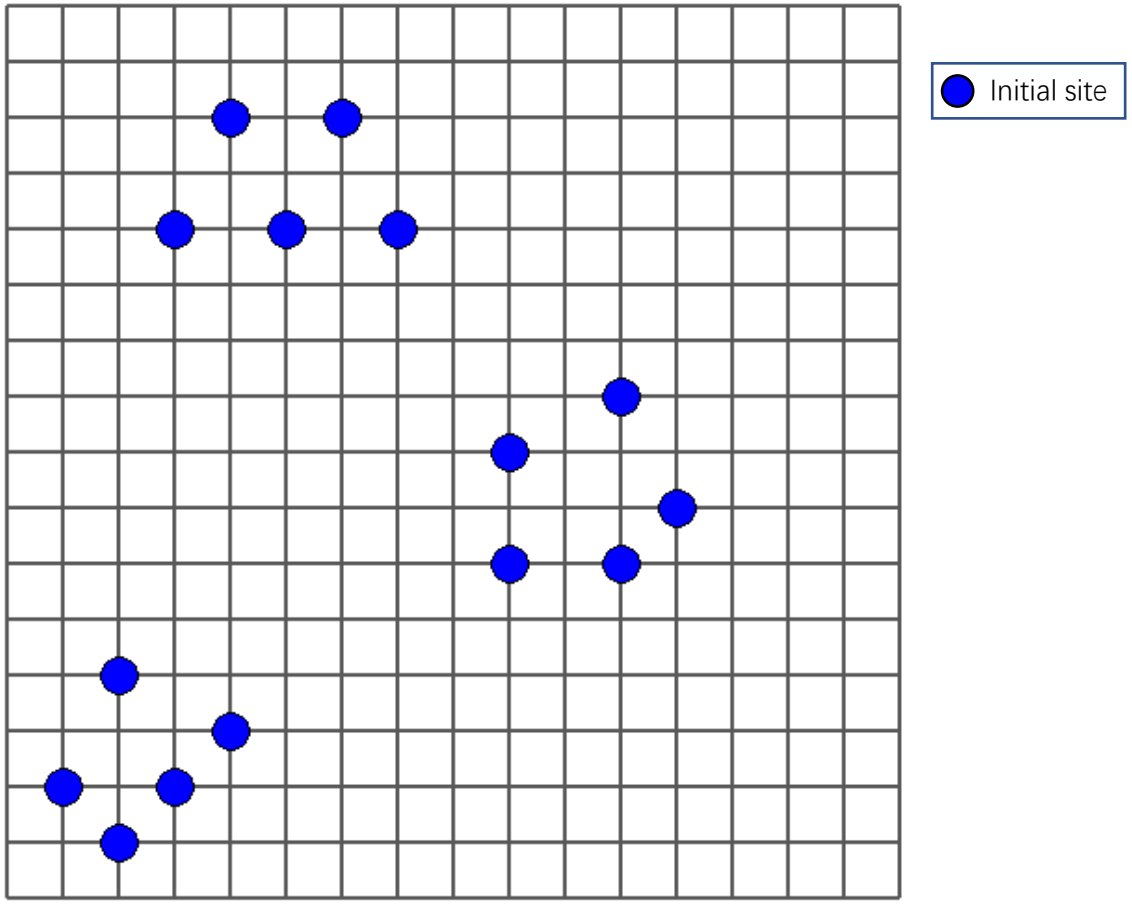
Advanced Algorithm
Assignment 4
Center Selection Problem

