

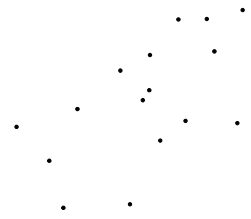
Ball Trees and Metric Trees

The Auton Lab
Carnegie Mellon University

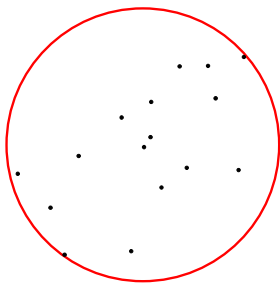


www.autonlab.org

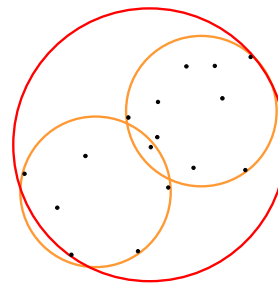
A Set of Points
in a metric
space



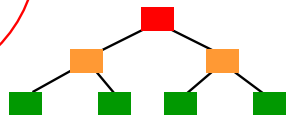
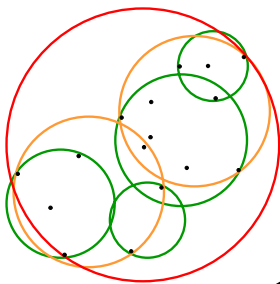
Ball Tree root
node



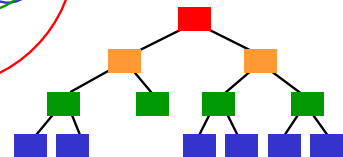
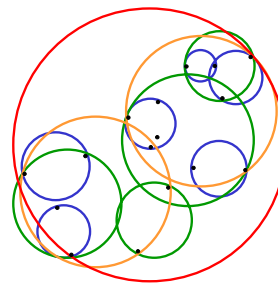
A Ball Tree

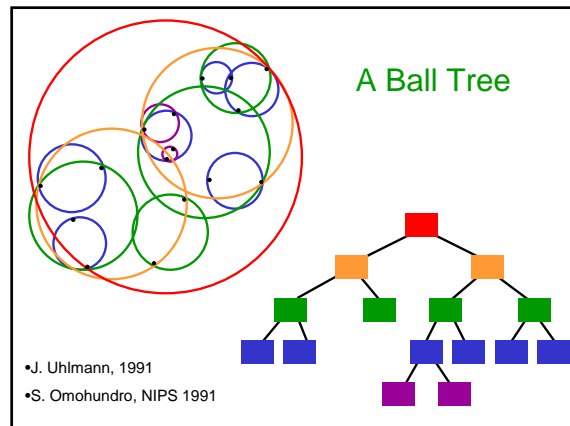
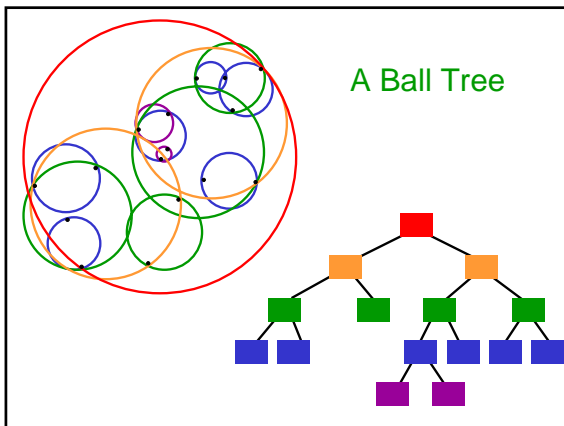


A Ball Tree



A Ball Tree





Ball-trees: properties

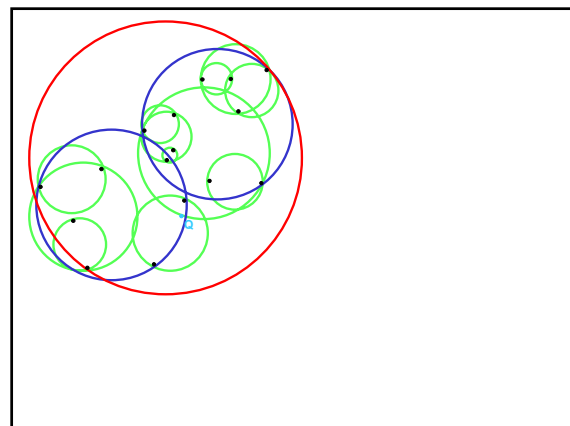
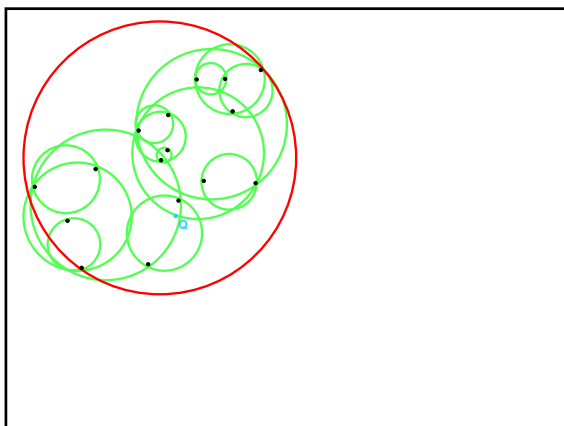
Let Q be any query point and let x be a point inside ball B

