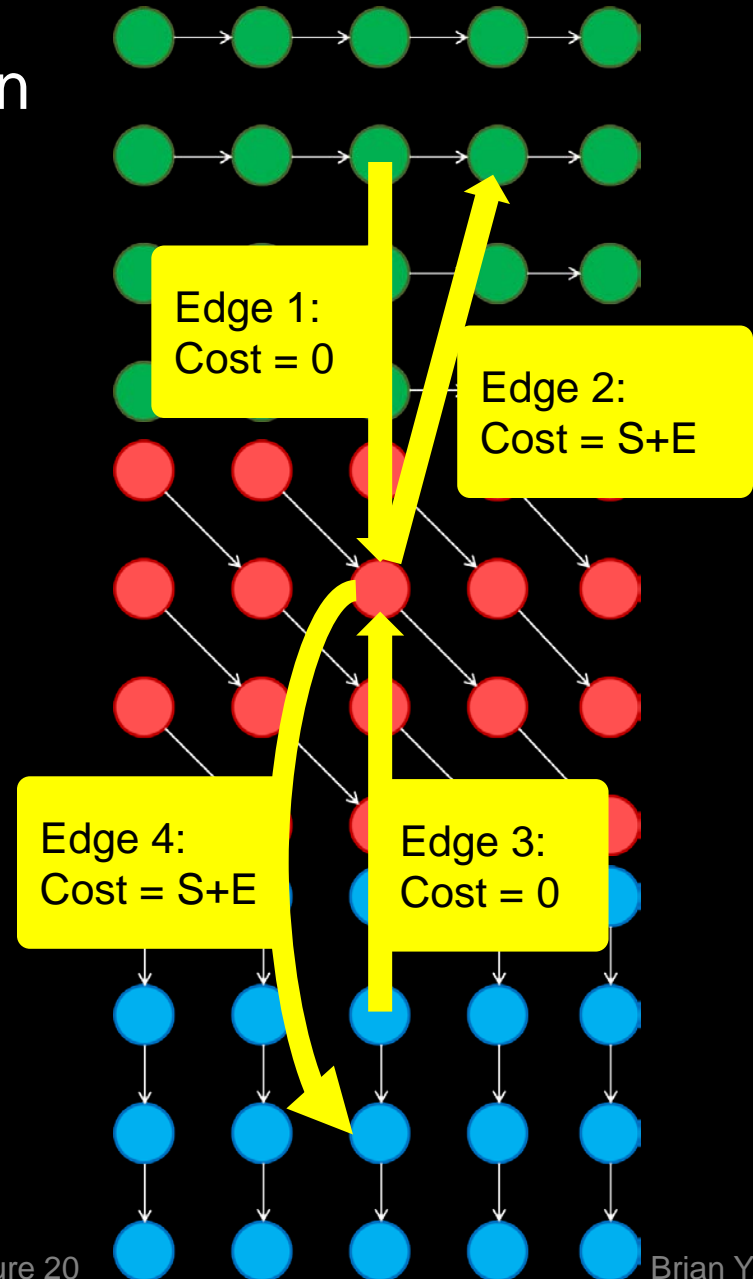


# Edges connecting M, X, Y are gap starters

- Every node in M is connected to two nodes in X, and two nodes in Y.

Start cost = S, Extension = E

- Lets look at M[2, 1]:
  - Edge 1 comes from X[2, 1]
    - “done with horizontal gap”
  - Edge 2 goes to X[3, 1]
    - “starting horizontal gap”
  - Edge 3 comes from Y[2, 1]
    - “done with vertical gap”
  - Edge 4 goes to Y[2,2]
    - “starting vertical gap”
- Inter-graph edges start and end gaps.
- Not Shown: Edges from X to Y
  - “cross sequence gaps”



