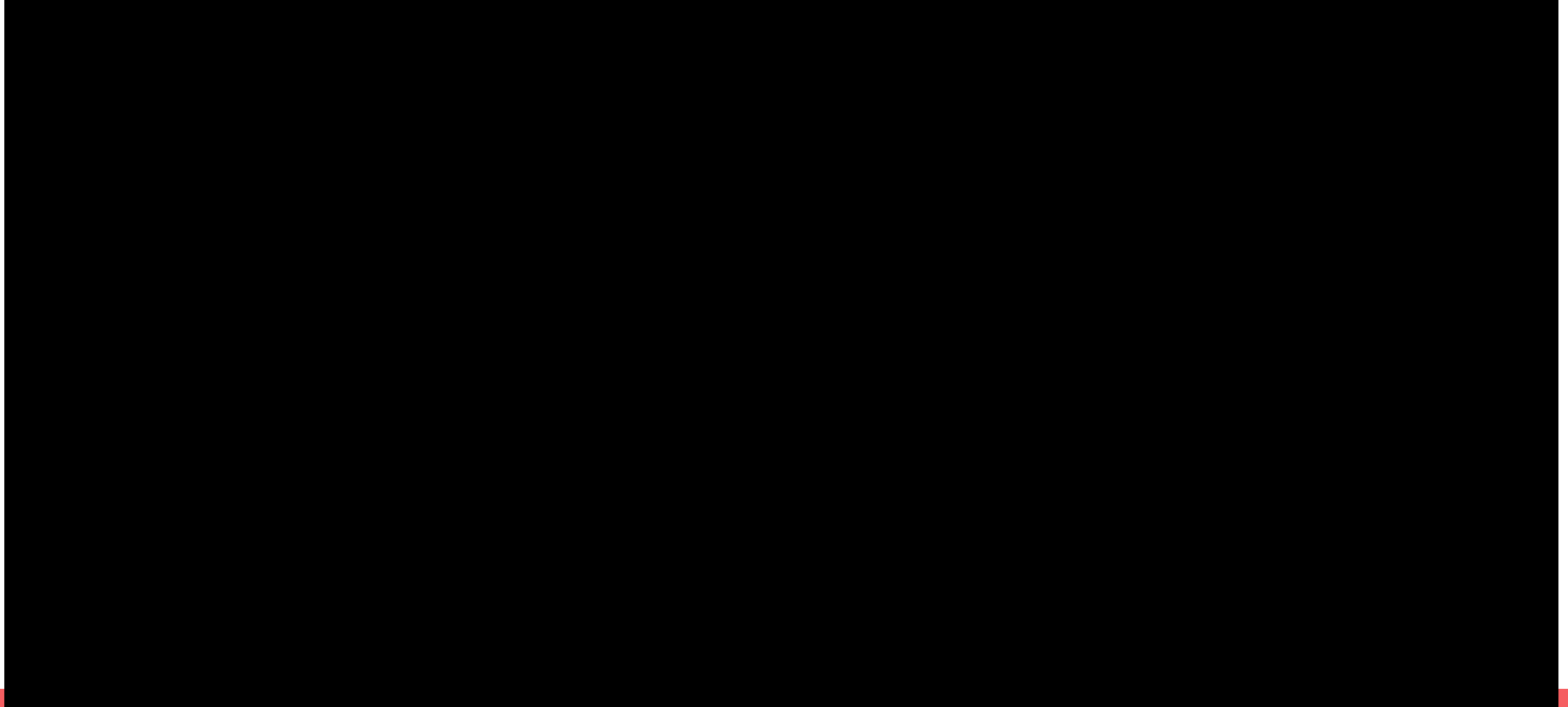


Global Alignment with Affine Gap Penalties

Jocelyn Hansson

Constant vs. Affine Gap Penalties



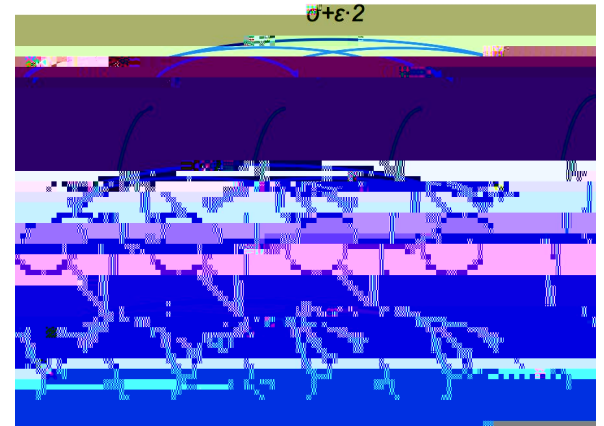
Constant vs Affine Gap Penalties

ATGGAGGGCGAGCTCGCAGCCAACCTGGAGCACCCGAGGCAGTCAACTCCAG
CGCCGCGCCGCGGGGGCAGAGGGCAACTGCACCCGCGGACCCCGCGCG
CGCAACGAGGCCCTGGCCGCGTGGAGGTGGCGGTGCTCTGTCTCATCT
GTTTCTGGCGCTGAGCGGGAATGCGTGTGTGCTGGCGCTGCGCACCA
CGCGCCACAAGCACTCGCCGCTCTTCTTTCATGAAGCACTGAGCATCG
CCGACCTGGTGGGCACTTCCAGGTGCTGCCGAGTTGCTGTGGGAC
ATCACCTCCGCTTCTACGGGCCGACCTGCTGTGCCGCTGGTCAAGTAC
TTGCAGGTGGTGGGCATGTTCCGCTCCACCTACCTGCTGCTACTCATGTCC
CTGGACCGCTGCTGGCCATCGCCAGCCGCTGCGCTCGCTGCGCCGCGG
CAGGACCCGCTGGCAGTGTCTGCCAGTGGCTGGCTGCGCTGGTGGCCA
GCGCGCCGAGGTGCACATCTTCTCTGCGCGAGGTGGCTGACGGCGCT
TTCGACTGCTGGCCGCTTTCATCCAGCCCTGGGGACCCAAGGCCATAC
ACGTGGATCACGCTAGCTGTCTACATCGTCCGCTCATCGTCTGCTGCTG
TGCTATGGCCTTATCAGCTCAAGATCTGGCAGAATTTGGCGCTCAAGACCG
