

Genetische Algorithmen
 Übungen zur Kodierung von Lösungskandidaten

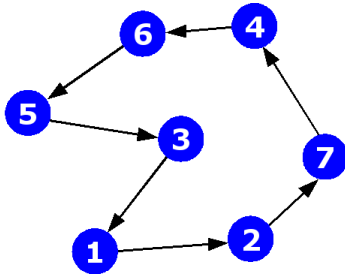


Abb. 1 – Eine Rundreise

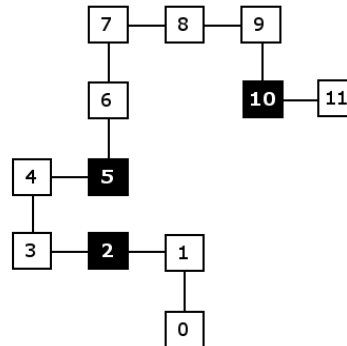


Abb. 2 – Eine Konformation im 2d HP-Modell

Aufgabe 1

* Berechnen Sie die Hamming-Distanz zwischen den Zahlen 2891 und 3927.

Aufgabe 2

** Wie gross ist die Wahrscheinlichkeit, dass die **fett unterstrichen** markierte Gruppe von Genen durch Crossover aufgebrochen wird?

1010010101000101101101010100101011001011110111

Aufgabe 3

** Wie wird die Rundreise in Abbildung 1 mit Pfadkodierung, Kantenkodierung bzw. mit Ordinalkodierung dargestellt?

Aufgabe 4

Betrachten Sie die Konformation in Abbildung 2:

* Wie wird diese Sequenz in absoluten Gitterkoordinaten (X_i, Y_i) dargestellt?

** Wie wird diese Sequenz in relativen Gitterkoordinaten (Oben, Unten, Links, Rechts) dargestellt?

** Wie wird diese Sequenz in relativen Konformationskoordinaten (Geradeaus, Links, Rechts) dargestellt?

*** Wie kann man bei diesem Problem Epistasie messen ?

Die Antwort sollte es Ihnen ermöglichen, folgende Frage zu beantworten: Eine Konformation A wird auf zwei verschiedene Arten mutiert und man erhält die Konformationen B und C. Welche dieser beiden Konformationen ist nun ähnlicher zu A?

** Welche Kodierung weist die geringste Epistasie auf? Betrachten Sie hierzu mögliche Mutationen an dem Element Nr. 1.

*** Überlegen Sie sich einen effizienten Algorithmus um eine Konformation in einer relativen Kodierung auf Überlappung zu prüfen.

*** Überlegen Sie sich eine mögliche Kodierung die sicherstellt, dass an jeder Gitterposition niemals mehr als ein Element der Sequenz platziert ist, und stellen Sie die obige Konformation in dieser Kodierung dar. Welche Probleme treten bei dieser Kodierung auf?

Lösungen

Aufgabe 1

2891 = 101101001011

3927 = 111101010111

Vier Bits unterscheiden sich, also beträgt die Hamming-Distanz 4.

Aufgabe 2

Es gibt 45 Stellen wo das Chromosom aufgebrochen werden kann, davon 32 innerhalb der markierten Gruppe von Genen, daher ist $p = 32/45 = 0,71111$ also etwa 71%

Aufgabe 3

Pfadbildung: (1 2 7 4 6 5 3)

Kantenkodierung: (2 7 1 6 3 5 4)

Ordinalkodierung mit der Annahme, dass die Referenz-Rundreise $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 7$ ist: (1 1 5 2 3 2 1)

Aufgabe 4

Annahme: Links unten ist (1,1)

Absolute Gitterkoordinaten: (3,1) (3,2) (2,2) (1,2) (1,3) (2,3) (2,4) (2,5) (3,5) (4,5) (4,4) (5,4)

Relative Gitterkoordinaten: (OLLOROORRUR)

Relative Konformationskoordinaten: (GLGRRLGRGRL)

Wie kann man Epistasie messen?

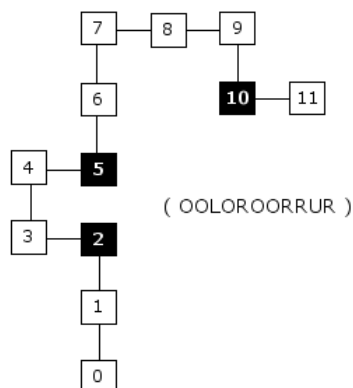
Beispielsweise durch die Anzahl der Elementpaare, deren relative Position zueinander durch die Mutation verändert wird, also mindestens 0 und höchstens $n*(n+1)/2$.

Relative Gitterkoordinaten:

(OLLOROORRUR) mutiert zu (ORLOROORRUR) und ist eine ungültige Konformation.

(OLLOROORRUR) mutiert zu (OULOROORRUR) und ist eine ungültige Konformation.

