A general notion of interiority

J. Orestes Cerdeira
Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa,
Portugal

Joint with M. João Martins, Pedro C. Silva and T. Monteiro-Henriques

cma FCT.UNL
centro de matemática e aplicações

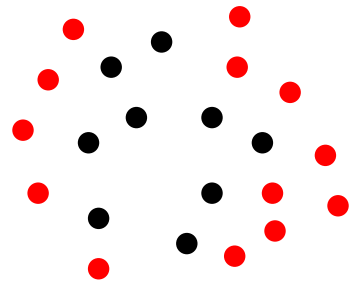
FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

two issues...

two issues...

- □ to what extent a set of entities $\{\bullet\}$ is separated from $\bullet$?

- to what extent a set of entities $\{\bullet\}$ is separated from $\bullet$?

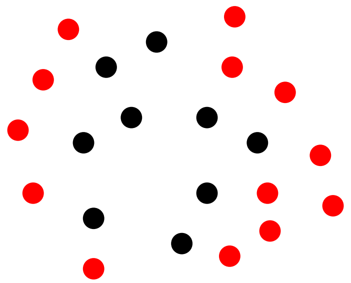


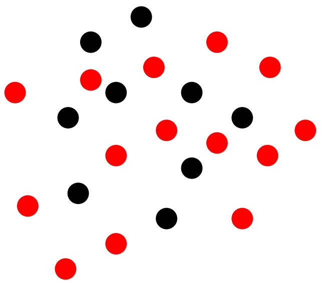
two issues...

- to what extent a set of entities {●} is separated from ● ?

two issues...

- □ to what extent a set of entities {•} is separated from • ?

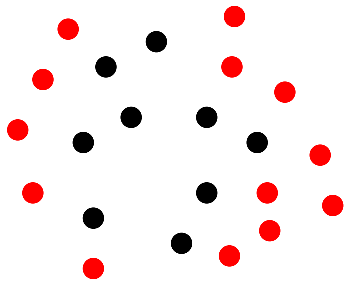


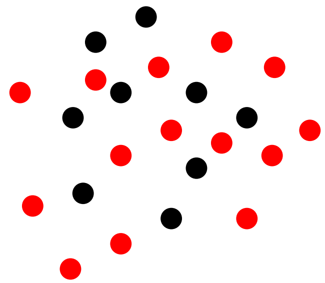


- □ how are the entities of a group arranged together ?

two issues...

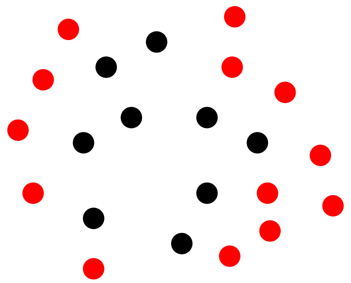
- □ to what extent a set of entities {•} is separated from • ?

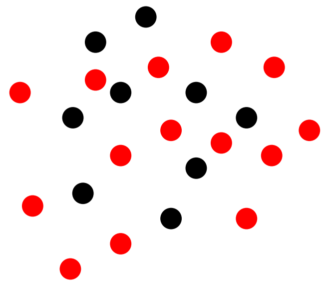




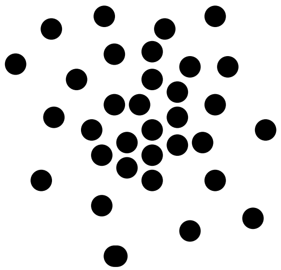
- □ how are the entities of a group arranged together ?
(more concentrated to the "interior“ or dispersed on the "margins“ of their range)

- to what extent a set of entities {●} is separated from ● ?





- how are the entities of a group arranged together ?
(more concentrated to the "interior“ or dispersed on the "margins“ of their range)

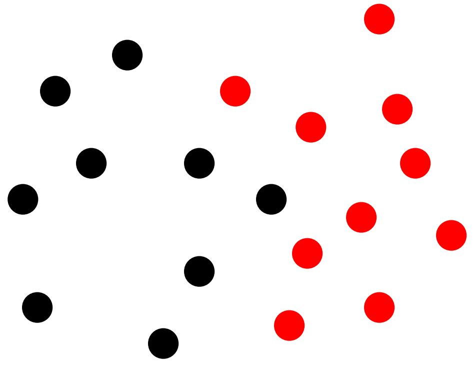


two issues...

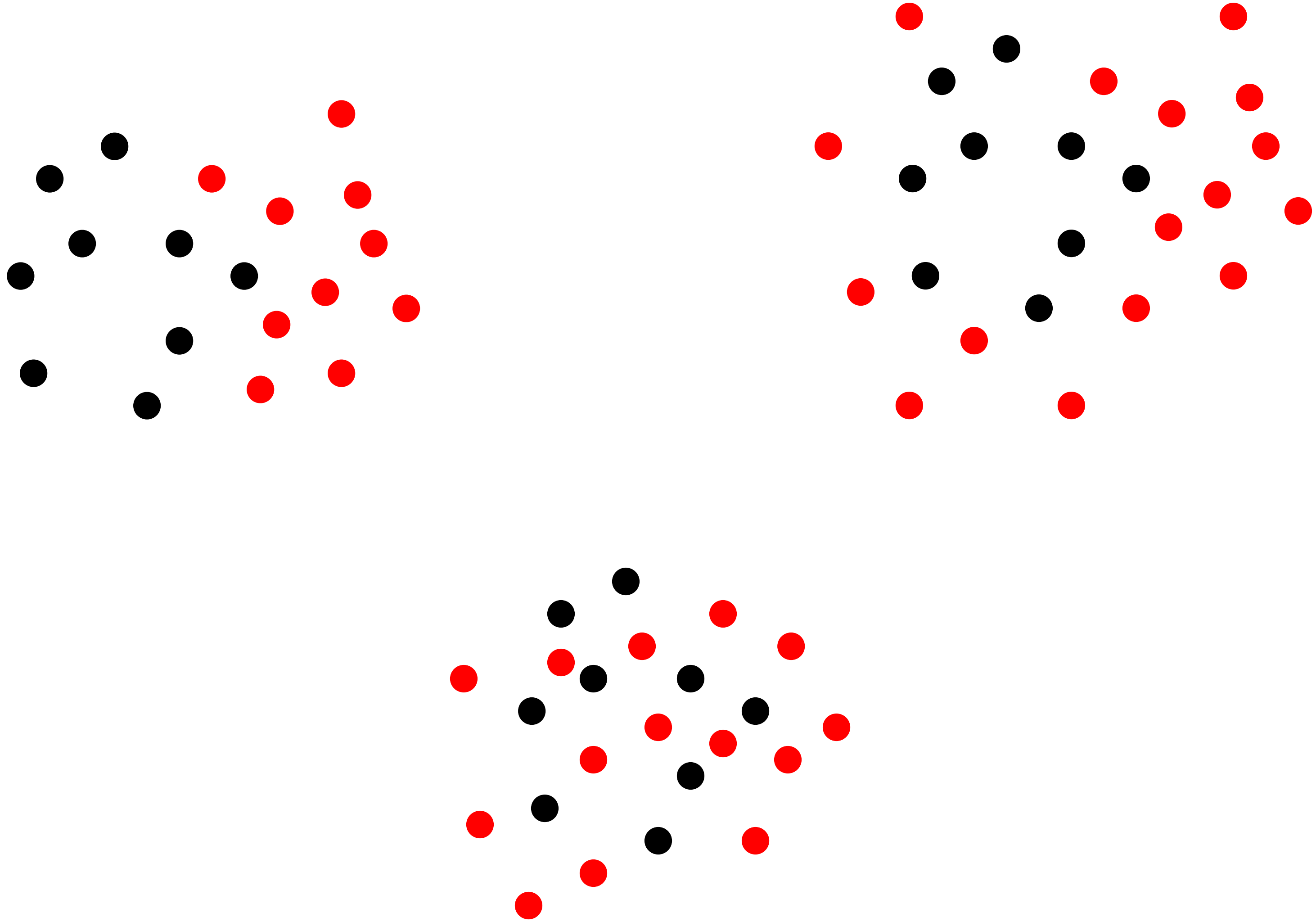
- to what extent a set of entities {●} is separated from ● ?

- how are the entities of a group arranged together ?
(more concentrated to the "interior" or dispersed on the "margins" of their range)

issue 1: is $\{\bullet\}$ well separated from $\{{\color{red}\bullet}\}$?



issue 1: is $\{\bullet\}$ well separated from $\{{\color{red}\bullet}\}$?



In many areas …

In many areas ...

Ecology, Chemistry, Molecular Biology, Physiology, Toxicology, Social Sciences ...

In many areas ...

Ecology, Chemistry, Molecular Biology, Physiology, Toxicology, Social Sciences ...

issue:

In many areas ...

Ecology, Chemistry, Molecular Biology, Physiology, Toxicology, Social Sciences ...

issue: to what extent a particular subset X of entities is separated from the other entities?

In many areas ...

Ecology, Chemistry, Molecular Biology, Physiology, Toxicology, Social Sciences ...

issue: to what extent a particular subset X of entities is separated from the other entities? arises

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

Dunn's index and Dunn-like indices, Davies-Bouldin's index, the silhouette index and Thornton's index.

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

Dunn's index and Dunn-like indices, Davies-Bouldin's index, the silhouette index and Thornton's index.

With these approaches separability is numerically quantified

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

Dunn's index and Dunn-like indices, Davies-Bouldin's index, the silhouette index and Thornton's index.

With these approaches separability is numerically quantified, but

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

Dunn's index and Dunn-like indices, Davies-Bouldin's index, the silhouette index and Thornton's index.

With these approaches separability is numerically quantified, but

- □ no detailed information on how each cluster is separated from the others is produced

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

Dunn's index and Dunn-like indices, Davies-Bouldin's index, the silhouette index and Thornton's index.

With these approaches separability is numerically quantified, but

- □ no detailed information on how each cluster is separated from the others is produced
- □ all, except Thornton's index, are expressions that explicitly involve numerical dissimilarities between entities

In the context of cluster analysis...

methods have been proposed that measure overlap/separability of clusters as part of assessing the quality of a given partition.

Dunn's index and Dunn-like indices, Davies-Bouldin's index, the silhouette index and Thornton's index.

With these approaches separability is numerically quantified, but

- □ no detailed information on how each cluster is separated from the others is produced
- □ all, except Thornton's index, are expressions that explicitly involve numerical dissimilarities between entities
- □ notions such as "perfect separation of cluster" are not comprise.

Silhouette index

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

Silhouette index

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

$S(i) \in [-1, 1]$.

Silhouette index

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

$S(i) \in [-1, 1]$.

$S(i) \simeq 1$, means i is "well-clustered" in X

Silhouette index

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

$S(i) \in [-1, 1]$.

> $S(i) \simeq 1$, means i is “well-clustered” in X
> $S(i) \simeq 0$, means i lies “equally far away” from X and $\bar{X}$

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

$S(i) \in [-1, 1]$.

> $S(i) \simeq 1$, means i is "well-clustered" in X
> $S(i) \simeq 0$, means i lies "equally far away" from X and $\bar{X}$
> $S(i) \simeq -1$, means i is "misclassified" as an X entity.

Silhouette index

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

$S(i) \in [-1, 1]$.

$S(i) \simeq 1$, means i is "well-clustered" in X
$S(i) \simeq 0$, means i lies "equally far away" from X and $\bar{X}$
$S(i) \simeq -1$, means i is "misclassified" as an X entity.

$S(X)$ is the average of the silhouettes of $i \in X$

Silhouette index

To construct the silhouette $S(X)$ (with respect to $\bar{X}$), for $i \in X$

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$a(i)$ - average dissimilarity of entity i to all other entities in X
$b(i)$ - average dissimilarity of entity i to all entities in $\bar{X}$

$S(i) \in [-1, 1]$.

$S(i) \simeq 1$, means i is "well-clustered" in X
$S(i) \simeq 0$, means i lies "equally far away" from X and $\bar{X}$
$S(i) \simeq -1$, means i is "misclassified" as an X entity.

$S(X)$ is the average of the silhouettes of $i \in X$

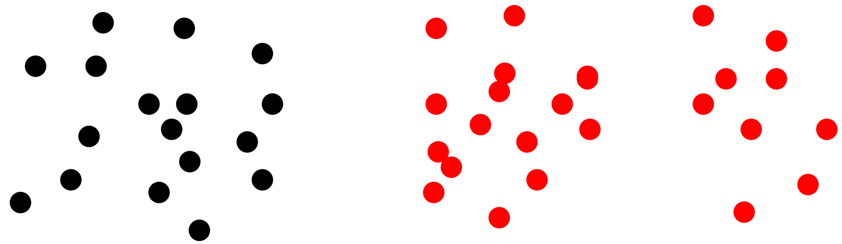
notions such as "perfect separation of cluster" are not comprise.