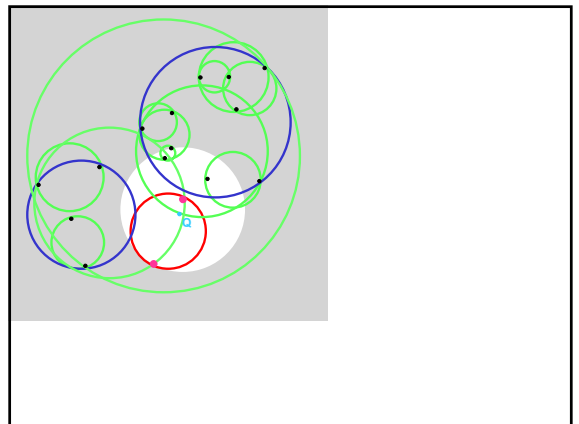
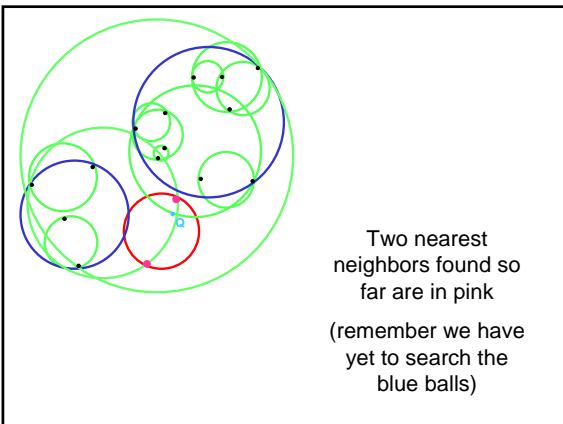
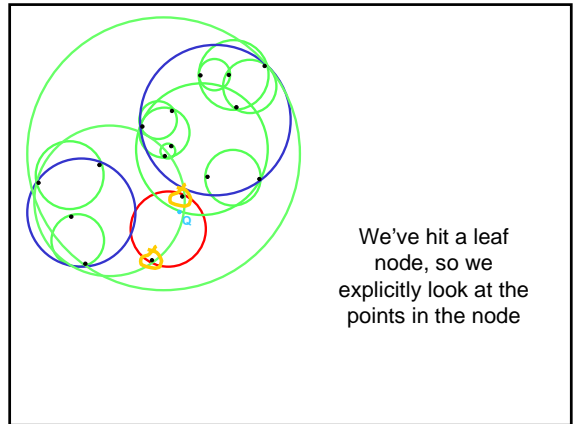
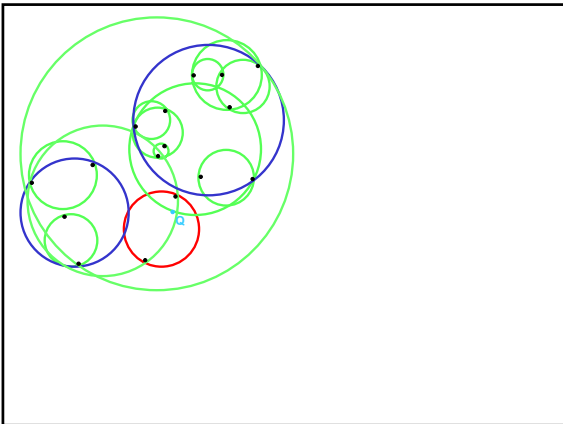
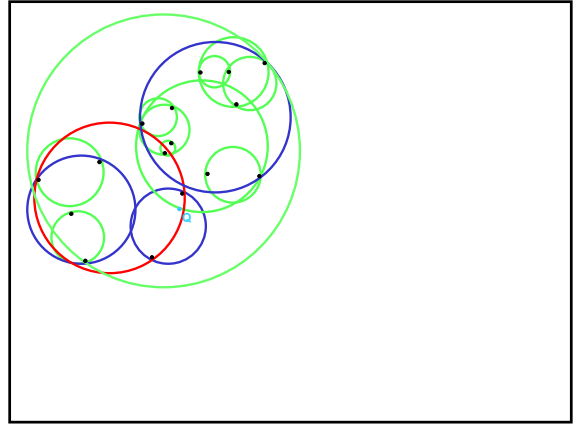
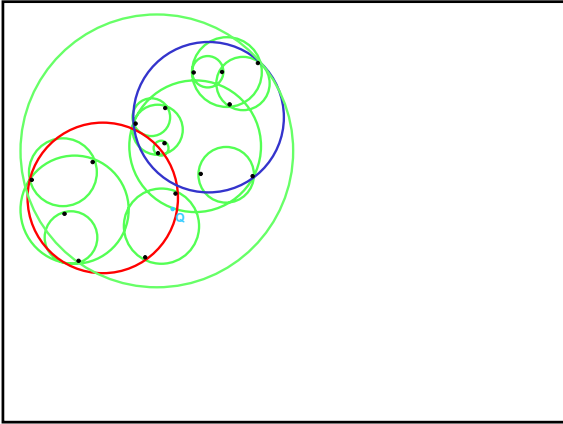
$$|x - Q| \geq |Q - B.\text{center}| - B.\text{radius}$$

