

# Introduction to Information Retrieval

<http://informationretrieval.org>

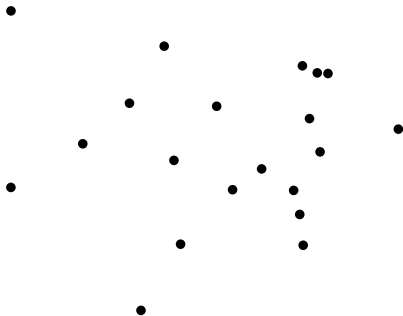
## IIR 16: K-Means Example

Hinrich Schütze

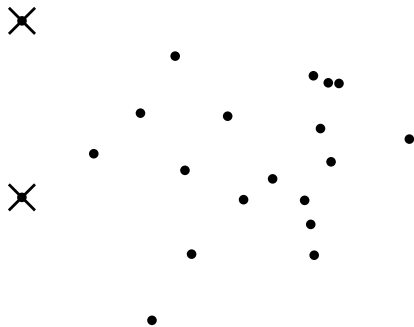
Institute for Natural Language Processing, Universität Stuttgart

2009.06.16

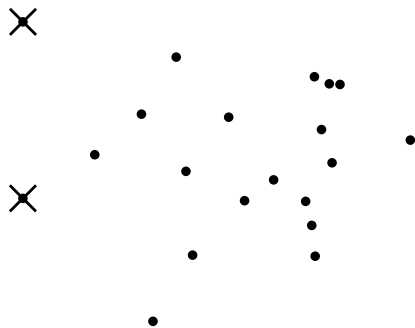
# Set of points to be clustered



# Random selection of initial cluster centers

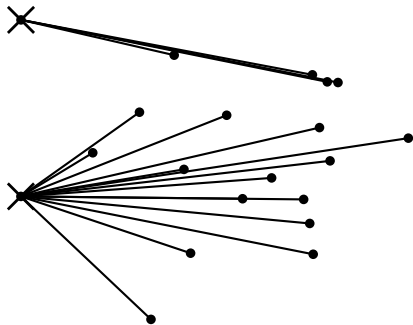


# Random selection of initial cluster centers

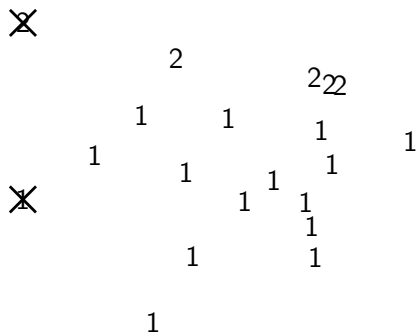


Centroids after convergence?