Thornton's index

Thornton's index

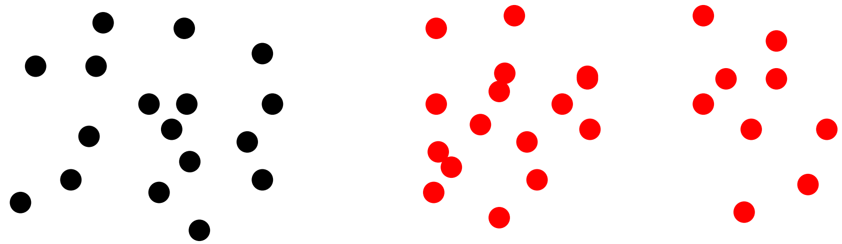
$$GSI = \frac{\sum_{i \in X} \delta_i}{|X|},\ \delta_i = 1(0) \text{ if the nearest neighbor of } i \text{ is (not) in } X$$

$$GSI = \frac{\sum_{i \in X} \delta_i}{|X|},\ \delta_i = 1(0) \text{ if the nearest neighbor of } i \text{ is (not) in } X$$



Thornton's index

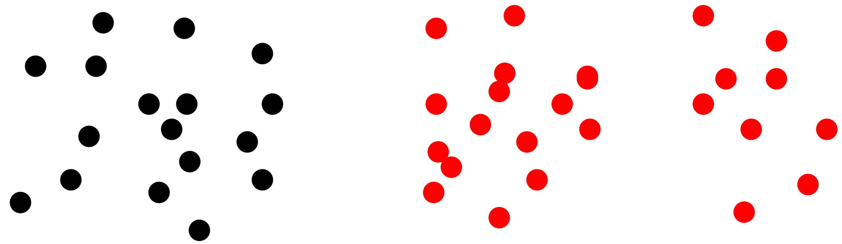
$$GSI = \frac{\sum_{i \in X} \delta_i}{|X|}, \ \delta_i = 1(0) \text{ if the nearest neighbor of } i \text{ is (not) in } X$$



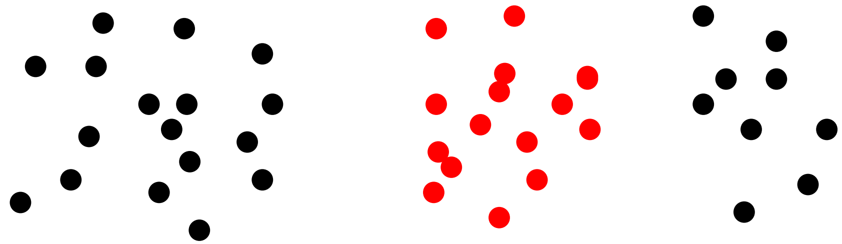
$$GSI = 1$$

Thornton's index

$$GSI = \frac{\sum_{i \in X} \delta_i}{|X|}, \ \delta_i = 1(0) \text{ if the nearest neighbor of } i \text{ is (not) in } X$$

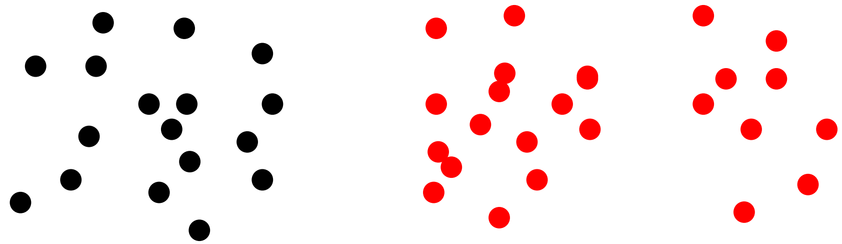


$$GSI = 1$$

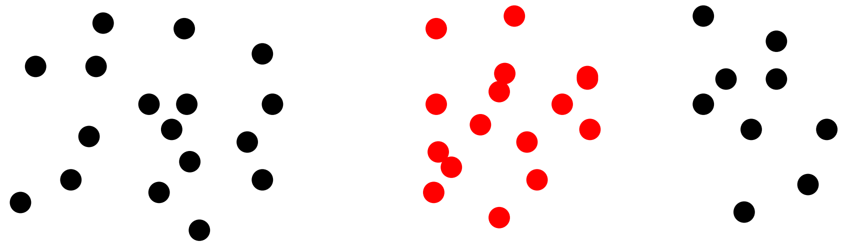


Thornton's index

$$GSI = \frac{\sum_{i \in X} \delta_i}{|X|}, \ \delta_i = 1(0) \text{ if the nearest neighbor of } i \text{ is (not) in } X$$



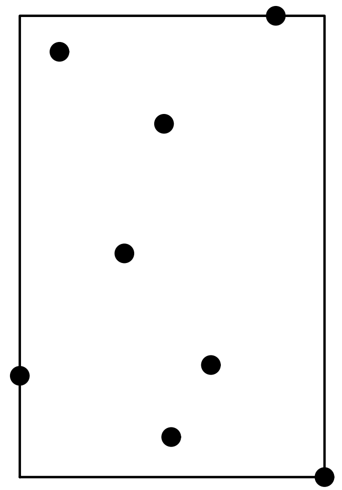
$GSI = 1$



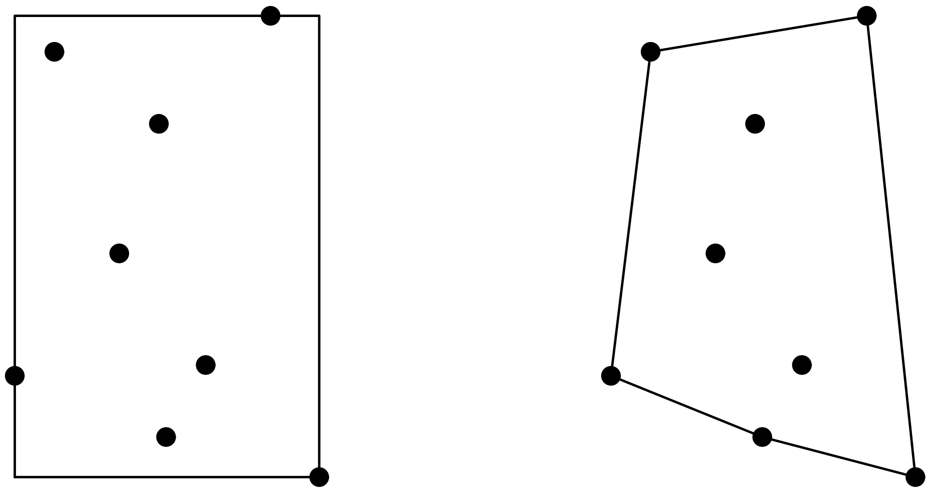
$GSI = 1$

other proposals rely on a notion of *territory*

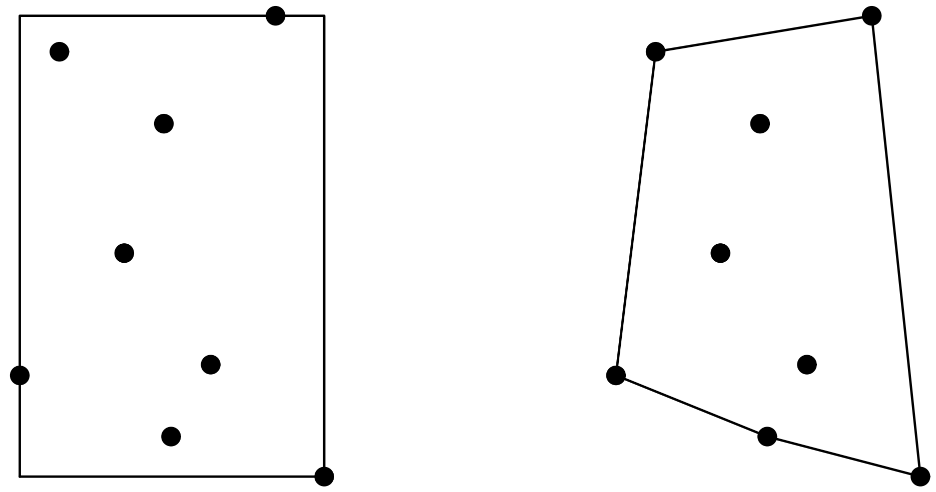
other proposals rely on a notion of *territory*



other proposals rely on a notion of *territory*



other proposals rely on a notion of *territory*



doesn't discriminate separability

Instead of deeming...

to what extent a set of entities X is separated from <u>other</u> entities?

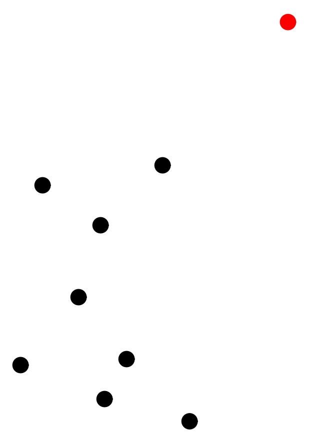
Instead of deeming...

to what extent a set of entities X is separated from <u>other</u> entities?

to what extent a set of entities X is separated from <u>ano</u>ther entity?

to what extend $\{\bullet\}$ is separated from $\bullet$?

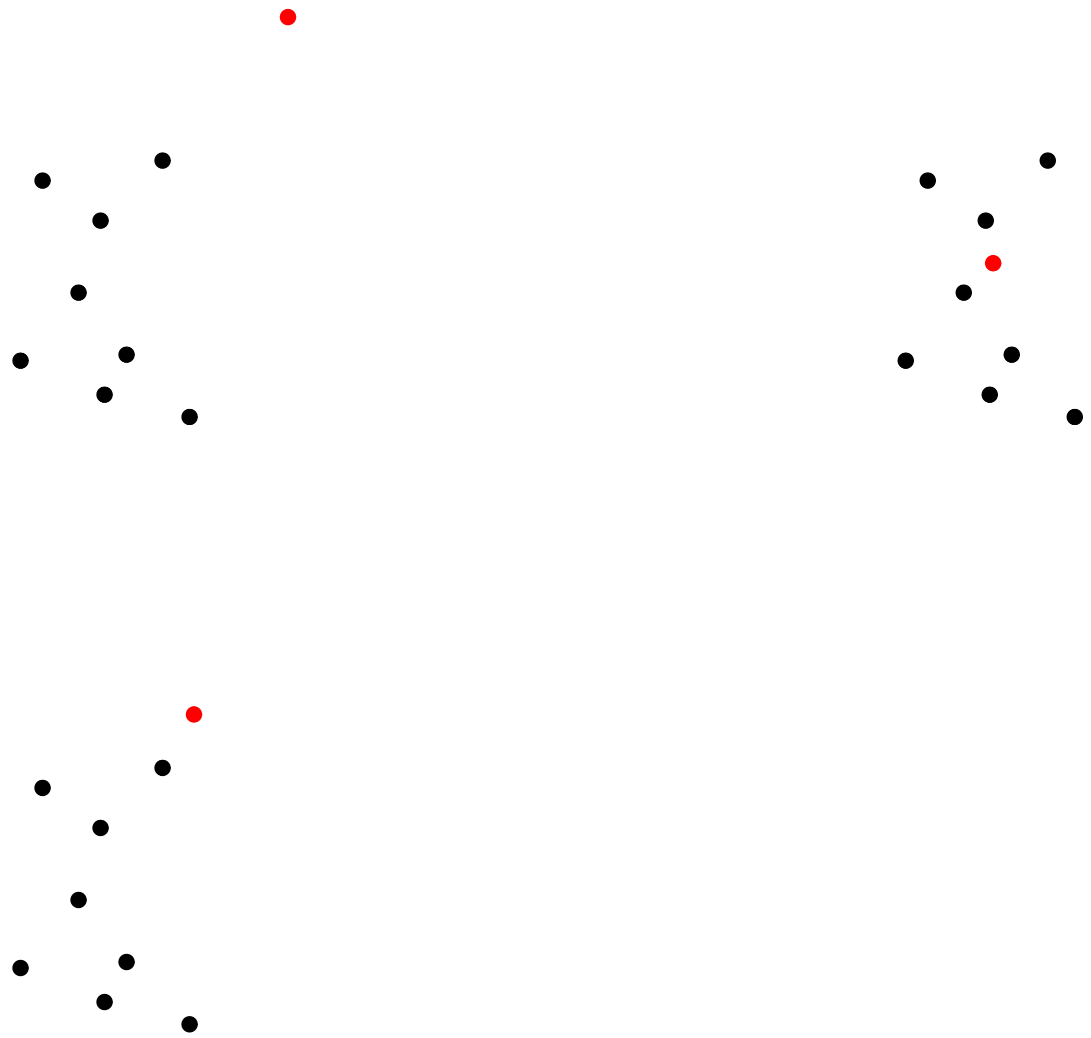
to what extend $\{\bullet\}$ is separated from $\bullet$?