12032189

Yuxi Liu

Exercise 4-1

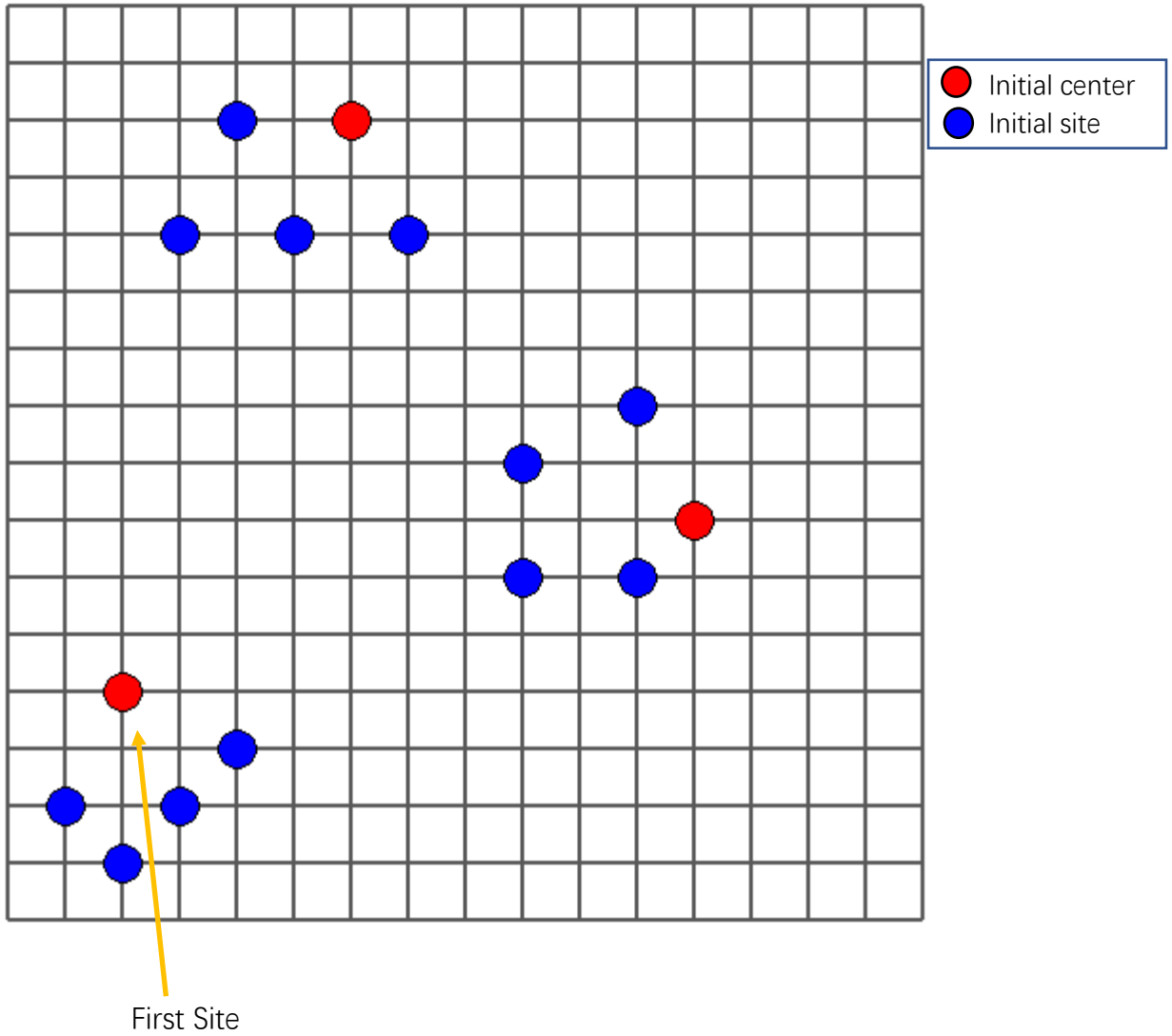


In the k-means algorithm, we can start with (i) using an initial partition {*S*1, *S*2, ..., *Sk*} or with (ii) using initial centers {*c*1, *c*2, ..., *ck*}. Design a good initialization method for k-means algorithm for (i) and also for (ii).

Greedy-Center-Selection

procedure CENTER-SELECT
    Assume $k \leq |S|$ (else define C=S)
    Select any site s and let $C = \{s\}$
    while $|C| < k$ do
        Select a site $s \in S$ that maximizes dist(s, C)
        Add s to C
    end while
    Return C as the selected set of sites
End procedure