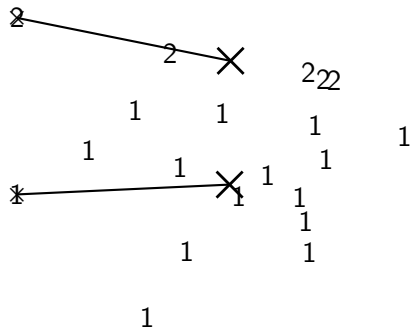
# Assign points to closest center



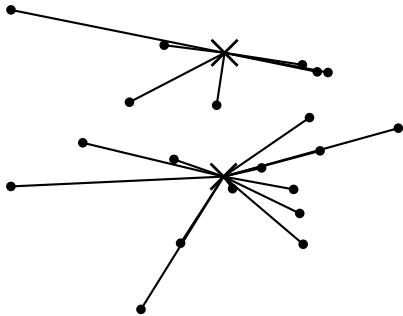
# Assignment



# Recompute cluster centroids

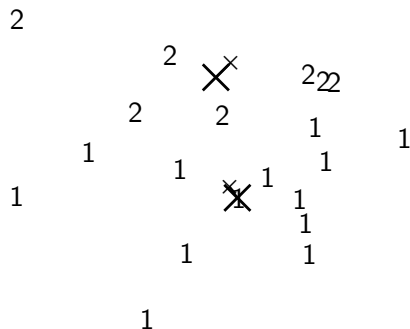


# Assign points to closest centroid

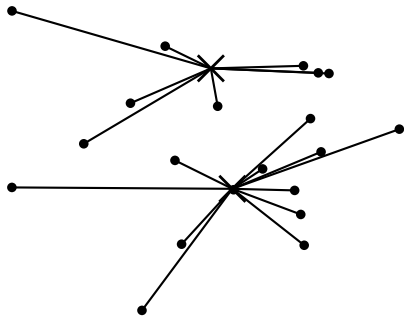




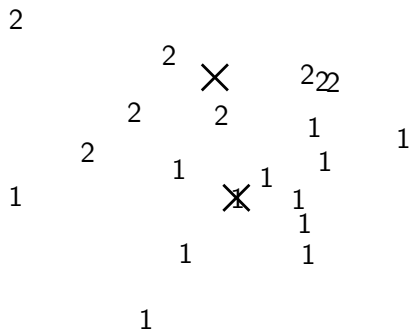
# Recompute cluster centroids



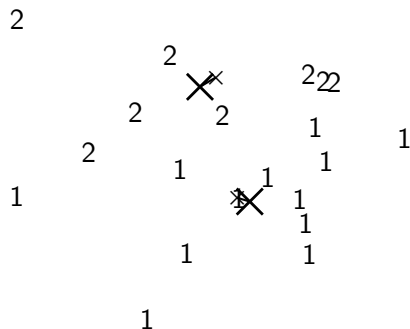
## Assign points to closest centroid



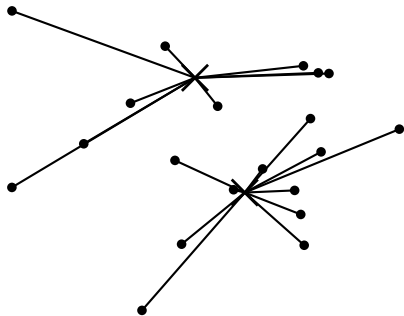
# Assignment



# Recompute cluster centroids

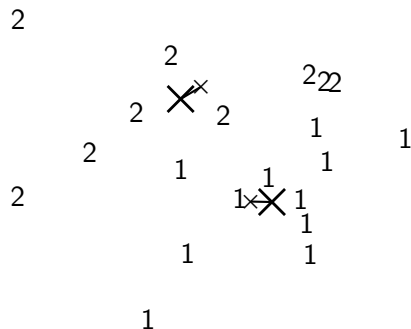


## Assign points to closest centroid

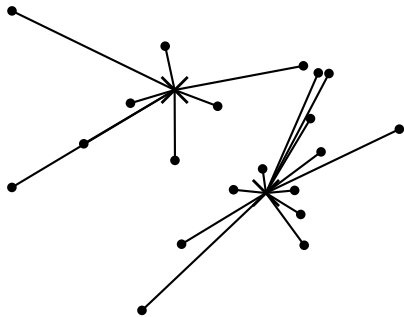




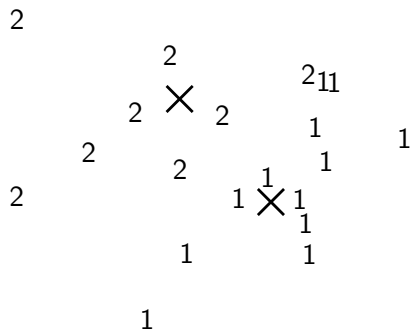
# Recompute cluster centroids



## Assign points to closest centroid

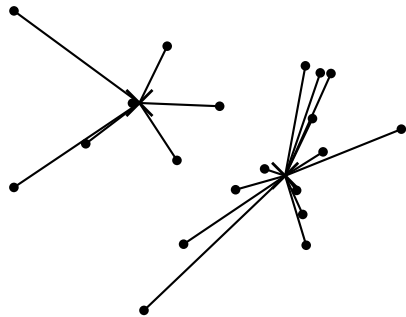


# Assignment

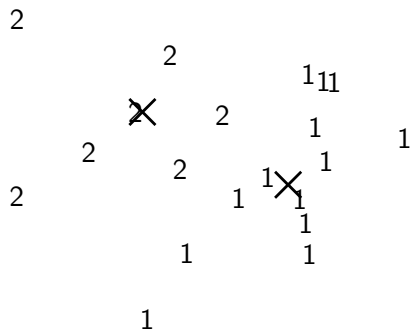




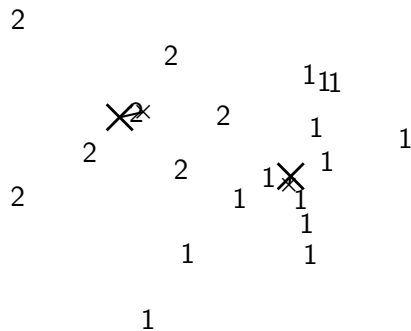
## Assign points to closest centroid



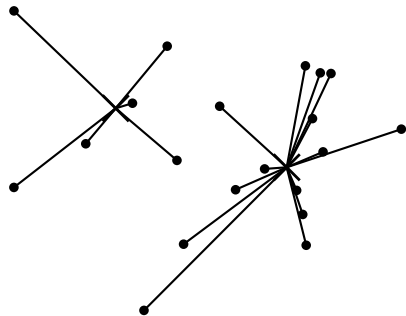
# Assignment



# Recompute cluster centroids

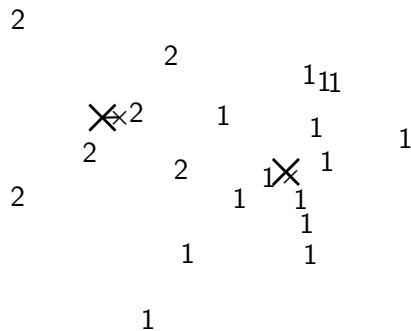


# Assign points to closest centroid





# Recompute cluster centroids



# Centroids and assignments after convergence