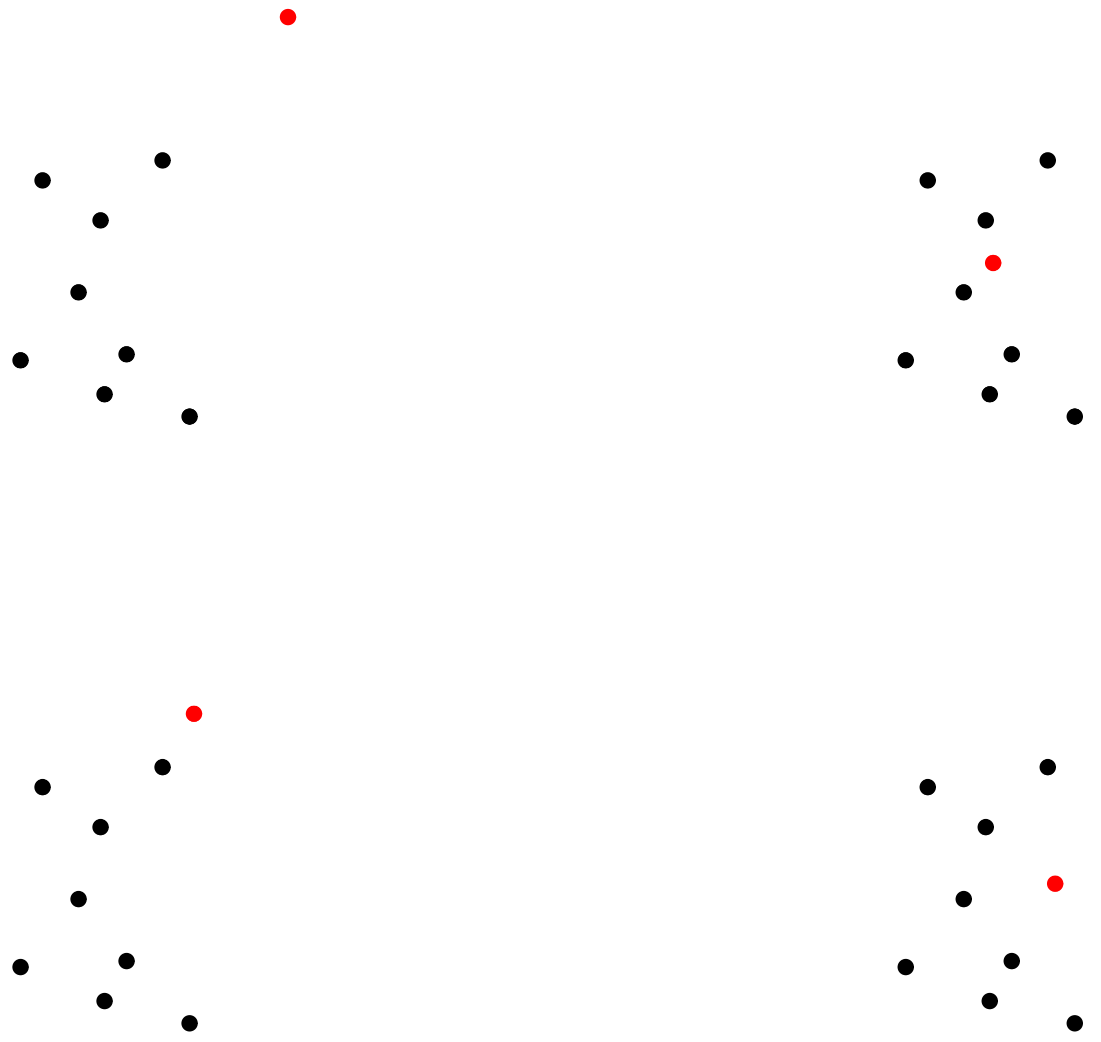
to what extend $\{\bullet\}$ is separated from $\bullet$?

▷ Introduction
Interiority
Separability index
Convexity criterion
Auto-interiority

to what extend $\{\bullet\}$ is separated from $\bullet$?



to what extend $\{\bullet\}$ is separated from $\color{red}{\bullet}$?



to formalize interiority needs fixing some aggregation criterion

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

□ if the sum of distances from each point to its cluster center is minimum

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

□ if the sum of distances from each point to its cluster center is minimum (k-means)

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

- ☐ if the sum of distances from each point to its cluster center is minimum (k-means)
- ☐ if the largest diameter of its clusters is minimum

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

- if the sum of distances from each point to its cluster center is minimum (k-means)
- if the largest diameter of its clusters is minimum (diameter)

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

- □ if the sum of distances from each point to its cluster center is minimum (k-means)
- □ if the largest diameter of its clusters is minimum (diameter)
- □ if clusters are the connected components of some minimum $(|E| - k)$-spanning forest

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

P_k is adequate

- □ if the sum of distances from each point to its cluster center is minimum (k-means)
- □ if the largest diameter of its clusters is minimum (diameter)
- □ if clusters are the connected components of some minimum $(|E| - k)$-spanning forest (single linkage)

to formalize interiority needs fixing some aggregation criterion

i.e., a rule that identifies, for a set of entities E and $1 \leq k \leq |E|$, the k-partitions of E which are *adequately* aggregated

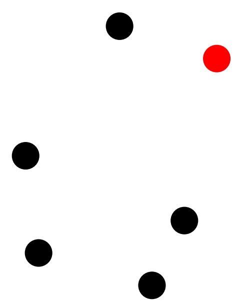
P_k is adequate

- □ if the sum of distances from each point to its cluster center is minimum (k-means)
- □ if the largest diameter of its clusters is minimum (diameter)
- □ if clusters are the connected components of some minimum $(|E| - k)$-spanning forest (single linkage)
- □ $\cdots$

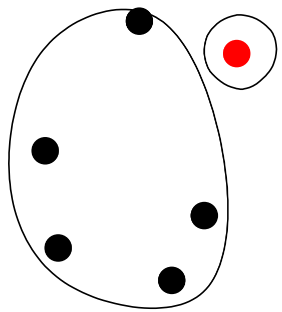
entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$

k-interior

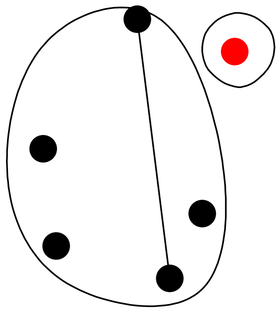
entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$

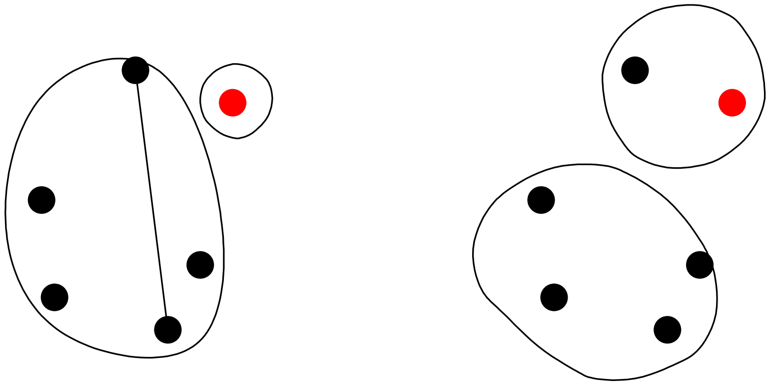


entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$

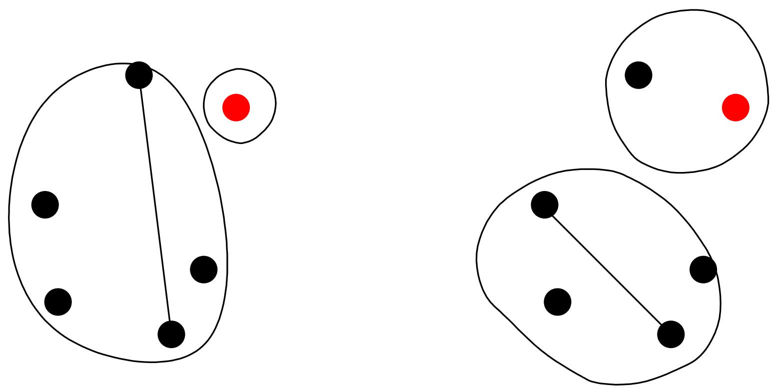


k-interior

entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$

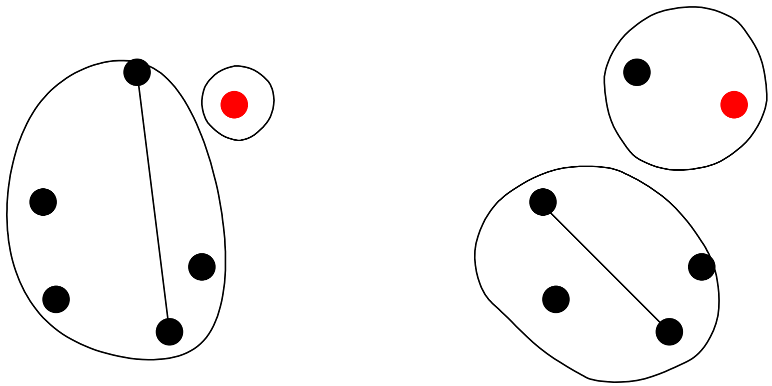


entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



k-interior

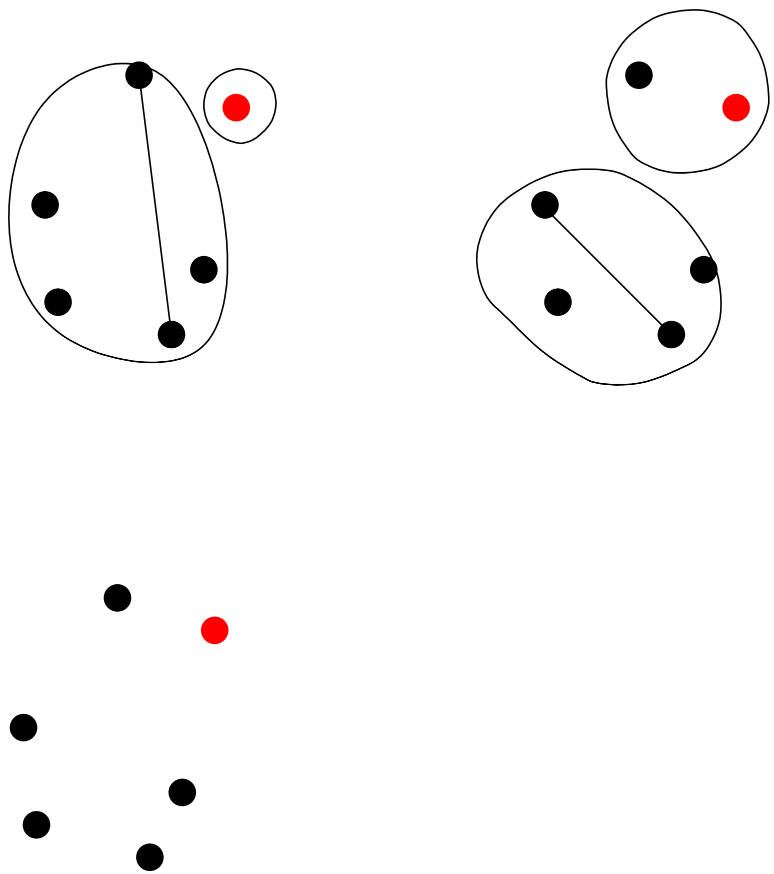
entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



● is 1-interior

k-interior

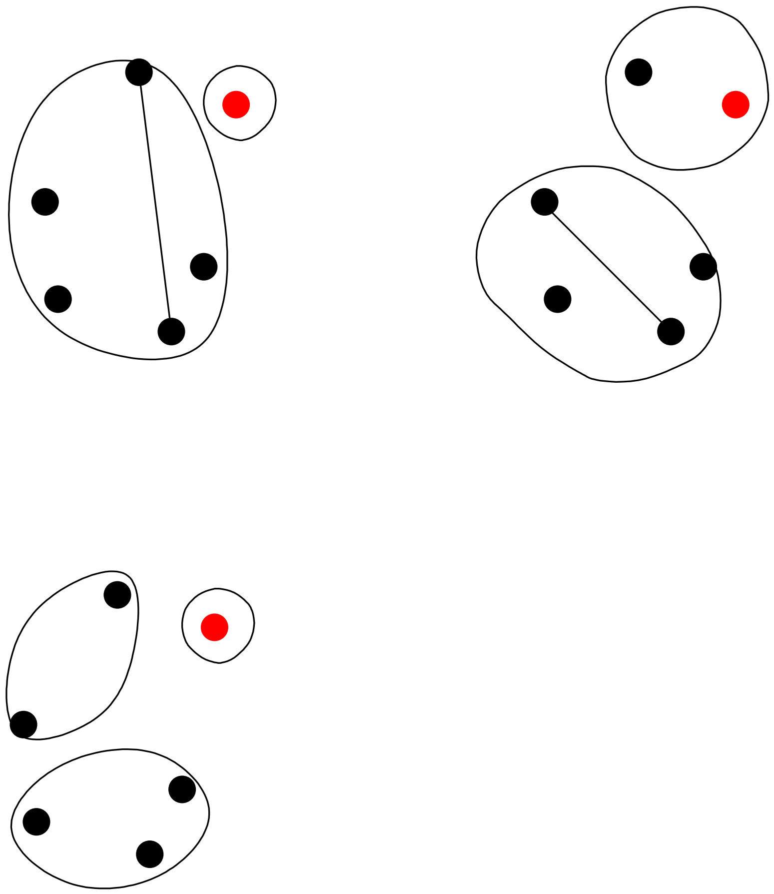
entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