CTGCAGCTGCGCGCGCAGAGGGCCAGAGGGCGCGGCGCTGGCGATGG
GGGGCGCATGGCCCTGGCAGTGTGACGACGCTCAAGCTCATCTCCAAGG
CCAAGATCCGACGGTCAAGATGACTTTCATCATCGTGTGGCCTTTCATCGT
GTGCTGAGCCCTTCTTCTTCTGTCAGATGTGGAGCGTCTGGGATGCCAA
CGCGCCCAAGGAAGTAGCCAGGGCTGGGAGACCCAGGAGGAGGGAGCC
TGGTGGCTGGGGGACGCCCTTATCTTGTGCTCAGAATGTCCAGGGGCT
TTGGACTTCTGGGGGATAAGAGGGTTTGAATCCACAGAGTCACTGTT
CCATCATCCCTGGCCAAAGTACTTGGGCAATTAACCTCCCTGAGCCTCCA
CTTCTCATCTGTGAGGTGGCAATAAGGATAAAAGTACCAACTGTCACTGGG
CATAGGGGCGAGCCATGAGAAAAATGACAGTAAAGTCTTTAGCACAGTCC
GGCTGCATATGGGCTGCATGGTTTACTGCGGTGGTGAACAGGTTCAAGGG
ACTCCATCTGCTTCCACGTTAGGAGGAGTGTAGGAGGAGATGG
CAGAGGTAAAGTTGGAAACCCCTGGCCAGGCTCAAACTCCTTCAAGTTAAAGT
TCACATTAAGCTGAACCTTCCACTTTAAGTTCAAGAAATCCATTTGAAGCCAA
AAGTCTGGTTTGGACAAGGACAGCCTGACAGGGGTGGGCGATTGTCCAGC
CAAGCCCTAGTAGTTGAAGGGGTGCTGGGGGCGAGGATTCAAGGG
AGAGGTGAAGACAAAATCCCGAACCCCTCATCGAGTGAAGGATGAGGCTC
CCATGCAGAGAGGTGAGTACTGCAAAAAGCGGTGGCAGGCTTAGGGAGG
AGAGTCCCAGGACTGAGACTTCACTTCTACTACTTGGGGA.....

