Universität Greifswald
Institute für Mathematik and Informatik
Lecturer: Marc Hellmuth Tutor: Nikolai Nøjgaard

7. Exercise "Bioinformatics", SS 17

Aufgabe 1: (2+3= 5 Credits)
Consider the RNA sequence

$$s = \texttt{GGGCACAUGGGGCAGUGCAGCCACUGAGCC}$$

with secondary structure

$$S = \{(1,30),(2,29),(4,17),(5,16),(6,15),(8,14),(9,13),(18,26),(19,25),(20,24)\}$$

and assume $\Theta = 0$.

(a) Draw the structure in dot-bracket notation and as another graphical representation of your choice.

(b) Prove or disprove: $S \cup bp_i$ is a secondary structure for s with $bp_1 = \{(10,22)\}$, $bp_2 = \{(10,12)\}$, $bp_3 = \{(10,13)\}$.

Aufgabe 2: (5+5=10 Credits)
Let $S(n)$ denote the number of possible secondary structures of size n and $S(n,k)$ denote the number of possible secondary structures of size n that have exactly k basepairs.

(a) Show that for all $n \geq 2$ holds:

$$S(n) \geq 2^{n-2}$$

(b) Let $S(n,0) = 1$ for all n and $S(n,k) = 0$ for $k \geq n/2$. Show that for all $n \geq 2$ holds:

$$S(n+1,k+1) = S(n,k+1) + \sum_{j=1}^{n-1}\left[\sum_{i=0}^{k} S(j-1,i)S(n-j,k-i)\right].$$

Aufgabe 3: (2.5+2.5 =5 Credits)
Let $\mathcal{A} = \{\texttt{A,C,G,U}\}$, $\mathcal{B} = \{\texttt{AU, UA, GC, CG, GU, UG}\} \cup \{\texttt{AA}\}$ and $S_1, \ldots, S_k$ secondary structures of size n ($\Theta = 0$). Prove or disprove:

(a) If $G(S_1, \ldots, S_k)$ is bipartite then there is a sequence $s \in \mathcal{A}^n$ realizing all secondary structures $S_1, \ldots, S_k$.

(b) If there is a sequence $s \in \mathcal{A}^n$ realizing all secondary structures $S_1, \ldots, S_k$ then $G(S_1, \ldots, S_k)$ is bipartite.

Deadline: Tuesday - May 23, 2017