(OLLOROORRUR) mutiert zu (OOLOROORRUR) und sieht wie folgt aus:



Die Abstände zwischen den Elementpaaren sind wie folgt:

	0	1	2	3	4	5	6	7	8	9	10	11
0												
1	1,00											
2	1,41	1,00										
3	2,24	2,00	1,00									
4	2,83	2,24	1,41	1,00								
5	2,24	1,41	1,00	1,41	1,00							
6	3,16	2,24	2,00	2,24	1,41	1,00						
7	4,12	3,16	3,00	3,16	2,24	2,00	1,00					
8	4,00	3,00	3,16	3,61	2,83	2,24	1,41	1,00				
9	4,12	3,16	3,61	4,24	3,61	2,83	2,24	2,00	1,00			
10	3,16	2,24	2,83	3,61	3,16	2,24	2,00	2,24	1,41	1,00		
11	3,61	2,83	3,61	4,47	4,12	3,16	3,00	3,16	2,24	1,41	1,00	

OLLOROORRUR

	0	1	2	3	4	5	6	7	8	9	10	11
0												
1	1,00											
2	2,00	1,00										
3	2,24	1,41	1,00									
4	3,16	2,24	1,41	1,00								
5	3,00	2,00	1,00	1,41	1,00							
6	4,00	3,00	2,00	2,24	1,41	1,00						
7	5,00	4,00	3,00	3,16	2,24	2,00	1,00					
8	5,10	4,12	3,16	3,61	2,83	2,24	1,41	1,00				
9	5,39	4,47	3,61	4,24	3,61	2,83	2,24	2,00	1,00			
10	4,47	3,61	2,83	3,61	3,16	2,24	2,00	2,24	1,41	1,00		
11	5,00	4,24	3,61	4,47	4,12	3,16	3,00	3,16	2,24	1,41	1,00	

OOLOROORRUR

Insgesamt gibt es 17 Elementpaare, deren Distanzen sich bei der Mutation von (OLLOROORRUR) nach (OOLOROORRUR) geändert haben:

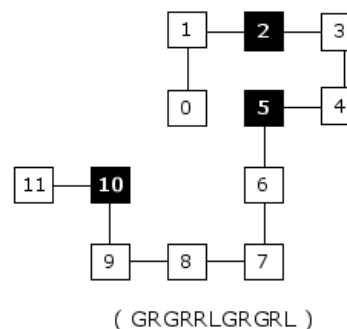
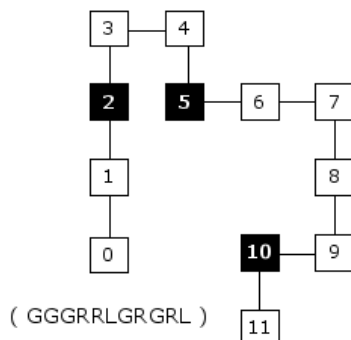
	0	1	2	3	4	5	6	7	8	9	10	11
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0,59	0	0	0	0	0	0	0	0	0	0	0
3	0	-0,59	0	0	0	0	0	0	0	0	0	0
4	0,33	0	0	0	0	0	0	0	0	0	0	0
5	0,76	0,59	0	0	0	0	0	0	0	0	0	0
6	0,84	0,76	0	0	0	0	0	0	0	0	0	0
7	0,88	0,84	0	0	0	0	0	0	0	0	0	0
8	1,1	1,12	0	0	0	0	0	0	0	0	0	0
9	1,26	1,31	0	0	0	0	0	0	0	0	0	0
10	1,31	1,37	0	0	0	0	0	0	0	0	0	0
11	1,39	1,41	0	0	0	0	0	0	0	0	0	0

17 geänderte Distanzen

Relative Konformationskoordinaten:

(GLGRRLGRGRL) mutiert nach (GGRRLLGRGRL)

(GLGRRLGRGRL) mutiert nach (GRGRLLGRGRL)



Die Abstände zwischen den Elementpaaren sind wie folgt:

	0	1	2	3	4	5	6	7	8	9	10	11
0												
1	1,00											
2	1,41	1,00										
3	2,24	2,00	1,00									
4	2,83	2,24	1,41	1,00								
5	2,24	1,41	1,00	1,41	1,00							
6	3,16	2,24	2,00	2,24	1,41	1,00						
7	4,12	3,16	3,00	3,16	2,24	2,00	1,00					
8	4,00	3,00	3,16	3,61	2,83	2,24	1,41	1,00				
9	4,12	3,16	3,61	4,24	3,61	2,83	2,24	2,00	1,00			
10	3,16	2,24	2,83	3,61	3,16	2,24	2,00	2,24	1,41	1,00		
11	3,61	2,83	3,61	4,47	4,12	3,16	3,00	3,16	2,24	1,41	1,00	