is 1-interior

k-interior

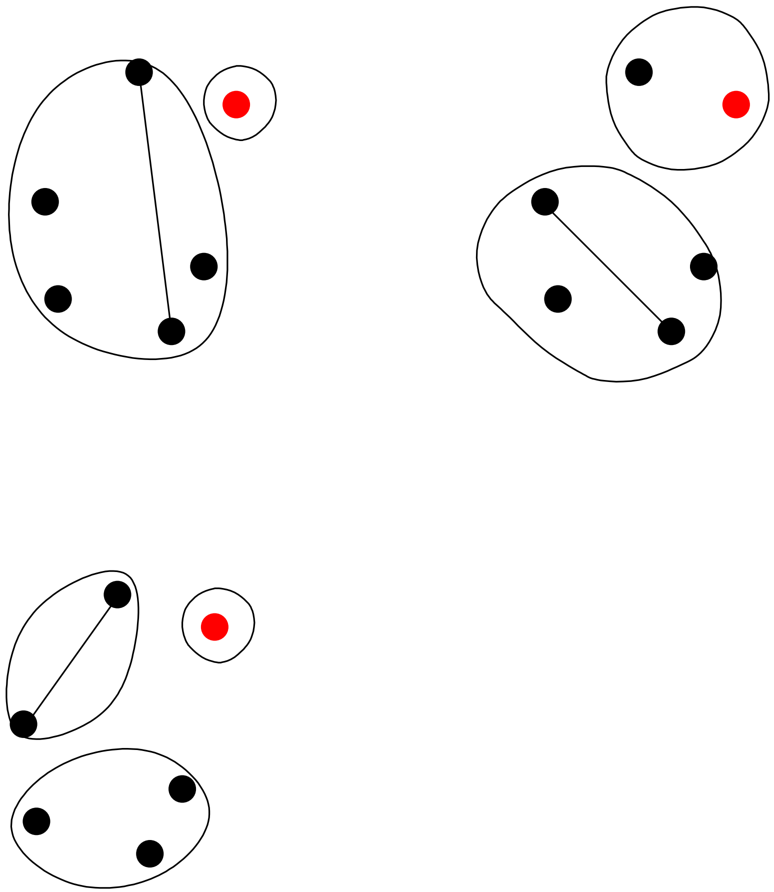
entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



is 1-interior

k-interior

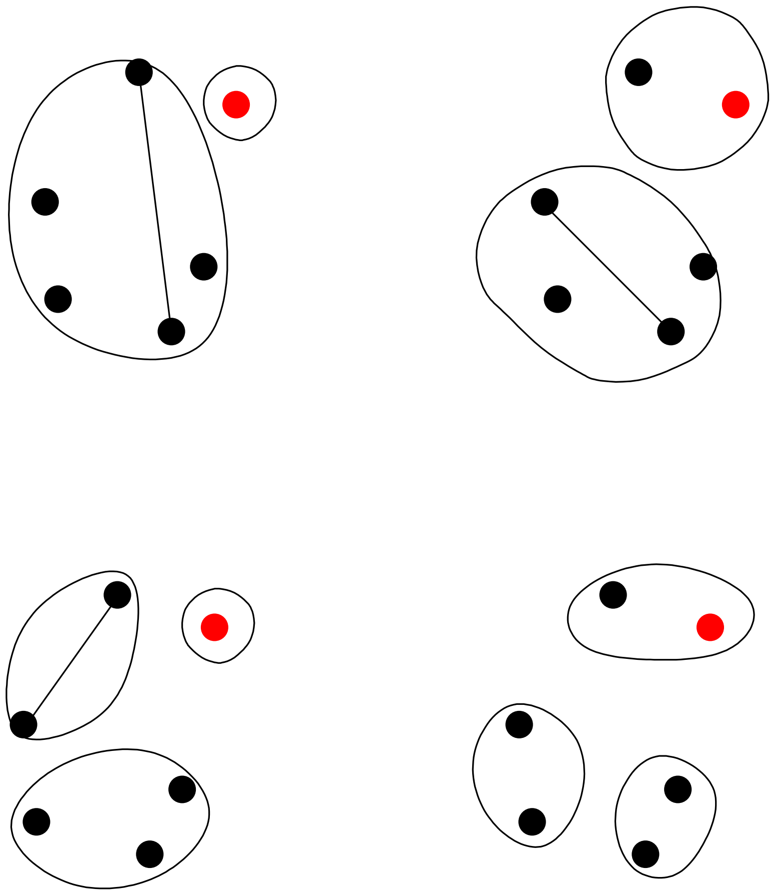
entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



• is 1-interior

k-interior

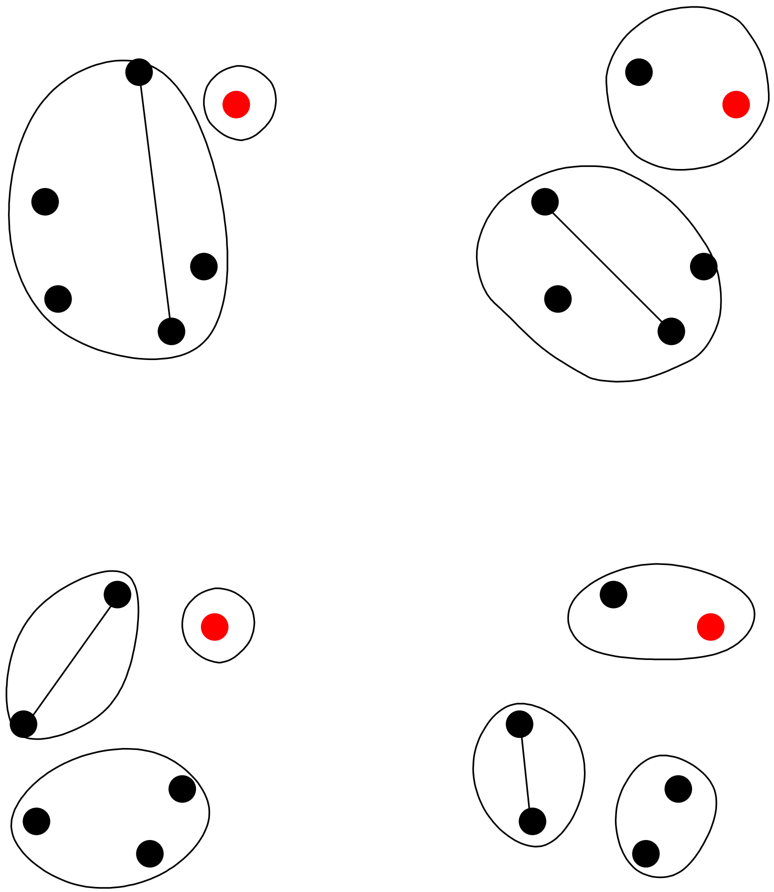
entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



● is 1-interior

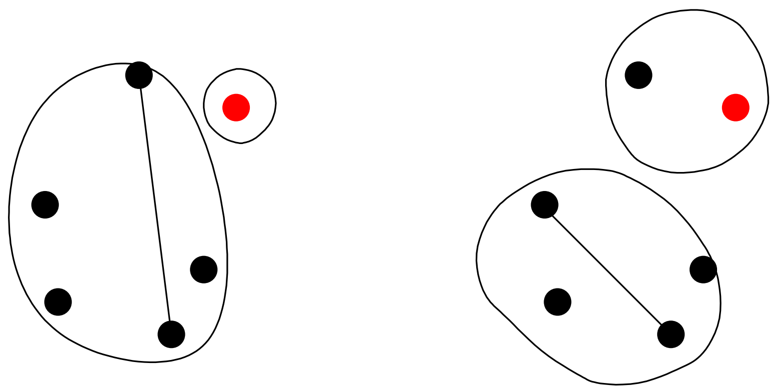
k-interior

entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$

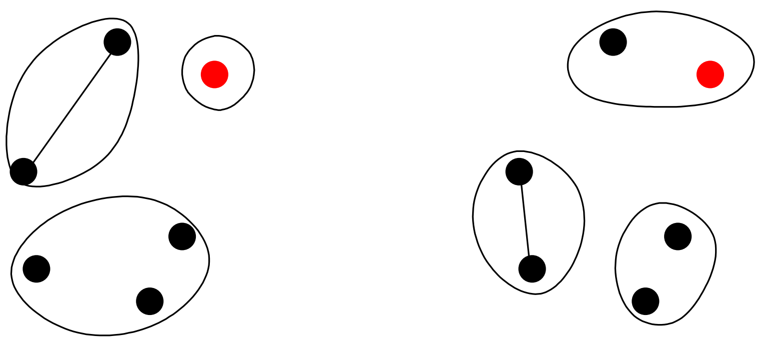


● is 1-interior

entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$

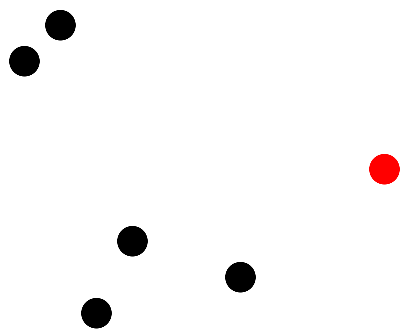


• is 1-interior



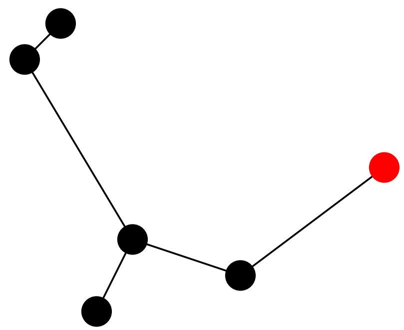
• is 2-interior

entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



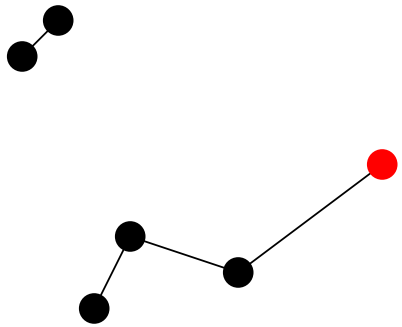
k-interior

entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



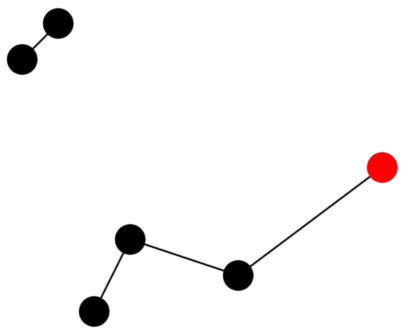
k-interior

entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



k-interior

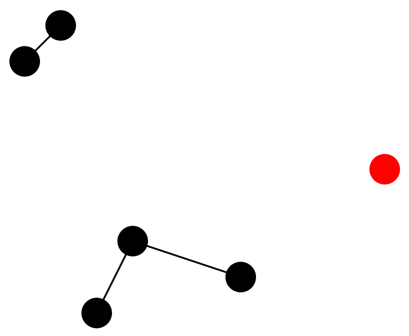
entity $\bar{x}$ is *k-interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



is 1-interior

k-interior

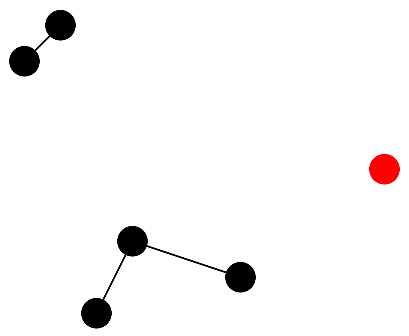
entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



● is 1-interior

k-interior

entity $\bar{x}$ is k-*interior* with respect to X if for every k-partition $\mathcal{P}_k$ of X, $\mathcal{P}_k \cup \{\{\bar{x}\}\}$ is not an adequate $(k+1)$-partition of $X \cup \{\bar{x}\}$



● is 1-interior   but not 2-interior

interiority degree

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

□ $d_{\bar{x}} = 0$

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

- $d_{\bar{x}} = 0$
  - diameter iff

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

- □ $d_{\bar{x}} = 0$
  - diameter iff $\min_{x \in X} dis(\bar{x}, x) \geq \max_{x, x' \in X} dis(x, x')$

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

□ $d_{\bar{x}} = 0$

- diameter iff $\min\limits_{x \in X} dis(\bar{x}, x) \geq \max\limits_{x, x' \in X} dis(x, x')$
- single linkage iff

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

□ $d_{\bar{x}} = 0$

- diameter iff $\min_{x \in X} dis(\bar{x}, x) \geq \max_{x, x' \in X} dis(x, x')$
- single linkage iff
  $\min_{x \in X} dis(\bar{x}, x) \geq \max_{\{X', X''\}} \min_{x' \in X', x'' \in X''} dis(x', x'')$,
  where $\{X', X''\}$ is a non trivial bipartition of X

