Beispiel zur Befüllung der Matrix

Sei die gegebene RNA-Sequenz UUUCAGUAGCA
Diese besteht aus 11 Basen, also mit Index von 0 bis 10. Wir erzeugen eine 11x11 Matrix.
In dieser Matrix soll der Eintrag (i,j) die maximal Anzahl von Wasserstoffbrücken für die Teilsequenz von Base mit Index i bis zur Base mit Index j enthalten (für 0 <= i, j <= 10)

	j	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U											
1	U											
2	U											
3	C											
4	A											
5	G											
6	U											
7	A											
8	G											
9	C											
#	A											

i = Startposition
j = Endposition

M(2,6) enthält Anzahl an WBB für Sequenz
UCAGU

M(1,8) enthält Anzahl an WBB für Sequenz
UUCAGUAG

1. Startposition liegt immer vor Endposition
Da i die Startposition angibt ist der untere Teil der Matrix mit den Werten 0 zu belegen.
Hier würde die Endposition VOR der Startposition liegen.

	j	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U											
1	U	0										
2	U	0	0									
3	C	0	0	0								
4	A	0	0	0	0							
5	G	0	0	0	0	0						
6	U	0	0	0	0	0	0					
7	A	0	0	0	0	0	0	0				
8	G	0	0	0	0	0	0	0	0			
9	C	0	0	0	0	0	0	0	0	0		
#	A	0	0	0	0	0	0	0	0	0	0	

2. Primary Proximity Constraint

Da mindestens 3 Basen zwischen der beiden Basen, die eine Verbindung bilden,
liegen sollen, können Einträge für Sequenzen mit Länge <= 4 mit dem Wert 0 befüllt werden.

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0							
1	U	0	0	0	0	0						
2	U	0	0	0	0	0	0					
3	C	0	0	0	0	0	0	0				
4	A	0	0	0	0	0	0	0	0			
5	G	0	0	0	0	0	0	0	0	0		
6	U	0	0	0	0	0	0	0	0	0	0	
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0

3. Diagonale Matrixbefüllung

Wir beginnen nun an Position (0,4). Das ist eine Sequenz der Länge 5.
Es können sich also, wenn überhaupt, nur Anfangs- und Endbase paaren.
Base 0 (= U) mit Base 4 (= A) können 2 Wasserstoffbrücken ausbilden.
Somit bekommt das Feld M(0,4) den Wert 2. Auf diese Weise wird diagonal gefüllt.

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0	2						
1	U	0	0	0	0	0	2					
2	U	0	0	0	0	0	0	0				
3	C	0	0	0	0	0	0	0	0			
4	A	0	0	0	0	0	0	0	0	0		
5	G	0	0	0	0	0	0	0	0	0	3	
6	U	0	0	0	0	0	0	0	0	0	0	2
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0

Teilsequenz (0,4) : UUUCA
(U,A) -> 2 Wasserstoffbrücken

M(1, 5) = 2 (U,G)
M(2, 6) = 0 (U,U)
M(3, 7) = 0 (C,A)
M(4, 8) = 0 (A,G)
M(5, 9) = 3 (G,C)
M(6,10) = 2 (U,A)

4. Fallunterscheidung bei Längeren Sequenzen

Als nächstes wird die Länge um 1 vergrößert. Wir beginnen bei Eintrag (0,5).
1. Fall: Base 0 bindet mit Base 5, U mit G, das ergibt 2 Wasserstoffbrücken.
Hinzu kommt die maximale Anzahl für den inneren Teil der zu betrachtenden Sequenz (UUCA). Dieses Maximum wurde bereits berechnet und befindet sich im Eintrag M(1,4) = 0.

