

CSE 417: Algorithms and Computational Complexity
 Recurrence for predicting RNA secondary structure
 From a paper by Nussinov and Jacobsen, 1980

Let $R = R[1]R[2]\dots R[n]$ be an RNA sequence, i.e., a string over the alphabet $\{A,C,G,U\}$. Define $E_{i,j}$ to be the maximum number of base pairs in any fold of $R[i], R[i+1], \dots, R[j]$. Here, any of the pairs $\{A,U\}$, $\{C,G\}$, and $\{G,U\}$ can form a base pair. Then

$$E_{i,j} = \begin{cases} 0 & , \text{if } j - i < 4 \\ \max \left(\begin{array}{l} E_{i+1,j-1} + p(R[i], R[j]), \\ \max_k E_{i,k-1} + E_{k,j} \\ i < k \leq j \end{array} \right) & , \text{if } j - i \geq 4 \end{cases}$$

where

$$p(s, t) = \begin{cases} 1 & , \text{if } s, t \text{ can form a base pair} \\ 0 & , \text{otherwise} \end{cases} .$$