ATGGAGGGCGAGCTCGCAGCCAACCTGGAGCACCCGAGGCAGTCAACTCCA
GCGCCGCGCGCCGGGGGCG-----C---G-A-G-GG-C-----AAC-----TG-
C-----A---C---C---G-C-C---G---G---A-C-----C---C---CGC-G-GC-
---GC---A-A-C---G-----A-G---GC-CCTG-GC---GC-G-C-G-TGG---AG-
---G-TG---GC-G---GT-GCT-----C-T-----G-T-CT-C---A-T-C---C---TG-T---
T-T-C---TGG---C-G-C-T-----G---A---G-C---GG-G-----A---A---
--TGCG-T-G-TG-----TG-C-TGCT-G---G---C-G-CTGC-G---C-A-C-
C---A---C-----GCGC---C-AC-----A---GC-----AC---
---T-C---G---C-GC---CT-CT---TC-T---TC-T-T-C-----ATG---A-AG
C---A---C---C---T-G---A-G-C---A-T-CG-C-C-GA-C-C-----TGG-
-TGG-----TGGC---A-GT-C-----T---TC-C---A-GG---T-GC---T-G---
C---C---G-C---A-GT-TGCTG-----T-----G-G-GA---C---AT-C-A-
C-C---T-----T---CC-G-----C-T-T-CT---ACG-----G-GCC-C
-GA---CC---T---G-CT---G---T-G-----CC---G-TG-TG-----G-
---T-C-A---A-----G---T-ACTT-----G---CA---GG-T---G-----GT-
-GGGACT-G---T---T-C-G---C---C---TCC---A---C-C-T---A
C-C-TG-----C---T---G---CT-A-CT-C---A-TG-T-----C-----
---CC-TGG-A-C-C-----G---C---T-GC---CT-----G---G---CC
AT-----C---T---G---C---CAGCC-GC---TG-CG---C-----T-----
---C-G---C---T---G---C---C---G---CC---G-----C-C---GC
A---C-G---G---A-C-----C---G-C---CT-----GG-----
---C---AG---T---G---CT-----C-G---C---C---A-----C
-G-TG-G-C---T-----C---G---G-C-TG-CC---T-GG-----T-G-G---
C-C-A-G---C-G---C-G-C-C---G-C-----A-GGTGC-----A-
-CA-----T-C-T-TC-T-C--T-C-----T---G---C---G---AGG-T
G---GCTG-----A-----C---G-G-C-G--T-C-----TTC---G---
-A---C---T---GCT-----G---G---GC---G---GT---C-TTCA---T-
C-C-A---G---C---C-C-TG---GGGA-----C-CCA-----A-G-G-C---
CT-----A---CA---TC-A-C---G---T---G---G-A-T---C-A---C---G-C
-T---A---G---C-TGT---C---T---ACA-T-C-G---T---G---CC---G-G---
---T-CA---TC---G---T-G-----GT-C---G-C---T---G---C-C-TG---
-----C---TA-T-G-G-C-CT---TA---T---C-A---G---C-T---T---C-AA
GA-----TC---T---G-G---C---AG-A-A-T-T-----TG---C-G-----