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

- $d_{\bar{x}} = 0$
  - diameter iff $\min_{x \in X} dis(\bar{x}, x) \geq \max_{x, x' \in X} dis(x, x')$
  - single linkage iff
    $\min_{x \in X} dis(\bar{x}, x) \geq \max_{\{X', X''\}} \min_{x' \in X', x'' \in X''} dis(x', x'')$,
    where $\{X', X''\}$ is a non trivial bipartition of X
- $d_{\bar{x}} = |X| - 1$ diameter and single linkage iff

interiority degree

the *interiority* degree of $\bar{x}$ is

$d_{\bar{x}} =$ the largest k such that $\bar{x}$ is k-interior

$d_{\bar{x}} = 0$ if $\bar{x}$ is not interior

- $d_{\bar{x}} = 0$
  - diameter iff $\min_{x \in X} dis(\bar{x}, x) \geq \max_{x, x' \in X} dis(x, x')$
  - single linkage iff $\min_{x \in X} dis(\bar{x}, x) \geq \max_{\{X', X''\}} \min_{x' \in X', x'' \in X''} dis(x', x'')$, where $\{X', X''\}$ is a non trivial bipartition of X
- $d_{\bar{x}} = |X| - 1$ diameter and single linkage iff $\min_{x \in X} dis(\bar{x}, x) < \min_{x, x' \in X} dis(x, x')$

this combinatorial concept of interiority...

this combinatorial concept of interiority...

defines, for each aggregation criterion, a particular partition of the entity space into a finite number of iso-interiority regions

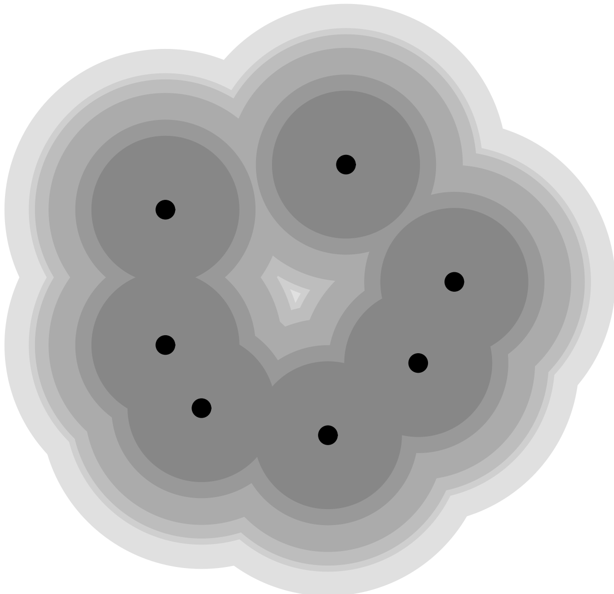
this combinatorial concept of interiority...

defines, for each aggregation criterion, a particular partition of the entity space into a finite number of iso-interiority regions

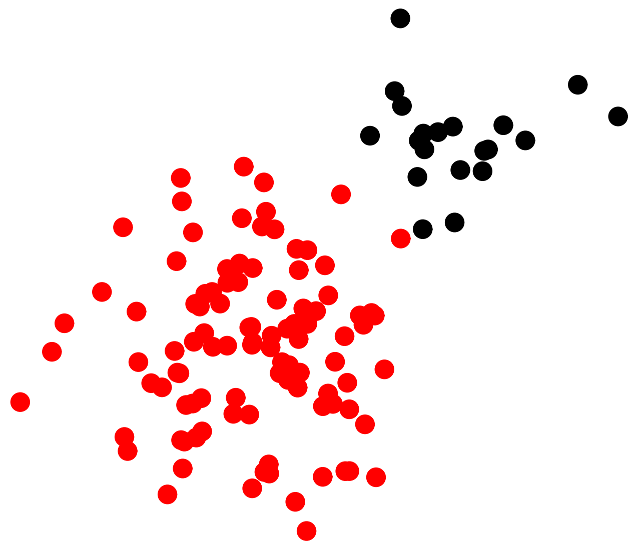
$d : \mathcal{E} \setminus X \to \{0, 1, \ldots, |X| - 1\}$

defines, for each aggregation criterion, a particular partition of the entity space into a finite number of iso-interiority regions

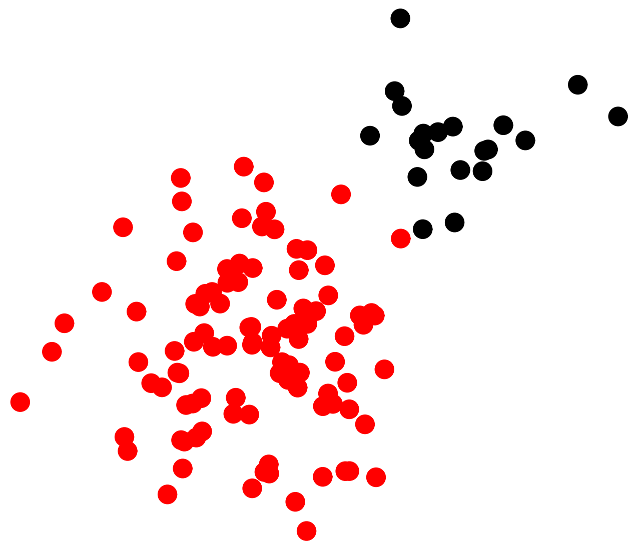
$$d : \mathcal{E} \setminus X \to \{0, 1, \ldots, |X| - 1\}$$



separability index

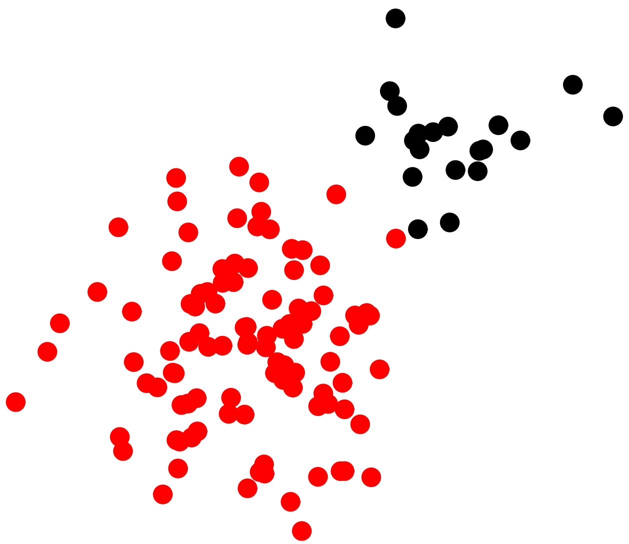


separability index



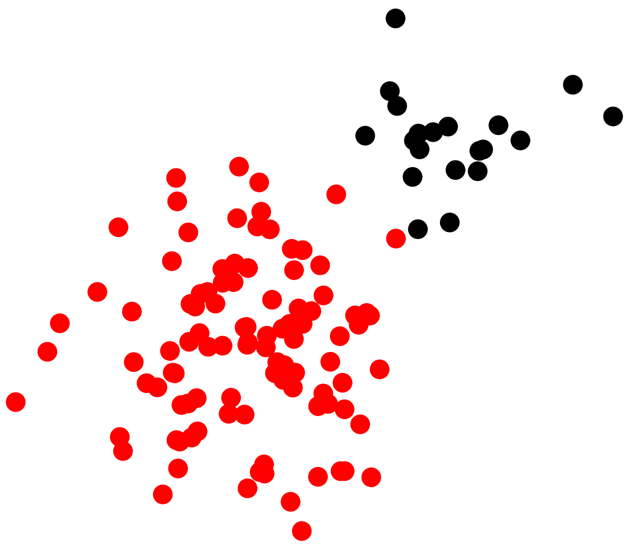
$$D_{\bar{X}} = \{d(\bar{x}) : \bar{x} \in \bar{X} = \{\bullet\}\}$$ *interiority* of $\bar{X}$

separability index



$$D_{\bar{X}} = \{d(\bar{x}) : \bar{x} \in \bar{X} = \{\bullet\}\}$$ *interiority* of $\bar{X}$

$D_{\bar{X}} \rightsquigarrow$ discriminate sep. index $DSI \in [0, 1]$

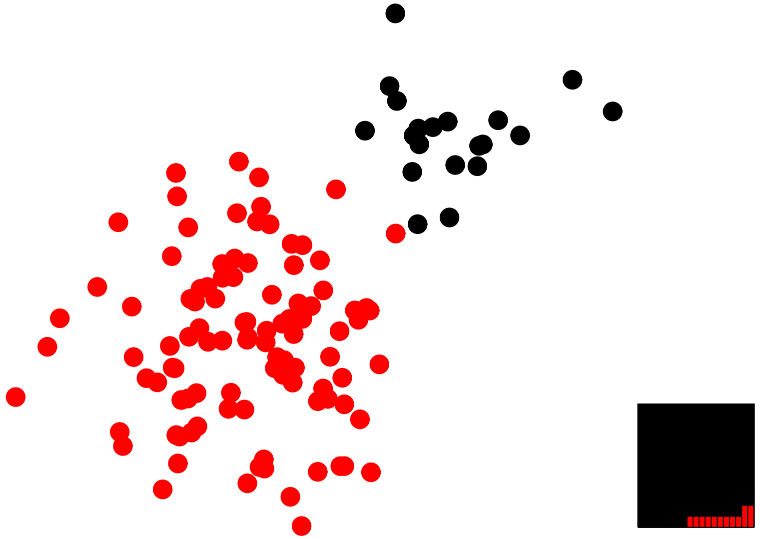


$D_{\bar{X}} = \{d(\bar{x}) : \bar{x} \in \bar{X} = \{\textcolor{red}{\bullet}\}\}$ *interiority* of $\bar{X}$

$D_{\bar{X}} \rightsquigarrow$ discriminate sep. index $DSI \in [0, 1]$

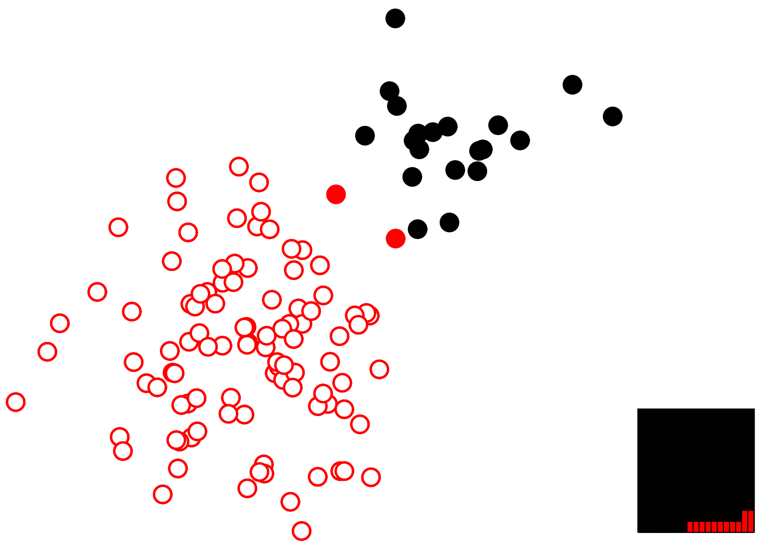
DSI of m points with interiority degree n = DSI of n points with interiority degree m

separability index



$D_{\bar{X}} = \{d(\bar{x}) : \bar{x} \in \bar{X} = \{\bullet\}\}$ *interiority* of $\bar{X}$

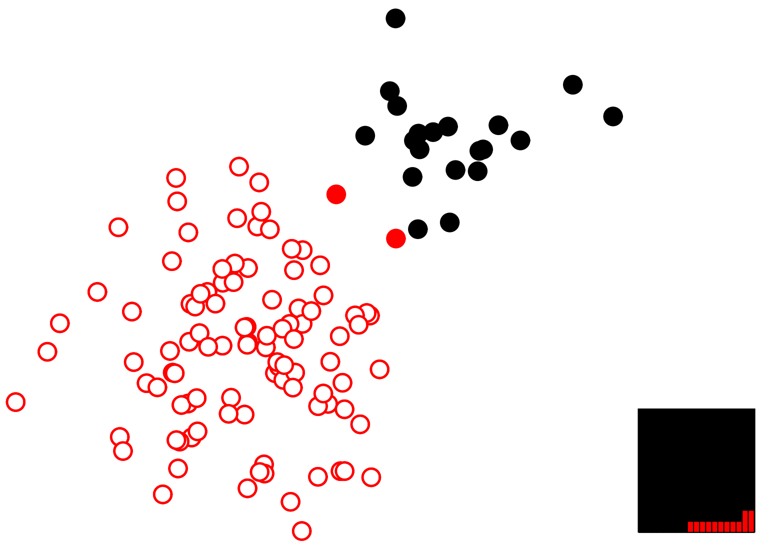
$D_{\bar{X}} \rightsquigarrow$ discriminate sep. index $DSI \in [0,1]$



$$D_{\bar{X}} = \{d(\bar{x}) : \bar{x} \in \bar{X} = \{\bullet\}\}$$ *interiority* of $\bar{X}$

