

D.8 Implementieren Sie einen Algorithmus `cluster`, der zu vorgegebener Distanzfunktion d über der Menge $U = \{0, \dots, m-1\}$ eine optimale Lösung von MAXIMUM l -CLUSTERING zurückliefert. Dabei sollen Sie `kruskal` aus der vorhergehenden Aufgabe unverändert verwenden und so das Problem MAXIMUM l -CLUSTERING auf MST reduzieren. Ihr Algorithmus besteht aus den folgenden Schritten:

1. Konstruktion eines vollständigen Graphen aus der Distanzfunktion
2. Aufruf von `kruskal`
3. Entfernen der $l - 1$ teuersten Kanten aus dem Spannbaum
4. Ausgabe einer Liste der Zusammenhangskomponenten des Graphen mit den verbliebenen Kanten

Führen Sie eine Laufzeitanalyse durch.

```

1
2 # l-CLUSTERING
3 from d07_kruskal import kruskal
4 from b12_comp import comp
5
6 def cluster(m,d,l):
7     # 1. vollständigen Graph mit m Knoten, m^2 Kanten und Kosten d[u,v] konstruieren
8     G = [{j:d[i,j] for j in range(m) if i!=j} for i in range(m)] # O(m^2)
9
10    # 2. Alg. kruskal unverändert aufrufen
11    (mst,_) = kruskal(G) # O(k*log(m))
12
13    # 3. die l-1 teuersten Kanten löschen
14    # Liste mit Kanten und dazugehörigen Kosten erstellen
15    li = []
16    for e in mst: # O(k)
17        (u,v) = e
18        dist = d[u,v]
19        li.append((dist,(u,v)))
20
21    # Liste reversed sortieren und ersten l-1 Elemente "entfernen"
22    edges = sorted(li, reverse=True)[l-1:] # O(k*log(k))
23
24
25    # 4. Zush.komponenten bestimmen, via B.12 (comp)
26    # dazu Graph mit allen Knoten konstruieren und Kanten einf.
27    Graph = [set() for _ in range(m)]
28    for e in edges: # O(k)
29        (_, (u,v)) = e
30        Graph[u].add(v)
31        Graph[v].add(u)
32
33    # Zusammenhangskomponenten bestimmen
34    comps = comp(Graph) # O(m+k)
35
36    return comps
37

```

Laufzeitanalyse

Es gilt $k \leq m^2$:

$$\begin{aligned}
 &\rightarrow O(m^2) + O(m^2 \cdot \log(m)) + O(m^2) + O(m^2 \cdot \log(m^2)) + O(m^2) + O(m^2 + m) \\
 &\quad \in O(m^2 \cdot \log(m))
 \end{aligned}$$

$\text{Comp}(\text{Graph})$ in $O(m)$, weil der Graph genauso viele Kanten hat, wie der Spannbaum und der Spannbaum ist in $O(m)$, weil Spannbaum ist ein Baum und im Baum gilt $k = m - 1$. Also gilt für Comp:

$$O(m + k) = O(m + m) = O(m)$$

Gleiches würde für „edges“ gelten. Da müsste dann wahrscheinlich auch $O(m \cdot \log m)$ hin