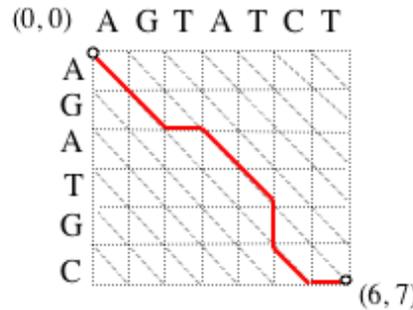


# 4. Comparaison de séquences

- Comment prédire les fonctions des gènes/protéines ?
- Évolution et similarité de séquences
- Quantifier la similarité de deux séquences
- L'alignement de séquences devient un problème d'optimisation
- Un alignement de séquences vu comme un chemin dans une grille
- **Si un chemin est optimal, tous ses chemins partiels sont optimaux**
- Coûts et alignement
- Un algorithme récursif
- Éviter la récursivité : une version itérative
- Cet algorithme est-il efficace ?

# Une preuve par l'absurde

- Si un chemin de longueur  $L$  est optimal, alors le chemin de longueur  $L-1$  l'est également
- Faire l'hypothèse contraire, c'est-à-dire que le chemin de longueur  $L-1$  n'est pas optimal, aboutit à une contradiction



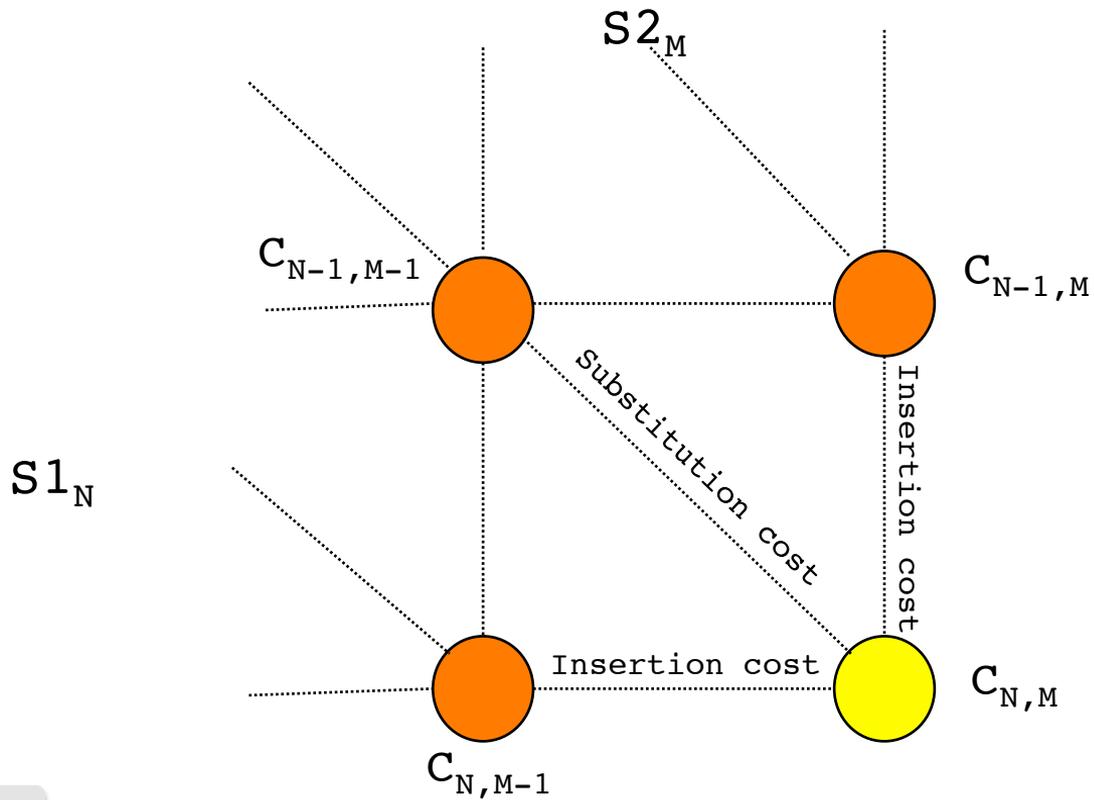
(a) **A G - A T G C -**  
**A G T A T - C T**

## Et maintenant par récurrence...

- Si un chemin de longueur  $L$  est optimal, alors le chemin de longueur  $L-1$  l'est également
- Faire l'hypothèse contraire, c'est-à-dire que le chemin de longueur  $L-1$  n'est pas optimal, aboutit à une contradiction
- Tous les chemins partiels de longueur  $1, 2, 3, \dots, L-2, L-1$  sont optimaux

# Vers un algorithme

- Il n'y a que trois chemins de longueur 1 qui conduisent au dernier nœud de la grille



# Vers un algorithme

- Il n'y a que trois chemins de longueur 1 qui conduisent au dernier nœud de la grille
- Supposons que les chemins (de longueur  $L-1$ ) qui aboutissent à ces 3 nœuds soient optimaux
- Alors, il est facile de déterminer le dernier chemin de longueur 1 et le coût optimal à affecter au dernier nœud

# Calcul du dernier chemin de longueur 1 et du coût à attacher au dernier nœud