initial centers selection algorithm

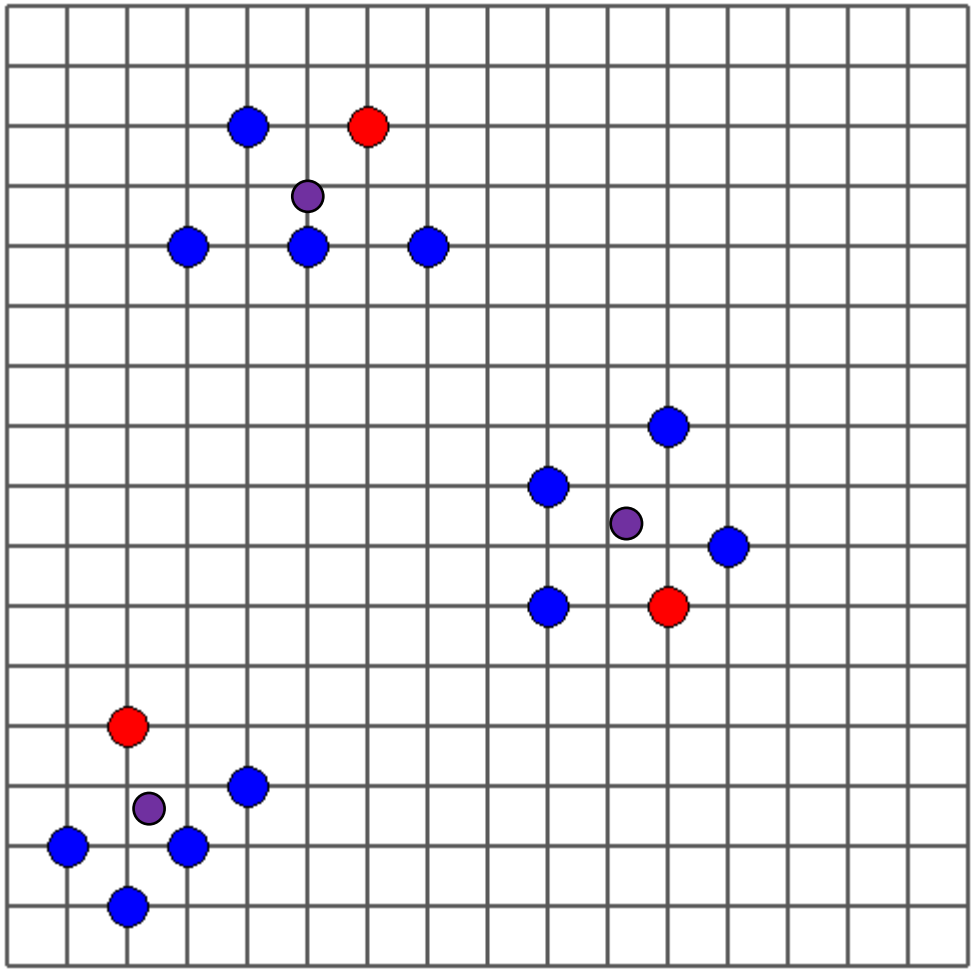


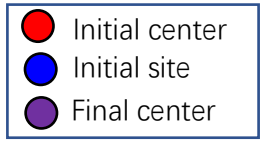
In the k-means algorithm, we can start with (i) using an initial partition {*S*1, *S*2, ..., *Sk*} or with (ii) using initial centers {*c*1, *c*2, ..., *ck*}. Design a good initialization method for k-means algorithm for (i) and also for (ii).

Greedy-Center-Selection

procedure CENTER-SELECT
    Assume $k \leq |S|$ (else define C=S)
    Select any site s and let $C = \{s\}$
    while $|C| < k$ do
        Select a site $s \in S$ that maximizes dist(s, C)
        Add s to C
    end while
    Return C as the selected set of sites
End procedure