$D_{\bar{X}} \rightsquigarrow$ discriminate sep. index $DSI \in [0, 1]$

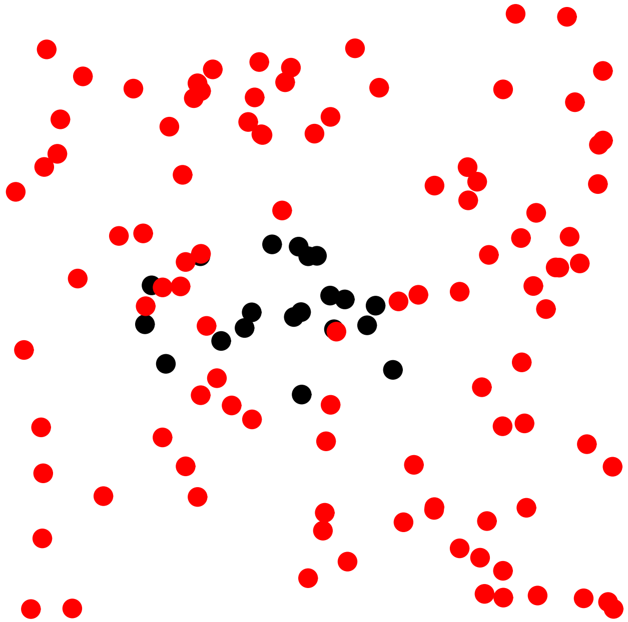
separability index

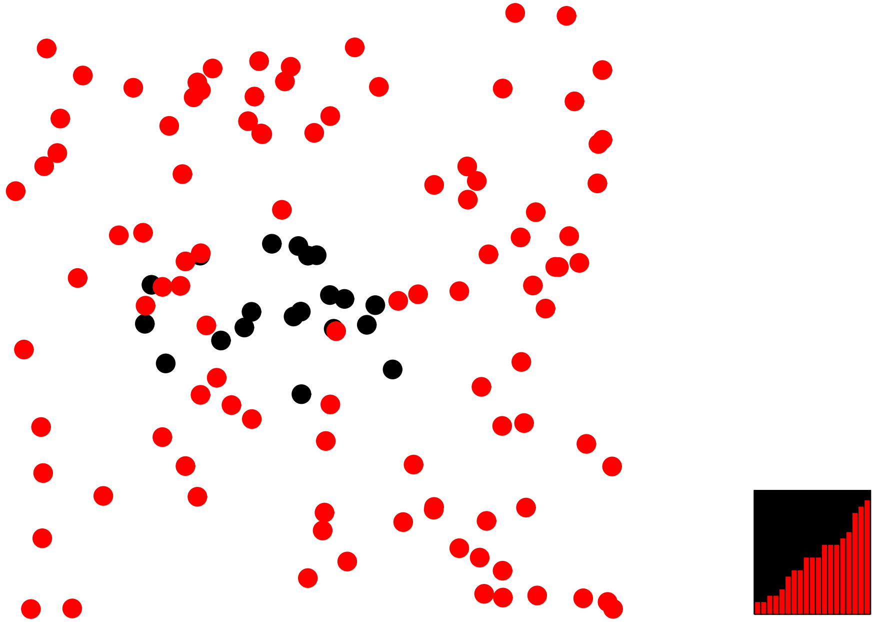


$D_{\bar{X}} = \{d(\bar{x}) : \bar{x} \in \bar{X} = \{\bullet\}\}$ *interiority* of $\bar{X}$

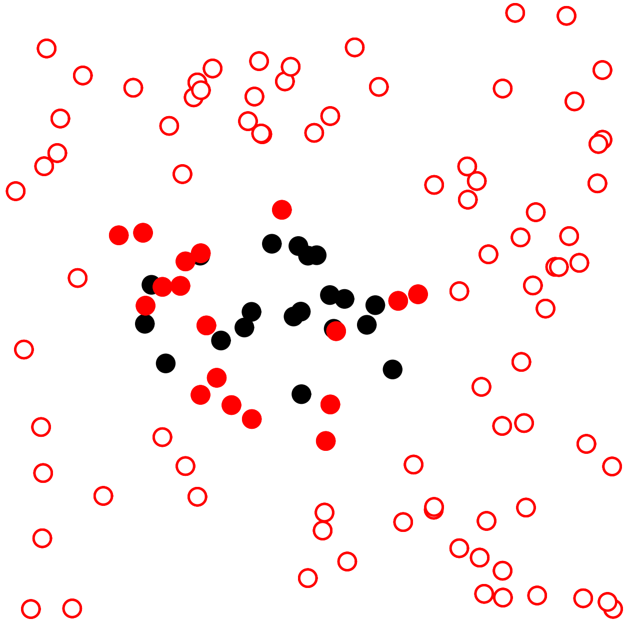
$D_{\bar{X}} \rightsquigarrow$ discriminate sep. index $DSI =$ black/(black+red)

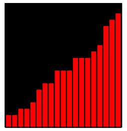
$DSI =$ black/(black+red)



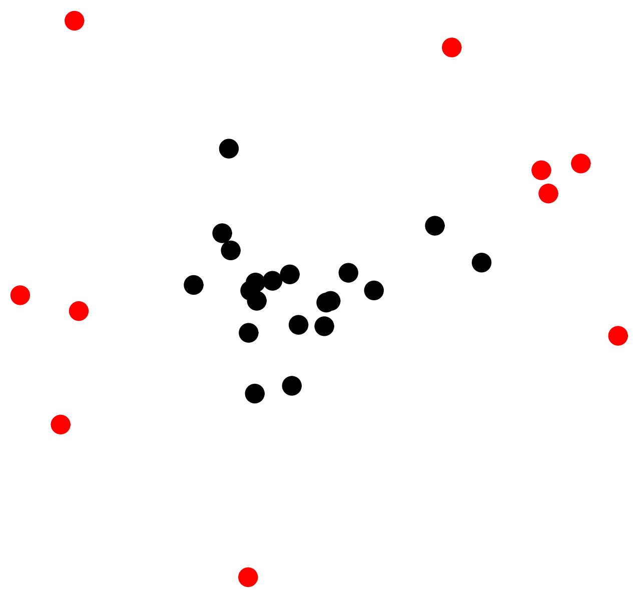


DSI = black/(black+red)

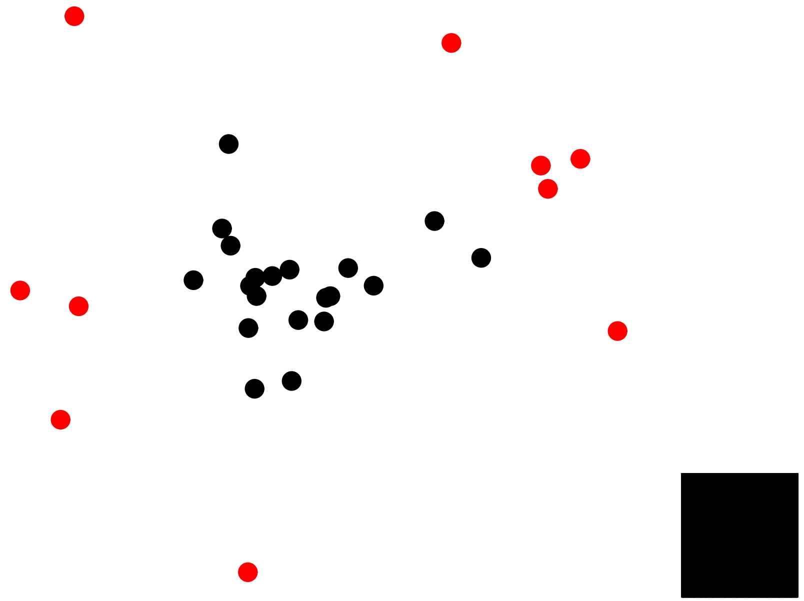




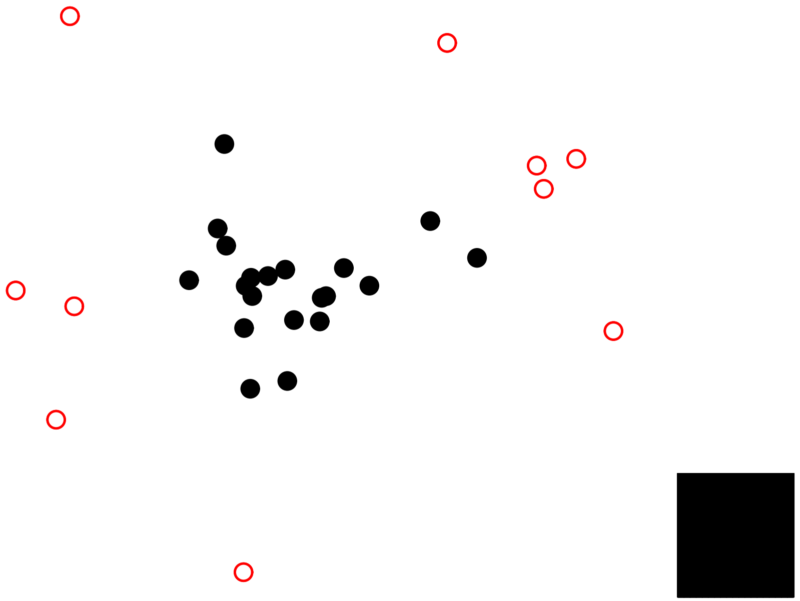
DSI = black/(black+red)



DSI = black/(black+red)



DSI = black/(black+red)

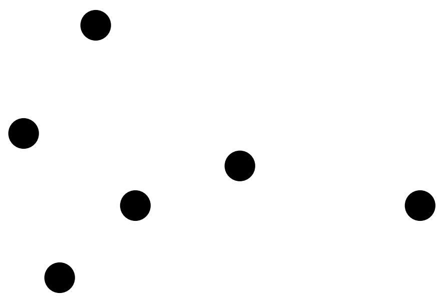


convexity criterion

convexity criterion

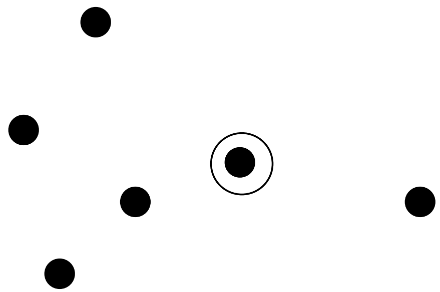
a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



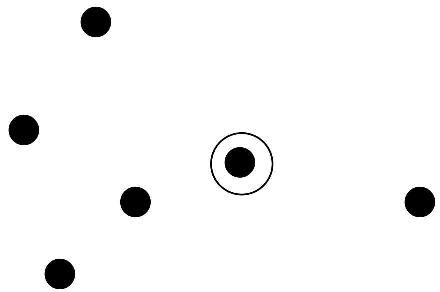
convexity criterion

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



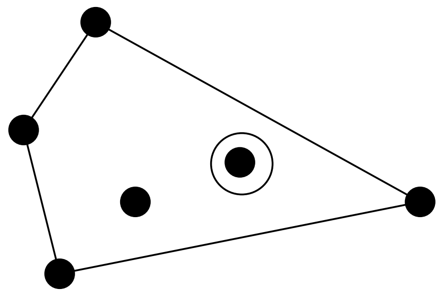
convexity criterion

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



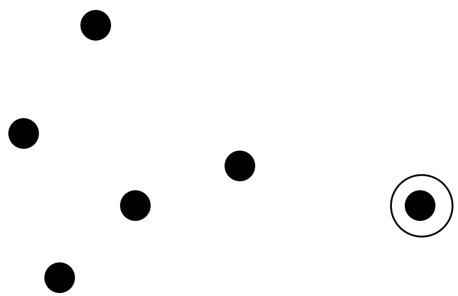
not adequate

a k-partition is adequate if it includes $k - 1$ singletons that are not in the interior of the convex hull of the remaining points

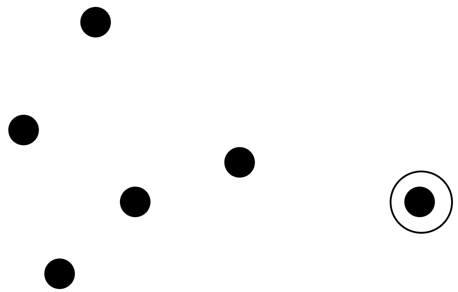


not adequate

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



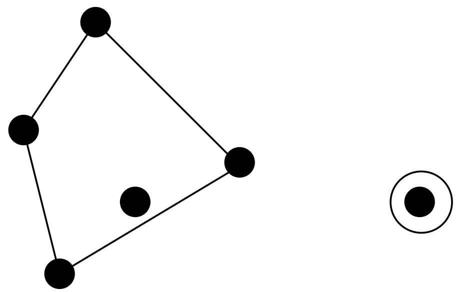
a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



adequate

convexity criterion

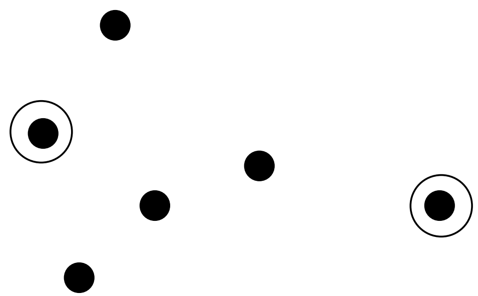
a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



adequate

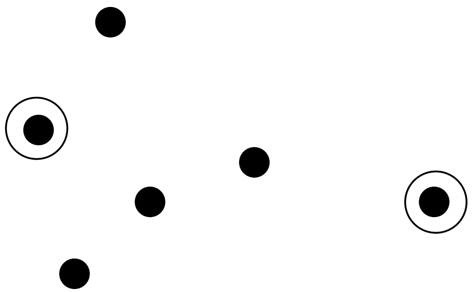
convexity criterion

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



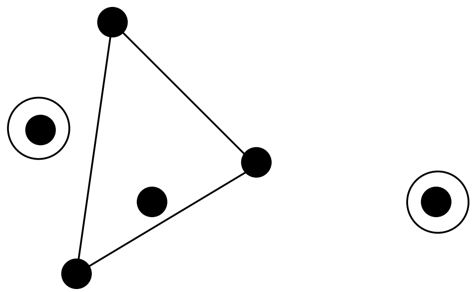
convexity criterion

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



adequate

a k-partition is adequate if it includes $k-1$ singletons that are not in the interior of the convex hull of the remaining points



