14.2.2
Edit distance as alignment

An Example

Example 14.4.

o		**c**	**u**	**r**	**r**	**a**	**n**	**c**	**e**
o	**c**	**c**	**u**	**r**	**r**	**e**	**n**	**c**	**e**

Cost $= \delta + \alpha_{ae}$

Alternative:

o		**c**	**u**	**r**	**r**		**a**	**n**	**c**	**e**
o	**c**	**c**	**u**	**r**	**r**	**e**		**n**	**c**	**e**

Cost $= 3\delta$

Or a really stupid solution (delete string, insert other string):

o	**c**	**u**	**r**	**r**	**a**	**n**	**c**	**e**											
									o	**c**	**c**	**u**	**r**	**r**	**e**	**n**	**c**	**e**

Cost $= 19\delta$.

What is the edit distance between...

What is the minimum edit distance for the following two strings, if insertion/deletion/change of a single character cost **1** unit?

374

473

- 1
- 2
- 3
- 4
- 5

What is the edit distance between...

What is the minimum edit distance for the following two strings, if insertion/deletion/change of a single character cost **1** unit?

373

473

- 1
- 2
- 3
- 4
- 5

What is the edit distance between...

What is the minimum edit distance for the following two strings, if insertion/deletion/change of a single character cost **1** unit?

37

473

- 1
- 2
- 3
- 4
- 5

Sequence Alignment

Input Given two words X and Y, and gap penalty δ and mismatch costs α_{pq}

Goal Find alignment of minimum cost

Sequence Alignment in Practice

1. Typically the DNA sequences that are aligned are about $\mathbf{10^5}$ letters long!
2. So about $\mathbf{10^{10}}$ operations and $\mathbf{10^{10}}$ bytes needed
3. The killer is the 10GB storage
4. Can we reduce space requirements?

THE END

...

(for now)