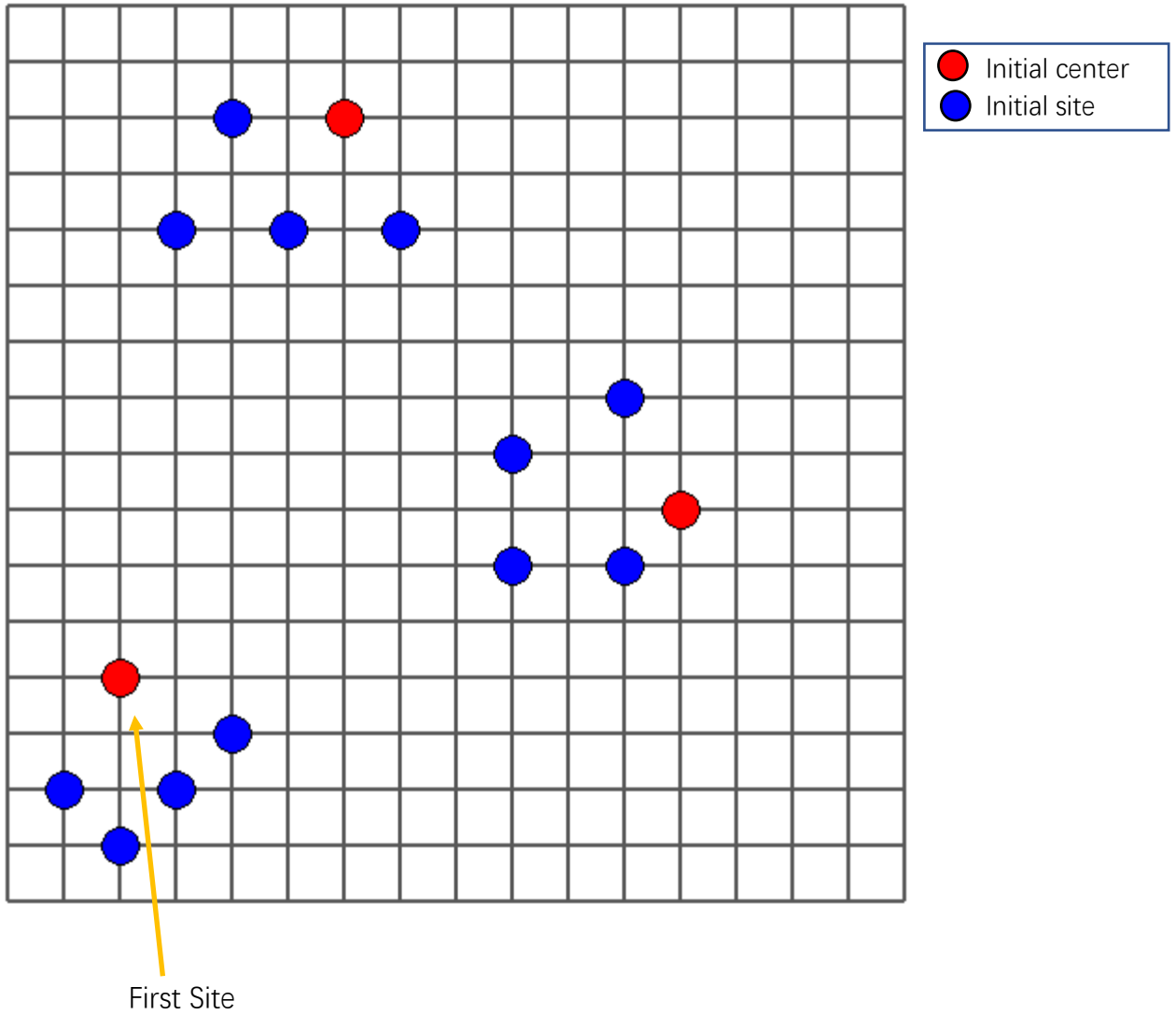
initial centers selection algorithm





$$r(C) = r(C^*) \approx 2.154$$

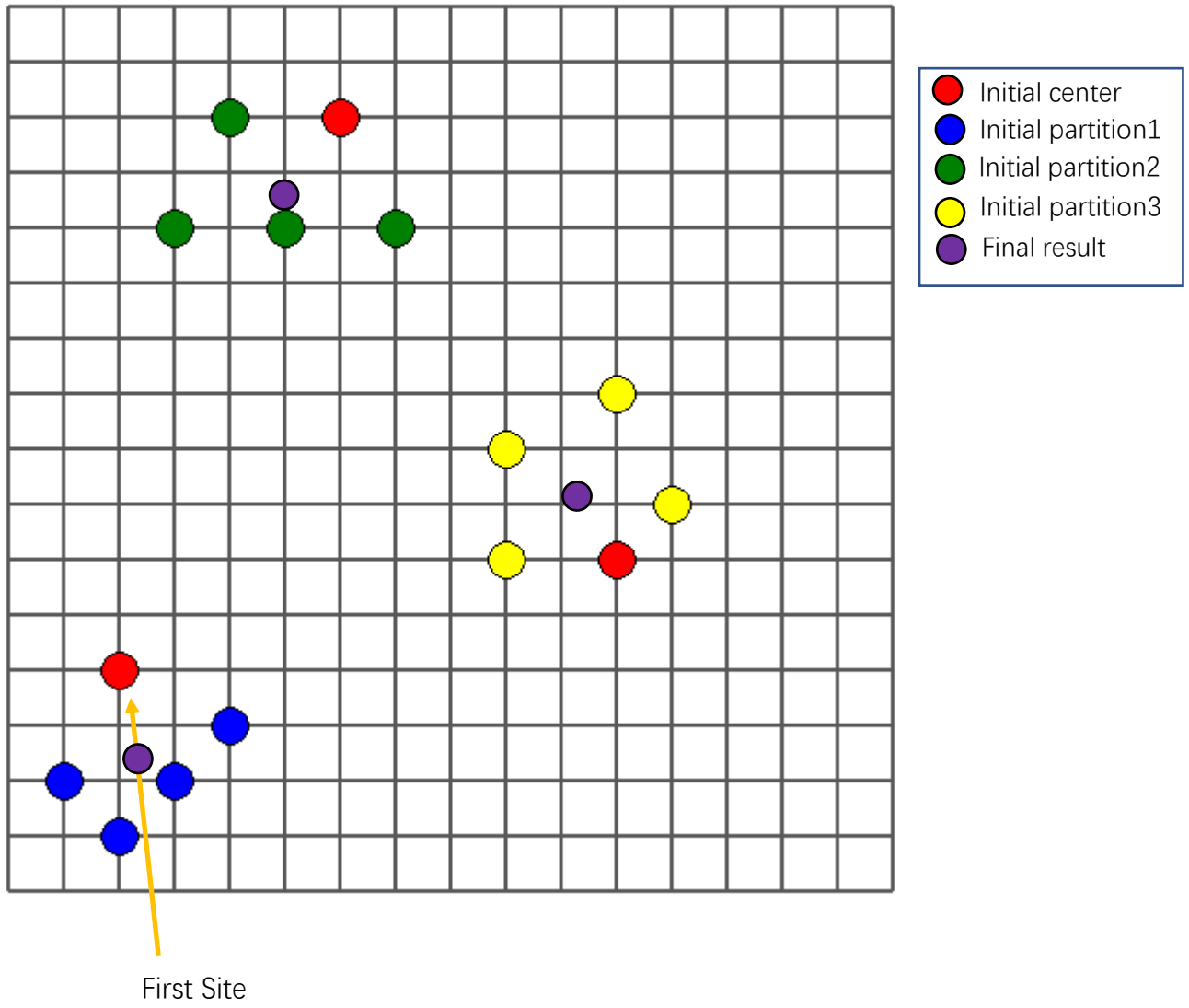
initial partition selection algorithm



Assume that we have N sites and divide them into K clusters. We calculate the distance between each site and these initial center. And partition the site by minimum distance.

procedure PARTITION-SELECT
    Assume $k \leq |S|$ (else define C=S)
    Set initial center as the initial cluster center C
    for s_i not in C
        put s_i into the cluster with minimum dist(s_i,C)
    Return C as initial partition
End procedure