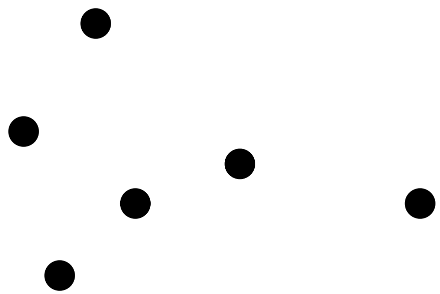
adequate

convexity criterion

$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X

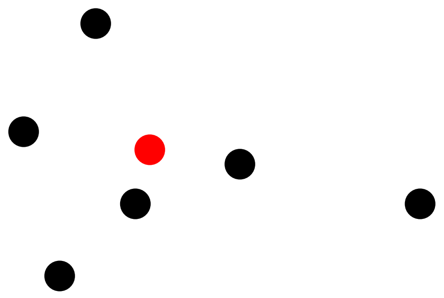
convexity criterion

$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X

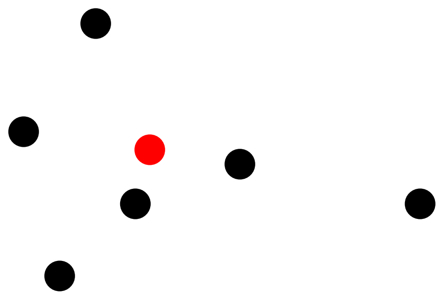


convexity criterion

$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X



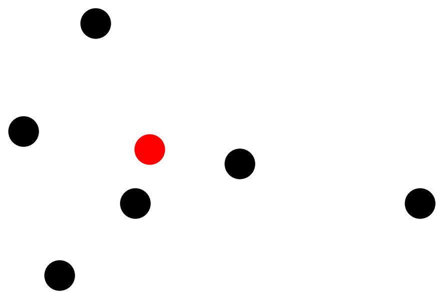
$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X



● is 1-interior

convexity criterion

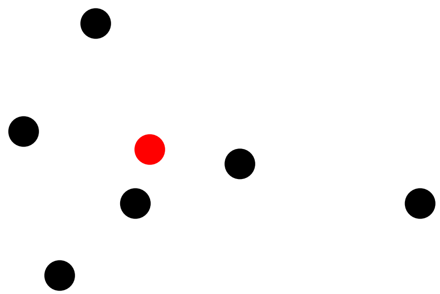
$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X



● is 1-interior, and 2-interior

convexity criterion

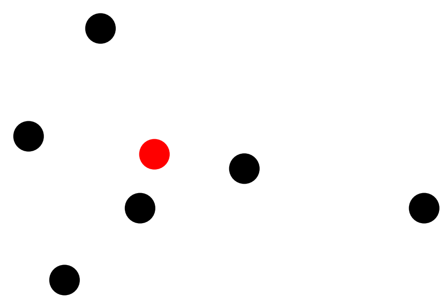
$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X



● is 1-interior, and 2-interior

$d(\bullet)$ is the minimum number of points to be removed from X such that ● is outside conv. hull of the remaining points

$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X



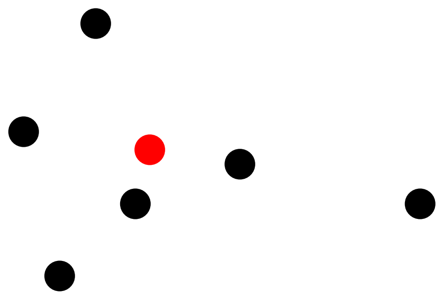
● is 1-interior, and 2-interior

$d(\bullet)$ is the minimum number of points to be removed from X such that ● is outside conv. hull of the remaining points

$d(\bullet) = 2$

convexity criterion

$\bar{x}$ is k-interior if it is inside the convex hull of any $|X| - k + 1$ points of X



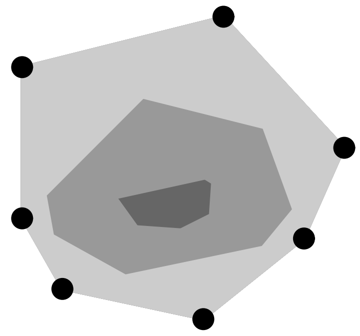
$\bullet$ is 1-interior, and 2-interior

$d(\bullet)$ is the minimum number of points to be removed from X such that $\bullet$ is outside conv. hull of the remaining points

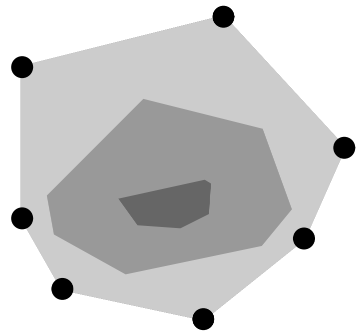
$d(\bullet) = 2$

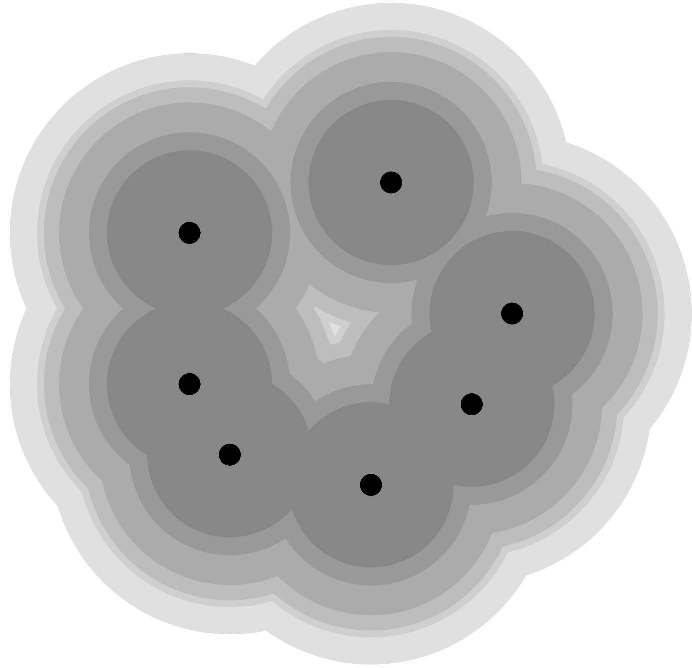
$d(\ )$ - *(halfspace, location, Tukey) depth*

iso-interiority regions



iso-interiority regions





depth $\hookrightarrow$ auto-interiority

depth $\hookrightarrow$ auto-interiority

for $x \in X$, $d(x)$ is the minimum number of points to be removed from X such that x is outside the conv.hull of the remaining points

for $x \in X$, $d(x)$ is the minimum number of points to be removed from X such that x is outside the conv.hull of the remaining points

$d(x) = 1$ iff x is a vertex of the conv.hull(X)

depth $\hookrightarrow$ auto-interiority

for $x \in X$, $d(x)$ is the minimum number of points to be removed from X such that x is outside the conv.hull of the remaining points

$d(x) = 1$ iff x is a vertex of the conv.hull(X)

$D_X = \{d(x) : x \in X\}$ *depth* of X

depth $\hookrightarrow$ auto-interiority

for $x \in X$, $d(x)$ is the minimum number of points to be removed from X such that x is outside the conv.hull of the remaining points

$d(x) = 1$ iff x is a vertex of the conv.hull(X)

$D_X = \{d(x) : x \in X\}$ *depth* of X

should distinguishes between configurations concentrated in the "interior" of the conv.hull, and those which occur mainly on the "margins"...

for $x \in X$, $d(x)$ is the minimum number of points to be removed from X such that x is outside the conv.hull of the remaining points

$d(x) = 1$ iff x is a vertex of the conv.hull(X)

$D_X = \{d(x) : x \in X\}$ *depth* of X

should distinguishes between configurations concentrated in the "interior“ of the conv.hull, and those which occur mainly on the "margins“...

This is relevant as it would allow detection the pattern of the cloud of points

for $x \in X$, $d(x)$ is the minimum number of points to be removed from X such that x is outside the conv.hull of the remaining points

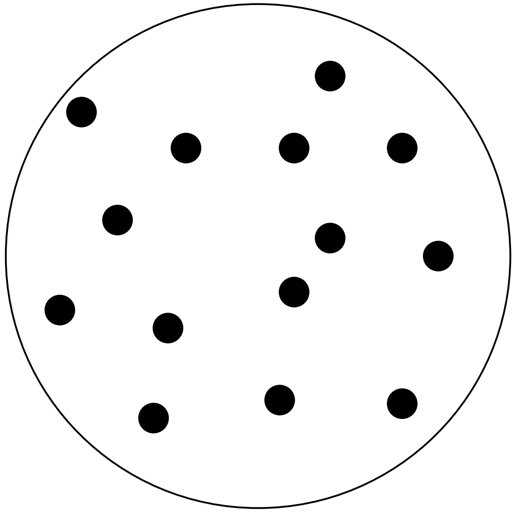
$d(x) = 1$ iff x is a vertex of the conv.hull(X)

$D_X = \{d(x) : x \in X\}$ *depth* of X

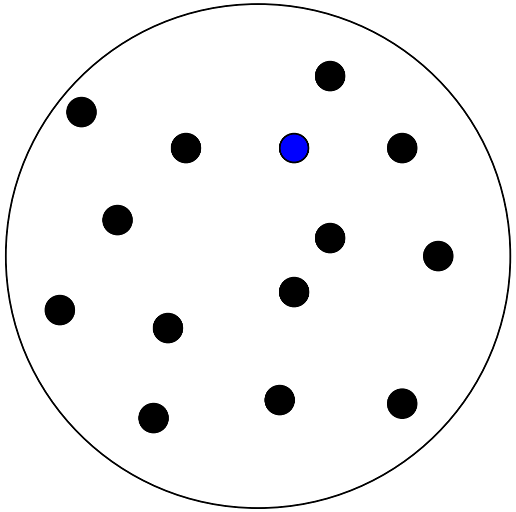
should distinguishes between configurations concentrated in the "interior“ of the conv.hull, and those which occur mainly on the "margins“...

This is relevant as it would allow detection the pattern of the cloud of points $\hookrightarrow$ **issue 2**

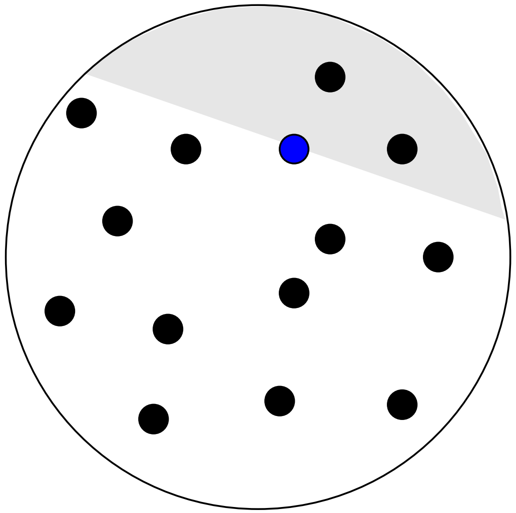
if X is uniformly distributed on an n-ball



if X is uniformly distributed on an n-ball

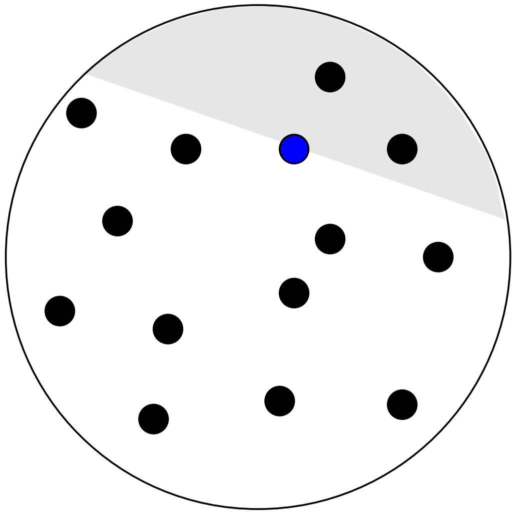


$$\frac{d(x)}{|X|} \rightsquigarrow$$



$$\frac{d(x)}{|X|} \rightsquigarrow \frac{vol(\;)}{vol(B_n)} = V_r(x)$$

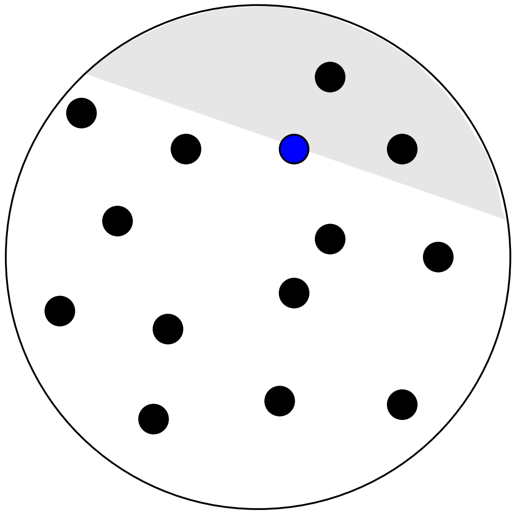
if X is uniformly distributed on an n-ball



