

Elenco domande



ARGORIENTO

Ricostruzione evoluzione

→ def. filogenesi perfette

→ esempio di molice in input binarie e albero ricostruito dalle molice

Perché l'albero filogenetico perfetto è un modello usato per l'evoluzione tumorale?

■ RicorTruzione operativi de
genotipi

def. genotipo

def. esplorato

■ Ricostruzione con filogenesi
perfette delle esplorazioni de
genotipi \Rightarrow esempi

■ Assemblaggio di operativi
de long reads

MEC (MINIMUM ERROR
CORRECTION)

def. Bloom filter

def. idee di base del
MIN HASH

def. Distanza di Jaccard

■ Per esempio di come
due gigliie viene utilizzate
per il confronto di due
sequenze

quale score usare per avere
una LCS (longest common
subsequence)
con le gigliie

PROBLEMA

Alignment - free sequence to graph

esempio

INPUT: graph DAG

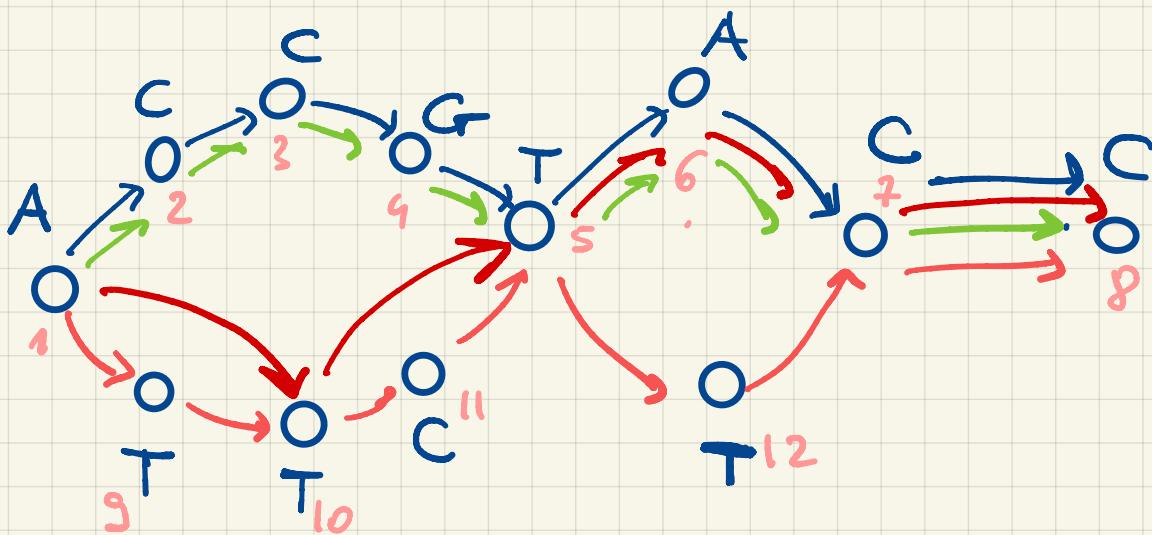


tabelle hash dei K-mer

per nodi $K = 3$

date le sequenze

$s = \text{CGT TCC}$

1 Acc ATT
 2 CCG '(■, ■)
 3 CGT (■, ■)
 4 GTA (■, ■)
 5 TAC / TTC

6 ACC

7 -

8 -

9 TTC (■, ■)

10 TCT TTA (■)
 11 CTT
 12 TTC

K-mer unici!

■ vediamo come K-mer di colori
opporsi coprono

$$s = \text{GGTTTCC}$$

4↑ (■, ■) 3↓ (■)

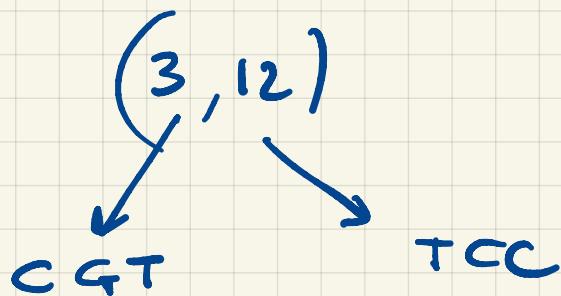
■ calcolare le posizioni di

s_i, j dove iniziava

in i e inizia in j due
K-mer UNICI che hanno nel loro

insieme nessun colore condizion

DANE i VERTICI Associazioni
elle copre di K-MEANS



- Spaced Seeds

INPUT: r , R

dove r differisce di:
UNO SNP o errore

de one sottoshingle di
 R , cioè $\exists r'$
sottoshingle di r tale
 $\text{Hamming}(r, r') = 1$

OUTPUT : dove le posizioni

di r che ha le SNP o
enone usando spaced seeds!

esempio $R = \text{A A A A C C G T T C T}$

$r = \text{C A G T T}$

$r' = \text{C C G T T}$

$$H(r, r') = 1$$

uso 1- approximate spaced seeds

de R con $K=3$ (cioè di lunghezza 3)

Hash * → 1

* AA

* AC

* CC

* CG

* GT

* TT

* TG

hash * → 2

A * A

A * C

C * G

G * T

T * G

hash * → 3

TT *

GT *

CG *

CG *

AC *

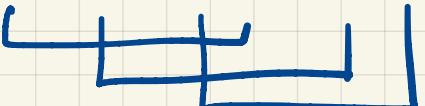
AA *

Prendo r e quindi

PRO w è vedere se $r[1] \rightarrow SNP/$
mettendoci $*$:

$r = * A G T T$ (testo ; k-mer
speciedi seeds con
 ℓ' hashing!)

testo r con $*$ in 2

$r = C * G \bar{T} T$ k-mer

($C * G,$
 $* G T,$
 $G T T'$)