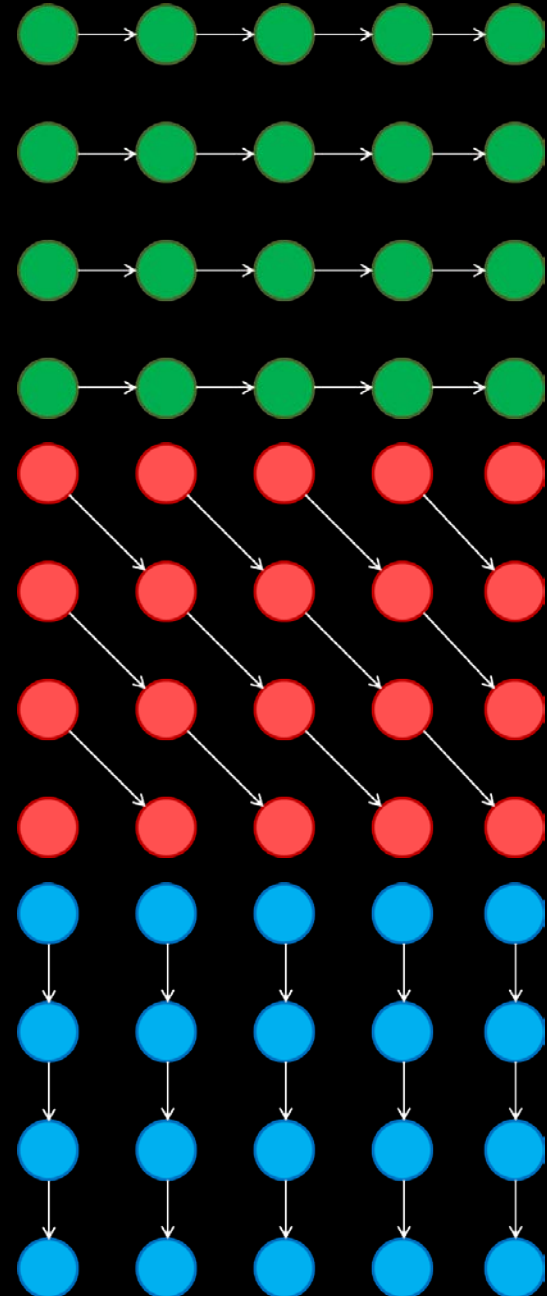
# Edges Values inside the 3 matrices

Edges in X have the weight of exactly one gap extension  
See previous slide for inter-subgraph edges

S = Gap opening, E = Gap extension

Edges in M have the weight exactly one match or mismatch (depending on the row/column),  
See previous slide for inter-subgraph edges

Edges in Y have the weight of exactly one gap extension. See previous slide for inter-subgraph edges



# Node values in the 3 matrices

M: scores of alignments  
ending in match, mismatch

X: scores of alignments  
ending in gaps in seq. x

Y: scores of alignments  
ending in gaps in seq. y

Score of a match  
or mismatch

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$



# Explaining the scores

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} \text{Score of a match or mismatch} \\ M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

- $M[i, j]$  is the score you get from a matching or mismatching pair of bases.
  - The alternate  $X$  and  $Y$  scores correspond to ending a gap for zero cost.
  - Some gaps end in a match or a mismatch, and this is how it ends – the  $X$  and  $Y$  scores here are the edges coming back from the other two matrices

# Explaining the scores

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

- $X[i, j]$  is the score you get for gapping in  $X$ .
  - The  $S + E$  cost in the  $M()$  component is the cost of starting a new gap ( $S$ ) from a matched pair, plus the first extension and whatever the score from  $M$  was.
  - The  $E$  cost alone is simply extending a gap, so you do not incur a new gap opening cost
  - The  $S + E$  cost in the  $Y$  component is a the cost of starting a new gap on the  $X$  sequence whereas before it was a gap on the  $Y$  sequence

# Explaining the scores

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$

- $Y[i, j]$  is the score you get for gapping in  $Y$ .
  - The  $S + E$  cost in the  $M()$  component is the cost of starting a new gap ( $S$ ) from a matched pair, plus the first extension and whatever the score from  $M$  was.
  - The  $E$  cost alone is simply extending a gap, so you do not incur a new gap opening cost
  - The  $S + E$  cost in the  $X$  component is a the cost of starting a new gap on the  $Y$  sequence whereas before it was a gap on the  $X$  sequence

# Major steps in Alignment with Affine Gaps

- Scoring
  - Scoring is a little complex, because it does additional inter-subgraph nodes. Otherwise it is the same scoring methods used in linear gap scores
- Backtracking
  - Backtracking the graph requires us to check all connected graph nodes, but is otherwise exactly the same as before
- Reference for the next slides:
  - X is the vertical sequence, Y the horizontal sequence
  - i iterates vertically, j iterates horizontally, when both used
  - S is the gap start penalty, -10.
  - E is the gap extension penalty, -.5.