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0	2						
1	U	0	0	0	0	0	2					
2	U	0	0	0	0	0	0	0				
3	C	0	0	0	0	0	0	0	0			
4	A	0	0	0	0	0	0	0	0	0		
5	G	0	0	0	0	0	0	0	0	0	3	
6	U	0	0	0	0	0	0	0	0	0	0	2
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0

1.Fall: Verbindung von erster mit letzter Base

Sequenz M(0,5) : U UUCA G

(U,G) + (UUCA)
2 + M(1,4)

Für den zweiten Fall teilen wir die Sequenz in alle möglichen Teilsequenzen und berechnen die Summe. Die maximale Anzahl der Wasserstoffbrücken für jede Teilsequenz kann aus der Matrix entnommen werden.

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0	2						
1	U	0	0	0	0	0	2					
2	U	0	0	0	0	0	0	0				
3	C	0	0	0	0	0	0	0	0			
4	A	0	0	0	0	0	0	0	0	0		
5	G	0	0	0	0	0	0	0	0	0	3	
6	U	0	0	0	0	0	0	0	0	0	0	2
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0

2.Fall: Summe der Teilsequenzen

Sequenz M(0,5) : UUUCAG

U + UUCAG	M(0,0) + M(1,5)	0+2 = 2
UU + UCAG	M(0,1) + M(2,5)	0+0 = 0
UUU + CAG	M(0,2) + M(3,5)	0+0 = 0
UUUC + AG	M(0,3) + M(4,5)	0+0 = 0
UUUCA + G	M(0,4) + M(5,5)	2+0 = 2

Das Maximum sind 2 Wasserstoffbrücken

Insgesamt ergibts sich sowohl aus dem 1. und 2. Fall ein maximaler Wert von 2 für M(0,5).

Als nächster Eintrag der Länge 6 kommt M(1,6) an die Reihe.

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0	2	2					
1	U	0	0	0	0	0	2					
2	U	0	0	0	0	0	0	0				
3	C	0	0	0	0	0	0	0	0			
4	A	0	0	0	0	0	0	0	0	0		
5	G	0	0	0	0	0	0	0	0	0	3	
6	U	0	0	0	0	0	0	0	0	0	0	2
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0

1. Fall: Erste mit Letzter Base

(U,U) + (UCAG) (U,U)+ M(2,5) 0+0 = 0

2. Fall: Teilsequenzen

U + UCAGU	M(1,1) + M(2,6)	0+0 = 0
UU + CAGU	M(1,2) + M(3,6)	0+0 = 0
UUC + AGU	M(1,3) + M(4,6)	0+0 = 0
UUCA + GU	M(1,4) + M(5,6)	0+0 = 0
UUCAG + U	M(1,5) + M(6,6)	2+0 = 2

Also M(1,6) = 2

Auf diese Weise berechnen wir Diagonal für Sequenzen der Länge 6:

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0	2	2					
1	U	0	0	0	0	0	2	2				
2	U	0	0	0	0	0	0	0	2			
3	C	0	0	0	0	0	0	0	0	3		
4	A	0	0	0	0	0	0	0	0	0	3	
5	G	0	0	0	0	0	0	0	0	0	3	3
6	U	0	0	0	0	0	0	0	0	0	0	2
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0

Nun wird die Sequenzlänge immer um 1 vergrößert, bis die Matrix diagonal bis zum Rand gefüllt ist. Als letztes wird der Wert M(0,10) eingetragen.

j		*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*#*
i		U	U	U	C	A	G	U	A	G	C	A
0	U	0	0	0	0	2	2	2	4	4	5	6
1	U	0	0	0	0	0	2	2	2	4	4	5
2	U	0	0	0	0	0	0	0	2	3	3	5
3	C	0	0	0	0	0	0	0	0	3	3	3
4	A	0	0	0	0	0	0	0	0	0	3	3
5	G	0	0	0	0	0	0	0	0	0	3	3
6	U	0	0	0	0	0	0	0	0	0	0	2
7	A	0	0	0	0	0	0	0	0	0	0	0
8	G	0	0	0	0	0	0	0	0	0	0	0
9	C	0	0	0	0	0	0	0	0	0	0	0
#	A	0	0	0	0	0	0	0	0	0	0	0