## Pre-cluster

.....

## Main C2G

```
Initiate Empty Queue;
Create Root Node;
Queue.push (Root Node);
While Queue is not empty:
Node id = Queue.pop();
Best Entropy = inf;
Best strategy = [];
     For all marker pairs:
          Compute Best Gate();
          Compute Entropy();
          If entropy < Best Entropy:</pre>
               Best Entropy = entropy;
               Best strategy = Current gates;
     End
     If Best strategy is not []:
          Create Children Nodes;
          Queue.push(All New Nodes);
     End
End
```

Compute\_entropy

## Compute\_Best\_Gate

```
Set: Grid Size;
Compute: x interval and y interval;
Grid X = (x-min(x))/x interval;
Grid Y = (y-min(y))/y interval;
Hash key = Grid X*100+Grid Y
Grid = unique(Hash key);
Initiate Num of Cell Grid;
%% Above is a n-by-m matrix. n is total number of grid and
m is total number of cell population. It's used to record
number of different cells in different grid.
For each cell:
     Add 1 to right place in Num of Cell Grid;
End
Compute Adjacent Matrix;
% overlap between 2 population is \sum_{All\ Grid} \sqrt{p_1p_2}. Where p_i is
number of cells in population i out of all cell in
population i.
MCL clustering.
%% Each MCL cluster is one gate we want to draw.
For each MCL cluster:
     If it has no target population, skip;
     Compute grid convex hull of the cluster;
     %%The union of above convex hull is the gate;
End
```

## Compute\_convex\_hull

End

```
cell. But the outlier level is unknown. We want to try
different values.
Best f score = 0;
Best Boundary = [];
For different ignore percentage:
     Boundary = [];
     For Each population in the cluster:
          Drop small grids;
          Compute Convex hull;
          Boundary = Union(Boundary, Convex hull);
     End
     Compute F score;
     If F score > Best f score:
          Best f score = F score;
          Best Boundary = Boundary;
     End
```

%% Here, we want to ignore certain percentage of outlier