- This spot starts as zero

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Matches, Mismatches only

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Vertical gaps only

- This means “score of best alignment between 0 bases of x and i bases of y that ends in a gap in x”
- In other words, x is all gaps against y, so score =  $S + i \cdot E$

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Matches, Mismatches only

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	.	.	.	.	.
T	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Vertical gaps only

- This row means “score of the best alignment between i bases of x and 0 bases of y, ending in a gap in x”.
- This is impossible; the bases of x cannot be a gap.
  - so the score is  $-\infty$

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Matches, Mismatches only

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Vertical gaps only

- This row means “score of the best alignment between  $i$  bases of  $y$  and 0 bases of  $x$ , ending in a gap in  $Y$ .”
- This is impossible; the bases of  $Y$  cannot be a gap.
  - so the score is  $-\infty$

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Matches, Mismatches only

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	.	.	.	.	.
T	.	.	.	.	.

Vertical gaps only



- This column means “score of best alignment between 0 bases of y and i bases of x that ends in a gap in y”
- In other words, y is all gaps against x, so score =  $S + i \cdot E$

	C	A	G	T	A
C	0	.	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Matches, Mismatches only

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	.	.	.	.
T	-11	.	.	.	.

Vertical gaps only

- This means: “score of best alignment between 0 bases of x and 1 base of y that ends in a match”. Impossible.
- Score:  $-\infty$ , a special case.

	C	A	G	T	A
C	0	$-\infty$	.	.	.
A	.	.	.	.	.
T	.	.	.	.	.

Matches, Mismatches only

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

$-\infty$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	.	.	.	.
T	-11.5	.	.	.	.

Vertical gaps only

- This point extends to all other values of  $M[0, i]$  and  $M[i, 0]$ .
- No alignment of 0 and i bases can end in a match.

		C	A	G	T	A
	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
C	$-\infty$	.	.	.	.	.
A	$-\infty$	.	.	.	.	.
T	$-\infty$	.	.	.	.	.

Matches, Mismatches only

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

$-\infty$

		C	A	G	T	A
	0	-10.5	-11	-11.5	-12	-12.5
C	$-\infty$	.	.	.	.	.
A	$-\infty$	.	.	.	.	.
T	$-\infty$	.	.	.	.	.

Horizontal gaps only

		C	A	G	T	A
	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
C	-10.5	.	.	.	.	.
A	-11	.	.	.	.	.
T	-11.5	.	.	.	.	.

Vertical gaps only

- $S + E + M(i, j-1) = -\infty$
- $E + X(i, j-1) = -\infty$
- $S + E + Y(i, j-1) = -21$
- This means you gapped the C in Y, then gapped the C in X.

C-  
-C

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

	C	A	G	T	A	
	0	-10.5	-11	-11.5	-12	-12.5
C	$-\infty$	-21	.	.	.	.
A	$-\infty$	.	.	.	.	.
T	$-\infty$	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	.	.	.	.
T	-11	.	.	.	.

Vertical gaps only

- $S + E + M(i-1, j) = -\infty$
- $S + E + X(i-1, j) = -21$
- $E + Y(i-1, j) = -\infty$
- This means you gapped the C in X, then gapped the C in Y.

-C  
C-

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$

	C	A	G	T	A	
	0	-10.5	-11	-11.5	-12	-12.5
C	$-\infty$	-21	.	.	.	.
A	$-\infty$	.	.	.	.	.
T	$-\infty$	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	.	.	.
T	-11	.	.	.	.

Vertical gaps only

- $M(i-1, j-1) = 0$
- $X(i, j) = -21$
- $Y(i, j) = -21$
- $\text{match}(i, j) = +1$  (C matches C)

C  
C

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	.	.	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	-21	.	.	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	.	.	.
T	-11	.	.	.	.