GLGRRLGRGRL

	0	1	2	3	4	5	6	7	8	9	10	11
0												
1	1,00											
2	2,00	1,00										
3	3,00	2,00	1,00									
4	3,16	2,24	1,41	1,00								
5	2,24	1,41	1,00	1,41	1,00							
6	2,83	2,24	2,00	2,24	1,41	1,00						
7	3,61	3,16	3,00	3,16	2,24	2,00	1,00					
8	3,16	3,00	3,16	3,61	2,83	2,24	1,41	1,00				
9	3,00	3,16	3,61	4,24	3,61	2,83	2,24	2,00	1,00			
10	2,00	2,24	2,83	3,61	3,16	2,24	2,00	2,24	1,41	1,00		
11	2,24	2,83	3,61	4,47	4,12	3,16	3,00	3,16	2,24	1,41	1,00	

GGGRRLGRGRL

	0	1	2	3	4	5	6	7	8	9	10	11
0												
1	1,00											
2	1,41	1,00										
3	2,24	2,00	1,00									
4	2,00	2,24	1,41	1,00								
5	1,00	1,41	1,00	1,41	1,00							
6	1,41	2,24	2,00	2,24	1,41	1,00						
7	2,24	3,16	3,00	3,16	2,24	2,00	1,00					
8	2,00	3,00	3,16	3,61	2,83	2,24	1,41	1,00				
9	2,24	3,16	3,61	4,24	3,61	2,83	2,24	2,00	1,00			
10	1,41	2,24	2,83	3,61	3,16	2,24	2,00	2,24	1,41	1,00		
11	2,24	2,83	3,61	4,47	4,12	3,16	3,00	3,16	2,24	1,41	1,00	

GRGRRLGRGRL

Insgesamt gibt es 9 Elementpaare, deren Distanzen sich bei der Mutation von (GLGRRLGRGRL) nach (GGGRRLGRGRL) geändert haben:

	0	1	2	3	4	5	6	7	8	9	10	11
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	-0,59	0	0	0	0	0	0	0	0	0	0	0
3	-0,76	0	0	0	0	0	0	0	0	0	0	0
4	-0,33	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0
6	0,33	0	0	0	0	0	0	0	0	0	0	0
7	0,52	0	0	0	0	0	0	0	0	0	0	0
8	0,84	0	0	0	0	0	0	0	0	0	0	0
9	1,12	0	0	0	0	0	0	0	0	0	0	0
10	1,16	0	0	0	0	0	0	0	0	0	0	0
11	1,37	0	0	0	0	0	0	0	0	0	0	0

9 veränderte Distanzen

Bei der Mutation von (GLGRRLGRGRL) nach (GRGRRLGRGRL) haben sogar nur 8 Elementpaare ihre Distanz zueinander geändert:

	0	1	2	3	4	5	6	7	8	9	10	11
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0,83	0	0	0	0	0	0	0	0	0	0	0
5	1,24	0	0	0	0	0	0	0	0	0	0	0
6	1,75	0	0	0	0	0	0	0	0	0	0	0
7	1,89	0	0	0	0	0	0	0	0	0	0	0
8	2	0	0	0	0	0	0	0	0	0	0	0
9	1,89	0	0	0	0	0	0	0	0	0	0	0
10	1,75	0	0	0	0	0	0	0	0	0	0	0
11	1,37	0	0	0	0	0	0	0	0	0	0	0

8 veränderte Distanzen

Zusammenfassung der Ergebnisse:

Mit relativen Gitterkoordinaten war nur eine Mutation möglich, dabei haben 17 Elementpaare ihre relative Position zueinander geändert.

Mit relativen Konformationskoordinaten waren zwei Mutationen möglich, dabei haben im Schnitt $(9 + 8) / 2 = 8.5$ Elementpaare ihre relative Position zueinander geändert.

In diesem Beispiel ist die Epistasie bei den relativen Konformationskoordinaten also geringer als bei den relativen Gitterkoordinaten.

Prüfung einer Konformation in relativen Koordinaten auf Überlappung:

Wandle die relativen Koordinaten in absolute Gitterkoordinaten um;

```

for (int i=0; i<n-2; i++)
    iPos = koordinateVon Element(i);
    for (int j=i+1; j<n-1; j++)
        jPos = koordinateVonElement(j);

        // vergleiche element i mit j
        if (iPos.equals(jPos)) konformationIstUngültig();

        // fast j skip
        j += iPos.Abstand(jPos) - 1;
    }
}

```

Kodierung, so dass an jeder Gitterposition niemals mehr als ein Element platziert ist:

Analog zu der Kantenkodierung bei der Rundreise, wobei hier die anstatt der Kantenummer eine Feldnummer verwendet wird. Annahme: Links unten ist die Position 1,1, der Code enthält die Symbole F = Frei, W = Weiß und S = Schwarz. Die Kodierung lautet:
 (FWWFFSSWWWWFFWFFFSWFFWF)

Problem Nr. 1: Die Länge eines Chromosom ist die Anzahl der Felder in dem Gitter und steigt somit quadratisch an.

Problem Nr. 2: Die Konformation ist nicht mehr zwingend zusammenhängend. Dadurch wird die Berechnung der Fitnessfunktion aufwendiger.