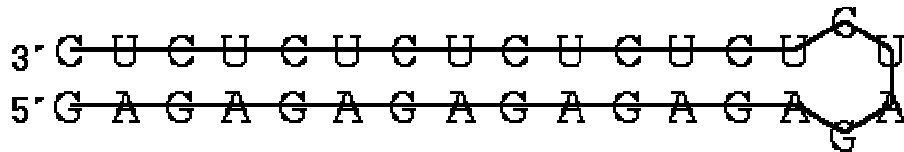


Folding binárných reťazcov

(Peter Orosi)

Binárny folding

- analógia s foldingom RNA



- možné väzby: GC, AU, GU
- stabilita sekundárnej štruktúry = množstvo vydanéj energie pri vytváraní foldingu
- cieľ väčšiny algoritmov = nájdenie konfigurácie párov s najnižšou voľnou energiou (najstabilnejší folding)

Binárny folding

- binárny folding = nahradíme (C, G, A, U) za (0, 1)
- možné väzby: medzi 0 a 1
- obmedzenia:
 - minimálna vzdialenosť medzi naviazanými elementami = 2
 - každý element max. jedna väzba (planárnosť grafu)
 - maximálny počet väzieb (najstabilnejšia štruktúra)

Možné reprezentácie foldingu

- vymenovanie spárovaných dvojíc a nespárovaných elementov

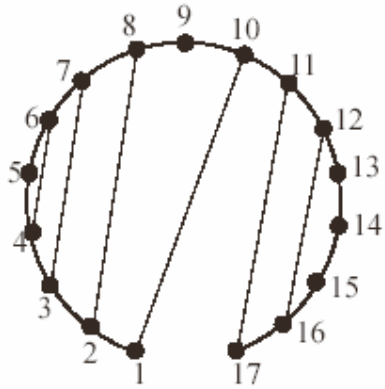
(1, 10), (2, 8), (3, 7), (4, 6), (11, 17), (12, 16), 5, 9, 13, 14, 15

- zátvorková reprezentácia

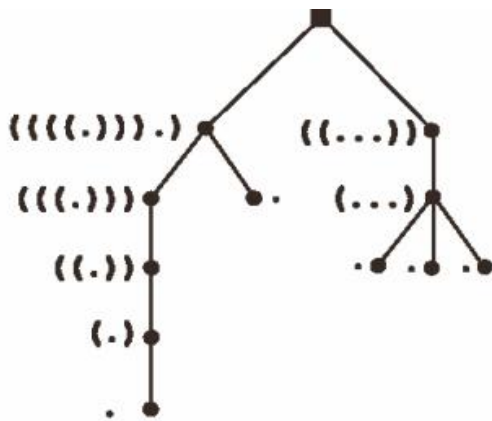
((((.))) .) ((. . .))

Možné reprezentácie foldingu

- kruhová reprezentácia



- binárnym stromom



Konštrukcia foldingu

- Kombinatorický prístup ~ exponenciálna zložitosť
- Dynamické programovanie
 - $S_{ii} = \text{'.'}$, $S_{i,i+1} = \text{'..'}$
 - rekurzia, pre $(j - i) > 2$

$$S_{i,j} = \max_{i \leq k \leq j-1} S_{ik} \oplus S_{k+1,j}$$

- Operácia sčítania reťazcov:
 - $..(\dots). + ..(\dots). = ..(\dots)..(\dots)$ – ak krajné elementy sú kompatibilné
 - $..(\dots). + ..(\dots). = ..(\dots)..(\dots)$ – ak krajné elementy nie sú kompatibilné
- Max = vybratie podreťazca s najväčším počtom väzieb (zátvoriek)
- Výsledok v S_{1n}

Algoritmus

```
for i:=1 to n S[i,i]:= '.';
for i:=1 to n-1 do S[i,i+1]:= '..';
for d:=2 to n-1 do
for i:=1 to n-d do
begin j:=d+i;
      S[i,j]:=max{i≤k≤j-1, S[i,k]⊕S[k+1,j]};
end;
folding:=S[1,n];
```

- Zložitosť ~ $O(n^3)$

Súvis foldingu s darwinizmom

- Genotyp (binárny reťazec)
- Fenotyp (telo - je určený foldingom binárneho reťazca)
- Fitness – možno vyjadriť ako funkciu fenotypu
- Neutrálne etapy, neutrálna stázia, neutrálna mutácia (bez zmeny fenotypu)