ATGGAGGGCGAGCTCGCAGCCAACCTGGAGCACCCGAGGCAGTCAACTCCA
CTCCAGCGCCGCGCGCCGGGGGCGGAGGGCAACTGCACCCGCG

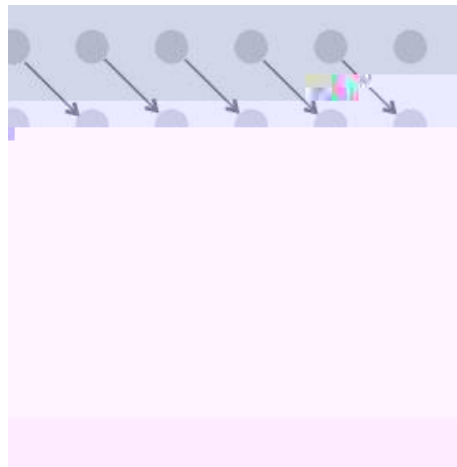
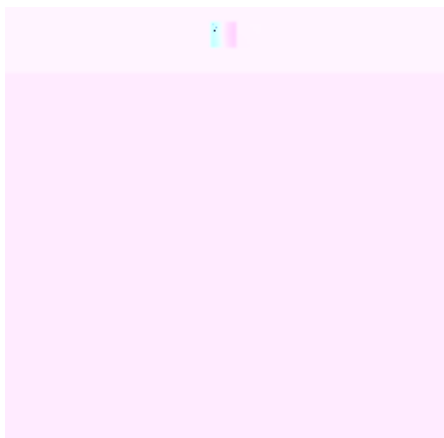
Long edges in graph



$$\text{lower}_{ij} = \max \begin{cases} \text{gap extension} \\ \text{Initialize gap} \end{cases} \quad \text{lower}_{i-1,j} - \text{middle}_{i-1,j}$$

ckef

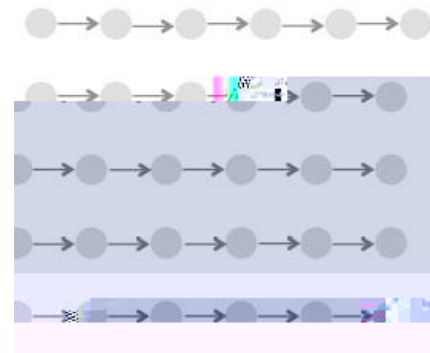
(Insertions on string 2)



aidd'e

(matches/mismatches)

ending gap on string 1
 match/mismatch
 ending gap on string 2



iddef

(Insertions on string 1)

gap extension upper_{i,j-1} -
 Initialize gap middle_{i,j-1} -

upper_{ij}
 middle_{i-1,j-1} + score
 lower_{ij}



Making 3 back track tables

$$\text{upper}_{ij} = \max \left\{ \begin{array}{ll} \text{gap extension} & \text{upper} \rightarrow \text{upper} \\ \text{Initialize gap} & \text{middle} \rightarrow \text{upper} \end{array} \right.$$

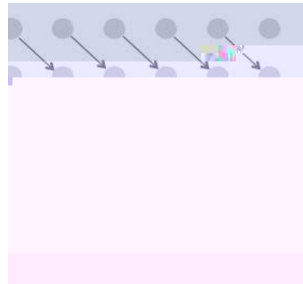
$$\text{middle}_{ij} = \max \left\{ \begin{array}{ll} \text{ending gap on string 1} & \text{upper} \rightarrow \text{middle} \\ \text{match/mismatch} & \text{middle} \rightarrow \text{middle} \\ \text{ending gap on string 2} & \text{lower} \rightarrow \text{middle} \end{array} \right.$$

$$\text{lower}_{ij} = \max \left\{ \begin{array}{ll} \text{gap extension} & \text{lower} \rightarrow \text{lower} \\ \text{Initialize gap} & \text{middle} \rightarrow \text{lower} \end{array} \right.$$

Back tracking

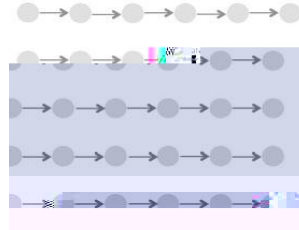
`ckef

(Insertions on string 2)



aidd`e

(matches/mismatches)



i ddef

(Insertions on string 1)

Data



Results