Vertical gaps only

- Now we do X:
  - i.e. suppose we start a new gap on x
- $S + E + M(i, j-1) = -9.5$
- $E + X(i, j-1) = -21.5$
- $S + E + Y(i, j-1) = -31.5$

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	.	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

	C	A	G	T	A	
X	0	-10.5	-11	-11.5	-12	-12.5
C	$-\infty$	-21	-9.5	.	.	.
A	$-\infty$	.	.	.	.	.
T	$-\infty$	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	.	.	.
T	-11	.	.	.	.

Vertical gaps only

- Now we do Y:
  - i.e. suppose we start a new gap on y
- $S + E + M(i-1, j) = -\infty$
- $S + E + X(i-1, j) = -21.5$
- $E + Y(i-1, j) = -\infty$

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	.	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$

		C	A	G	T	A	
y		0	-10.5	-11	-11.5	-12	-12.5
C		-∞	-21	-9.5	.	.	.
	A	-∞	.	.	.	.	.
	T	-∞	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
C	-10.5	-21	-21.5	.	.
A	-11	.	.	.	.
T	-11.5	.	.	.	.

Vertical gaps only



- Now, the next square:
- $\text{match}(i, j) = -4$  (mismatch)
  - $M(i-1, j-1) = -\infty$
  - $X(i, j) = -9.5$
  - $Y(i, j) = -21.5$

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	-9.5	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	-21	-9.5	.	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	-21.5	.	.
T	-11	.	.	.	.

Vertical gaps only

- Suppose a gap opened on X:
- $S + E + M(i, j-1) = -20$
- $E + X(i, j-1) = -10$
- $S + E + Y(i, j-1) = -32$

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	-9.5	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$X[i, j] = \max \begin{cases} S + E + M(i, j-1) \\ E + X(i, j-1) \\ S + E + Y(i, j-1) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	-21	-9.5	-10	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	-21.5	.	.
T	-11	.	.	.	.

Vertical gaps only

- Suppose a gap opened on Y:
- $S + E + M(i-1, j) = -\infty$
- $S + E + X(i-1, j) = -22$
- $E + Y(i-1, j) = -\infty$

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	-9.5	.	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$Y[i, j] = \max \begin{cases} S + E + M(i-1, j) \\ S + E + X(i-1, j) \\ E + Y(i-1, j) \end{cases}$$

	C	A	G	T	A	
C	0	-10.5	-11	-11.5	-12	-12.5
	$-\infty$	-21	-9.5	-10	.	.
	$-\infty$	.	.	.	.	.
	$-\infty$	.	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	-21.5	-22	.
T	-11	.	.	.	.

Vertical gaps only

- Back to M
- $\text{match}(i, j) = -4$ 
  - $M(i-1, j-1) = -\infty$
  - $X(i, j) = -10$
  - $Y(i, j) = -22$

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	-9.5	-14	.
T	$-\infty$	.	.	.	.

Matches, Mismatches only

$$M[i, j] = \text{match}(i, j) + \max \begin{cases} M(i-1, j-1) \\ X(i, j) \\ Y(i, j) \end{cases}$$

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	-21	-9.5	-10	.
T	$-\infty$	.	.	.	.

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	-21.5	-22	.
T	-11	.	.	.	.

Vertical gaps only

- I did a few more by hand:

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	$-\infty$	1	-13.5	-14	-15
T	$-\infty$	-13.5	2	-12.5	-13
	$-\infty$	-14	-12.5	-2	-11.5
	$-\infty$	-17			

Matches, Mismatches only

	C	A	G	T	A
C	0	-10.5	-11	-11.5	-12
A	$-\infty$	-21	-9.5	-10	-10.5
T	$-\infty$	-21.5	-20	-8.5	-9
	$-\infty$	-22	-20.5	-19	-12.5
	$-\infty$	-13			

Horizontal gaps only

	C	A	G	T	A
C	0	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-10.5	-21	-21.5	-22	-22.5
T	-11	-9.5	-20	-20.5	-21
	-11.5	-10	-8.5	-19	-19.5

Vertical gaps only

# So far:

- We have seen how we can do scoring for affine gap penalties
- Next time we will see how to do backtracking for affine gap penalties
- Also, next time we will have a question and answer session on Project 2