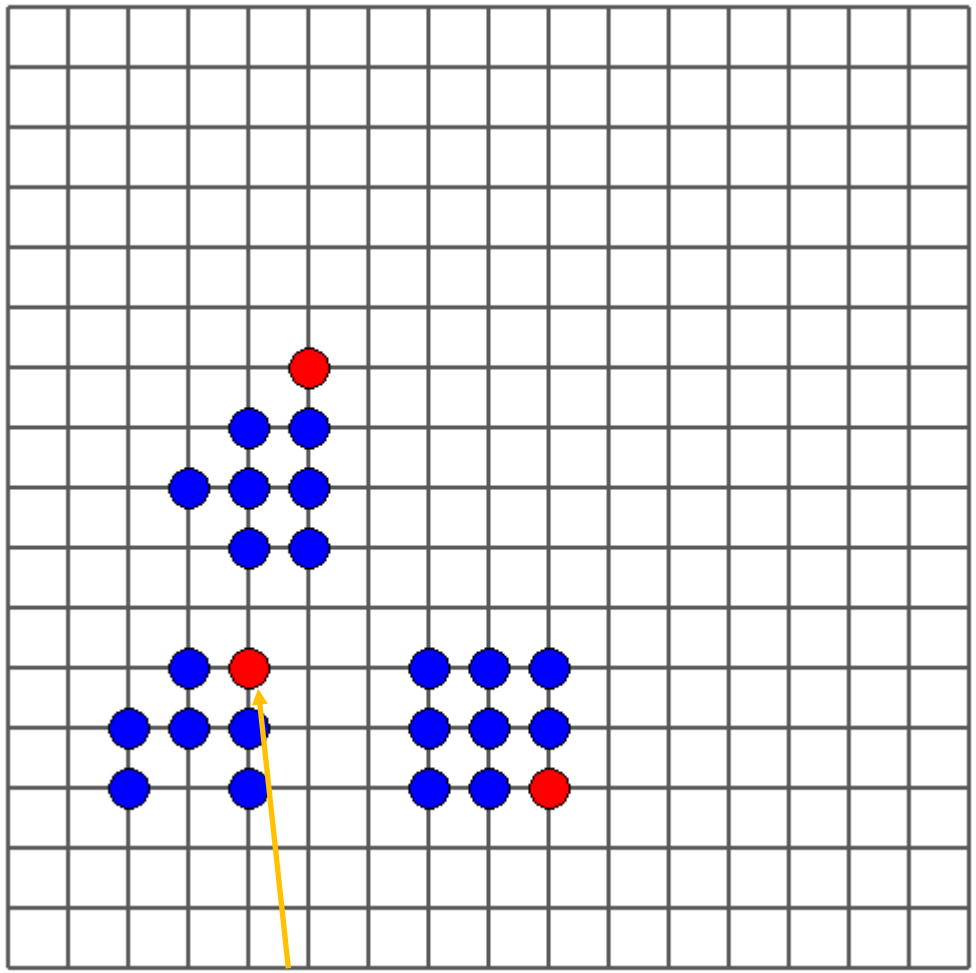
initial partition selection algorithm



$$r(C) = r(C^*) \approx 2.154$$

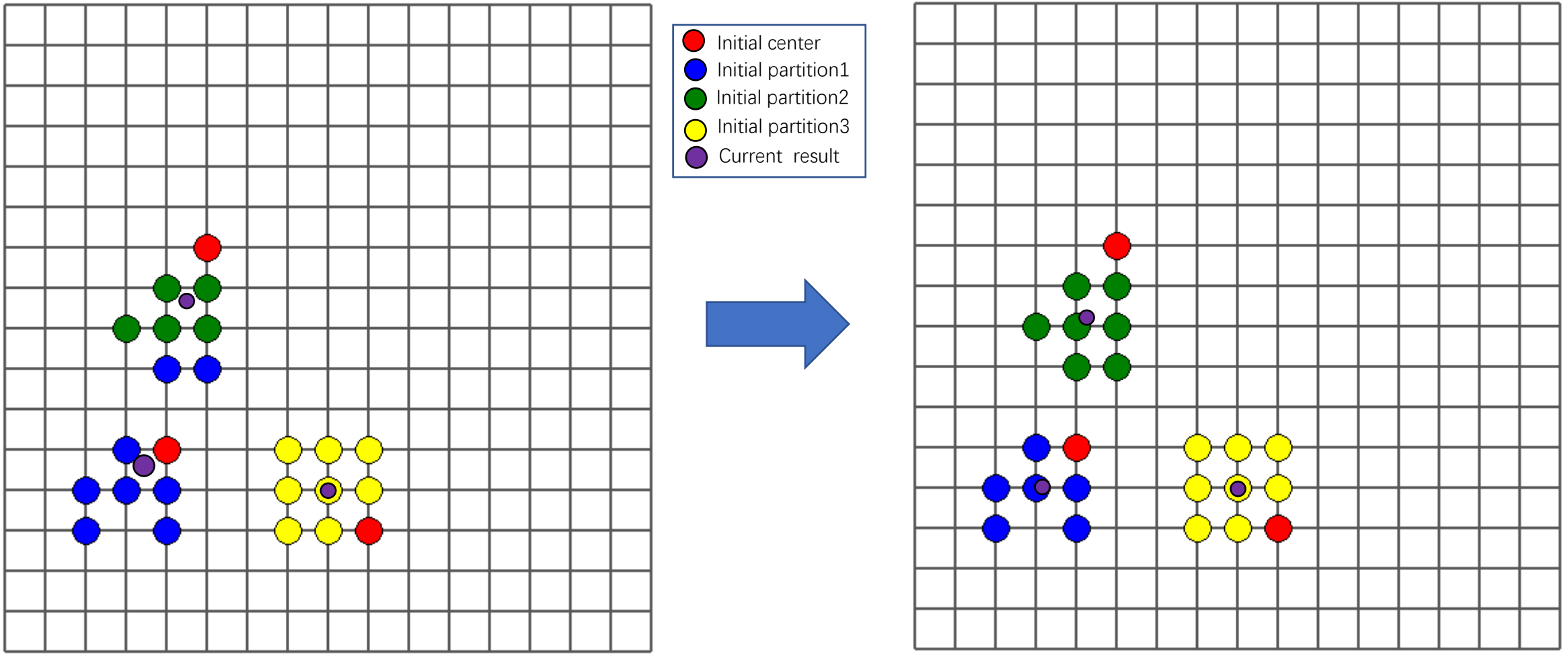
initial partition selection algorithm

Hard case：



initial partition selection algorithm

Hard case：



Thank you