$$\frac{d(x)}{|X|} \rightsquigarrow \frac{vol(\;)}{vol(B_n)} = V_r(x)$$

$$E[V_r(x)] = \int_{B_n} \frac{V_r(x)}{vol(B_n)}\, dx$$

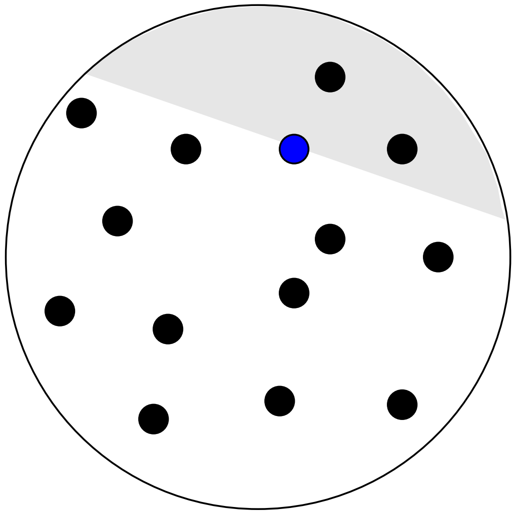
if X is uniformly distributed on an n-ball



$$\frac{d(x)}{|X|} \rightsquigarrow \frac{vol(\;)}{vol(B_n)} = V_r(x)$$

$$E[V_r(x)] = \int_{B_n} \frac{V_r(x)}{vol(B_n)}\, dx = \frac{1}{2^{n+1}}$$

if X is uniformly distributed on an n-ball



$$\frac{d(x)}{|X|} \rightsquigarrow \frac{vol(\quad)}{vol(B_n)} = V_r(x)$$

$$E[V_r(x)] = \int_{B_n} \frac{V_r(x)}{vol(B_n)}\,dx = \frac{1}{2^{n+1}}$$

Result holds if X is uniformly distributed on the closed region of $\mathbb{R}^n$ delimited by an ellipsoid

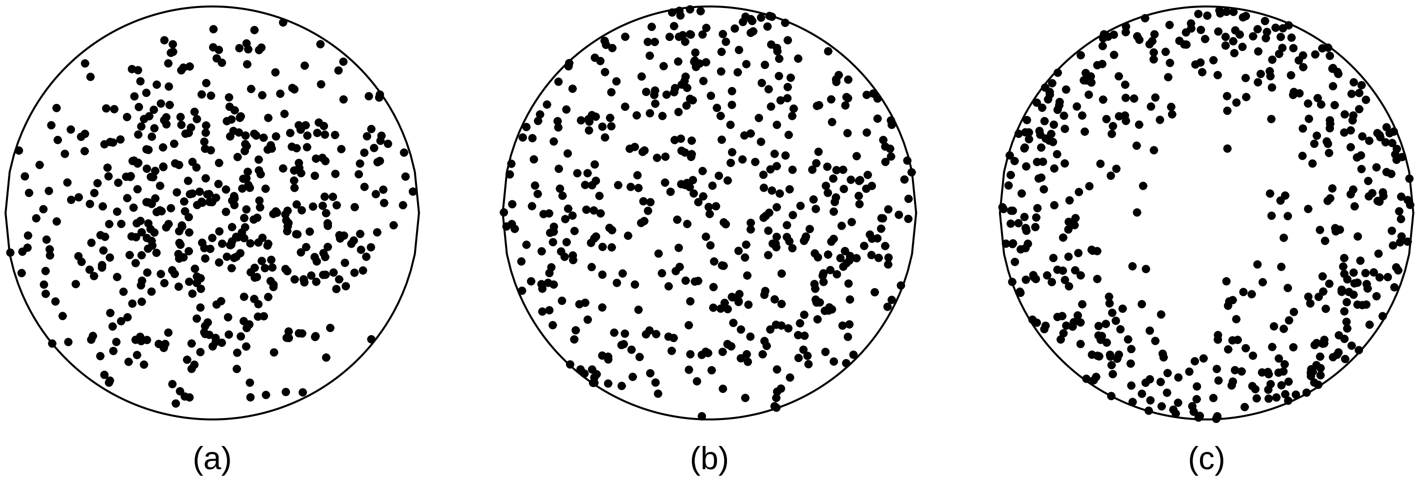
if X is uniformly distributed on an n-ball

$$E[V_r(x)] = \frac{1}{2^{n+1}}$$

if X is uniformly distributed on an n-ball

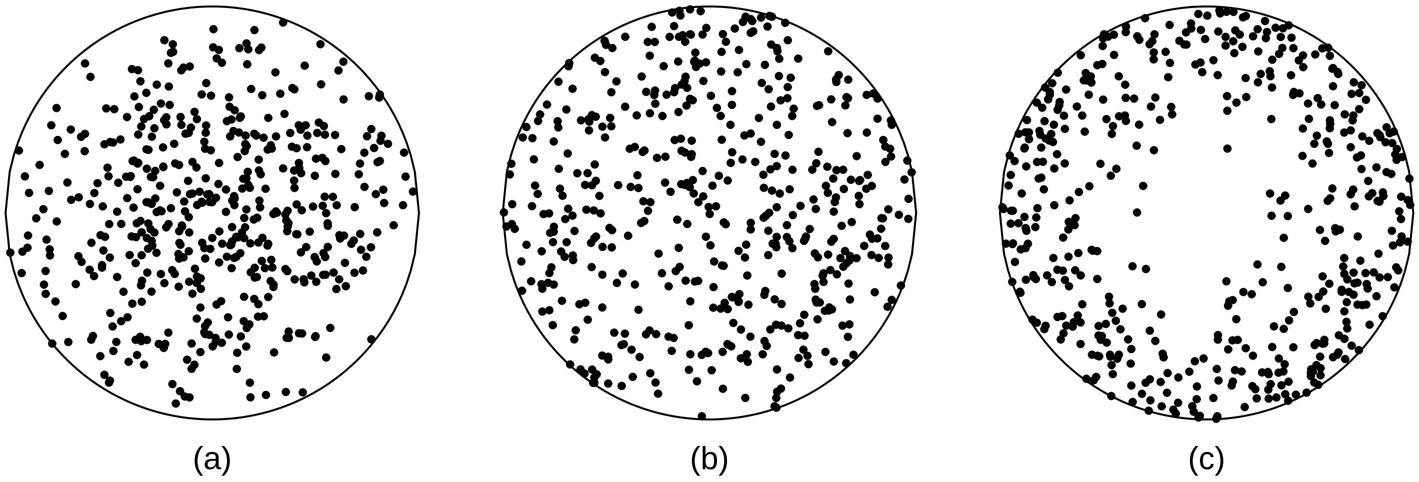
$$E[V_r(x)] = \frac{1}{2^{n+1}} \quad \Longrightarrow \quad \frac{1}{|X|}\sum_x V_r(x) \approx \frac{1}{2^{n+1}}$$

if X is uniformly distributed on an n-ball



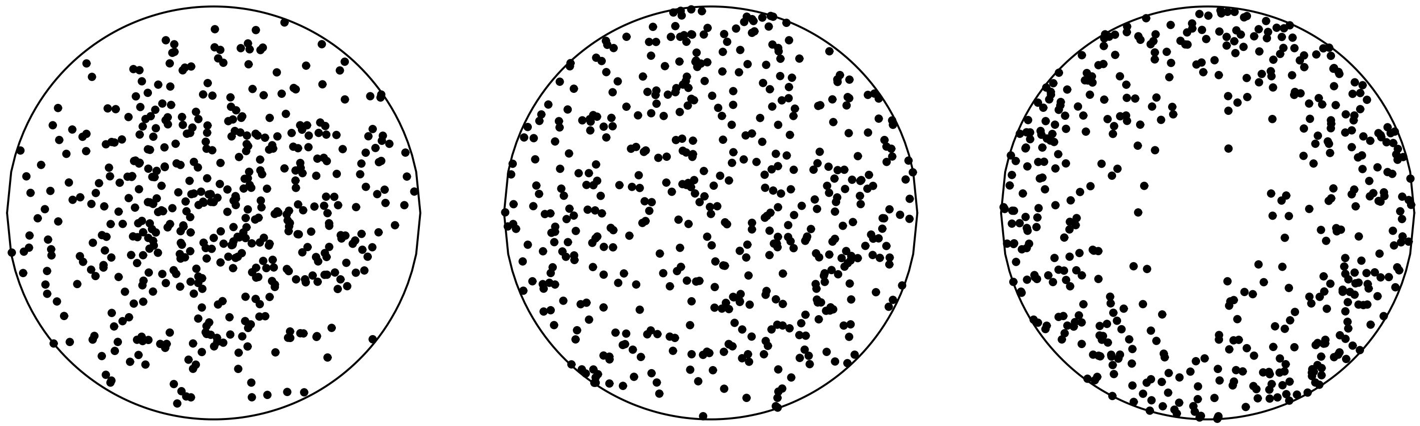
normal uniform LDI

if X is uniformly distributed on an n-ball



	normal	uniform	LDI
$\frac{1}{\lvert X \rvert}\sum_x V_r(x)$	0·1874	0·1247	0·0582

if X is uniformly distributed on an n-ball

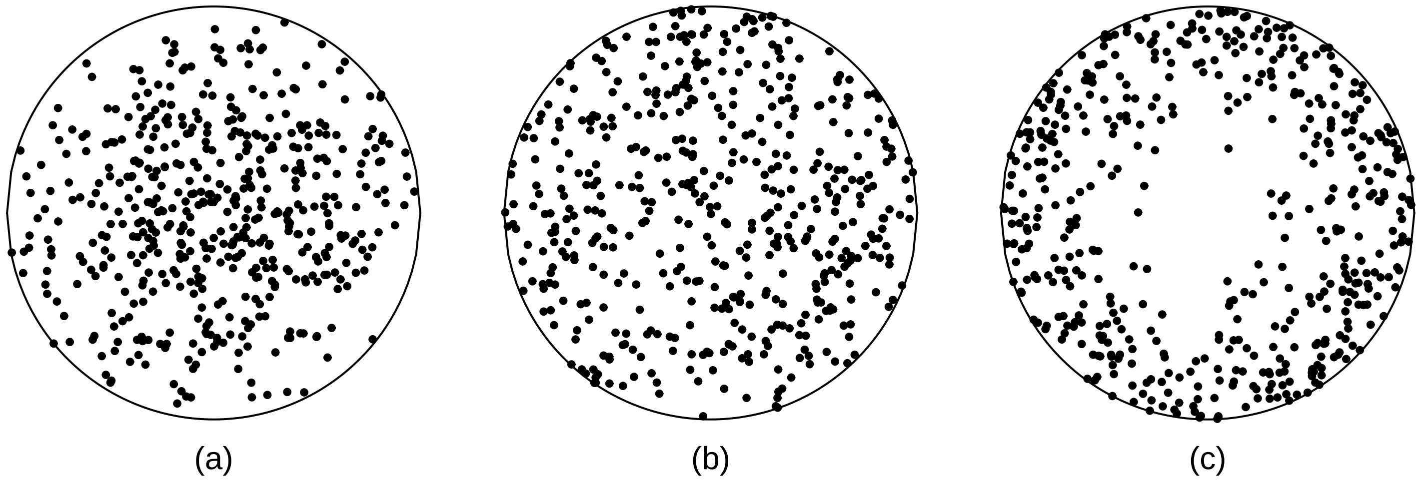


(a) (b) (c)

	normal	uniform	LDI
$\frac{1}{\lvert X\rvert}\sum_x V_r(x)$	0·1874	0·1247	0·0582

$$E[V_r(x)] = \frac{1}{2^{n+1}} = \frac{1}{2^3} = 0.125$$

if X is uniformly distributed on an n-ball



	normal	uniform	LDI
$\frac{1}{\lvert X\rvert}\sum_x V_r(x)$	0·1874	0·1247	0·0582

$$E[V_r(x)] = \frac{1}{2^{n+1}} = \frac{1}{2^3} = 0.125$$

even for small $|X|$, $\frac{1}{|X|}\sum_x V_r(x) \approx E[V_r(x)] = \frac{1}{2^{n+1}}$, for uniformly distributed X