

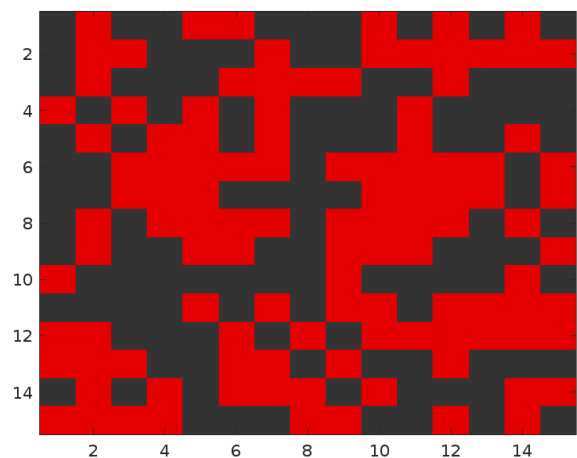
Percolazione

Studente: Alessio Russo **Matricola:** 376856
Corso di studio: Scienze Informatiche **Esame:** Modellazione e Simulazioni Numeriche

1 Algoritmo di Hoshen-Kopelman

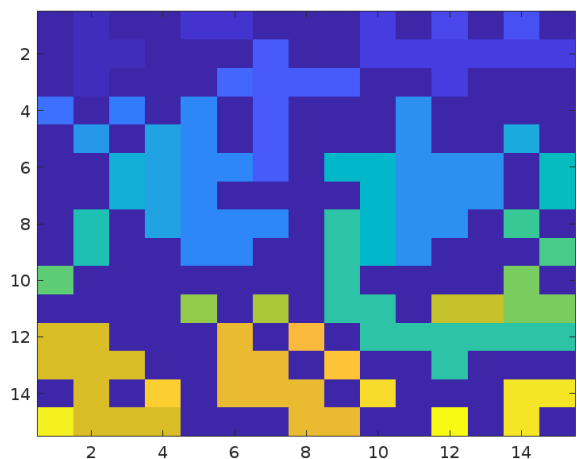
L'algoritmo di Hoshen-Kopelman (HK76) è una tecnica di etichettatura multipla dei cluster. Il reticolo viene visitato sito per sito per colonne, partendo dallo spigolo in alto a sinistra per arrivare a quello in basso a destra. Si prenda, ad esempio, il reticolo in figura

0	1	0	0	1	1	0	0	0	1	0	1	0	1	0
0	1	1	0	0	0	1	0	0	1	1	1	1	1	1
0	1	0	0	0	1	1	1	1	0	0	1	0	0	0
1	0	1	0	1	0	1	0	0	0	1	0	0	0	0
0	1	0	1	1	0	1	0	0	0	1	0	0	1	0
0	0	1	1	1	1	1	0	1	1	1	1	1	0	1
0	0	1	1	1	0	0	0	0	1	1	1	1	0	1
0	1	0	1	1	1	1	0	1	1	1	1	0	1	0
0	1	0	0	1	1	0	0	1	1	1	0	0	0	1
1	0	0	0	0	0	0	0	1	0	0	0	0	1	0
0	0	0	0	1	0	1	0	1	1	0	1	1	1	1
1	1	0	0	0	1	0	1	0	1	1	1	1	1	1
1	1	1	0	0	1	1	0	1	0	0	1	0	0	0
0	1	0	1	0	1	1	1	0	1	0	0	0	1	1
1	1	1	1	0	0	0	1	1	0	0	1	0	1	0



Durante la visita del reticolo, quando si incontra un sito occupato, allora: **(1)** Se il sito non è connesso ad altri siti occupato sopra o a sinistra, si inizia un nuovo cluster, a cui viene assegnata una **label** **(2)** Se c'è un primo vicino sopra o a sinistra occupato (uno solo dei due), il sito viene aggiunto al cluster del primo vicino occupato **(3)** Se i suoi primi vicini sono entrambi occupati, ma appartengono allo stesso cluster, il sito viene aggiunto al cluster dei primi vicini **(4)** Se i suoi primi vicini sono entrambi occupati, e non appartengono allo stesso cluster, il sito viene aggiunto al cluster con la **label** minore. Ad esempio, il cluster associati al reticolo precedente sono mostrati in figura

0	1	0	0	2	2	0	0	0	3	0	4	0	5	0
0	1	1	0	0	0	6	0	0	3	3	3	3	3	3
0	1	0	0	0	7	6	6	6	0	0	3	0	0	0
8	0	9	0	10	0	6	0	0	0	11	0	0	0	0
0	12	0	13	10	0	6	0	0	0	11	0	0	14	0
0	0	15	13	10	10	6	0	16	16	11	11	11	0	17
0	0	15	13	10	0	0	0	0	16	11	11	11	0	17
0	18	0	13	10	10	10	0	19	16	11	11	0	20	0
0	18	0	0	10	10	0	0	19	16	11	0	0	0	21
22	0	0	0	0	0	0	0	19	0	0	0	0	23	0
0	0	0	0	24	0	25	0	19	19	0	26	26	23	23
27	27	0	0	0	28	0	29	0	19	19	19	19	19	19
27	27	27	0	0	28	28	0	30	0	0	19	0	0	0
0	27	0	31	0	28	28	28	0	32	0	0	0	33	33
34	27	27	27	0	0	0	28	28	0	0	35	0	33	0



Tuttavia, quando si incontra un caso come quello descritto nel punto (4), occorre memorizzare che i due cluster sono in realtà lo stesso cluster. Questo viene fatto usando un vettore chiamato **Label of Label** (LofL), che contiene tutta l'informazione necessaria sui label dei cluster. In particolare: per un *good label*, memorizza la taglia del cluster; per *bad label*, memorizza qual è il vero cluster label a cui questo label appartiene. Questa distinzione viene fatta attraverso i segni dei numeri interi contenuti in LofL. Di seguito è riportato il LofL corrispondente al reticolo preso in esame

Label	1	2	3	4	5	6	7	8	9	10	11	12
Rank	4	2	10	-3	-3	24	-6	1	1	-6	33	1

Label	13	14	15	16	17	18	19	20	21	22	23	24
Rank	-10	1	-10	-11	2	2	-11	1	1	1	-11	1

Label	25	26	27	28	29	30	31	32	33	34	35
Rank	1	-11	11	8	1	1	-27	1	3	-27	1

Nota. Questo compito viene svolto dal modulo `HKclass`.

Alla fine, l'algoritmo HK (a meno che non venga fatta una rilabelizzazione successiva) non garantisce che tutti i siti di un fissato cluster abbiano lo stesso valore, ma restituisce in modo corretto le taglie dei cluster, che sono l'unica quantità a cui siamo interessati per la nostra analisi.