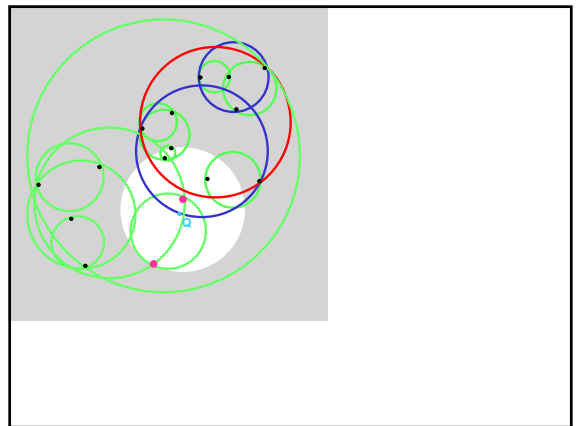
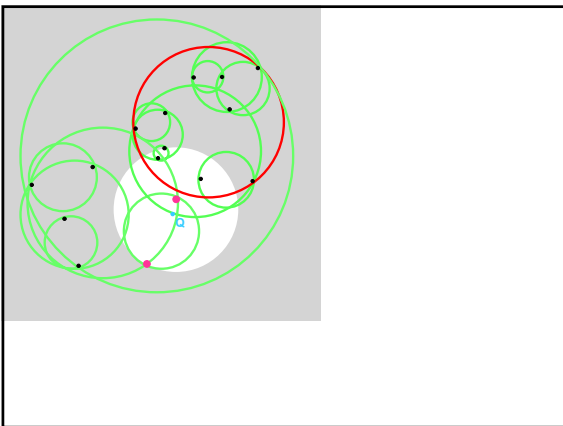
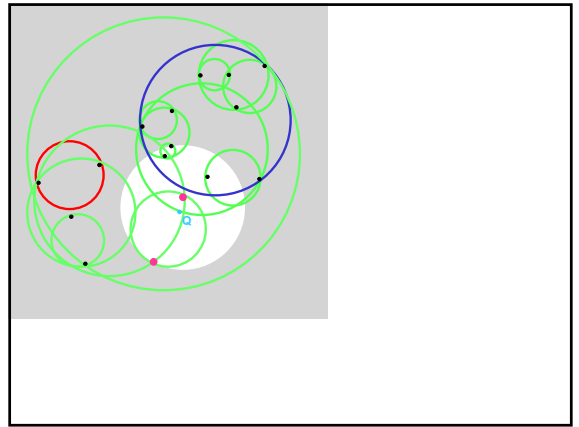
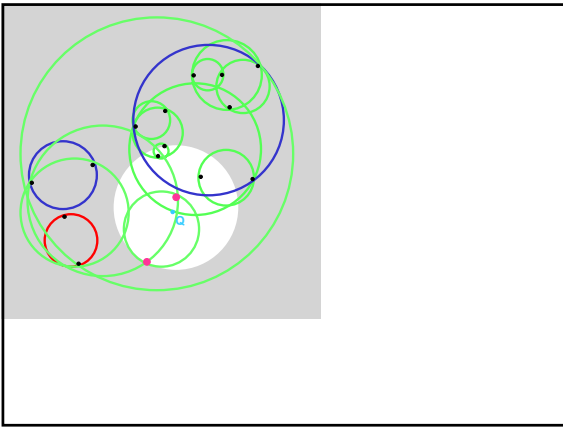
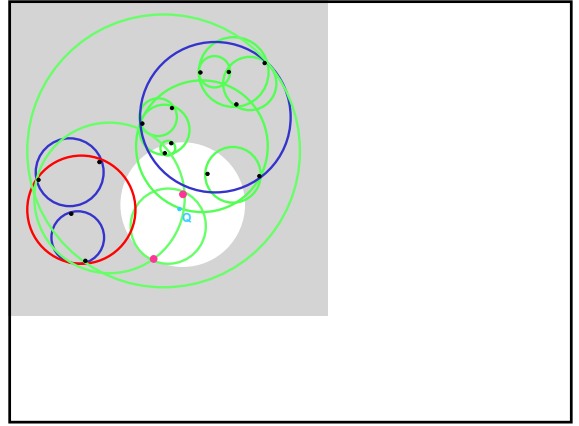
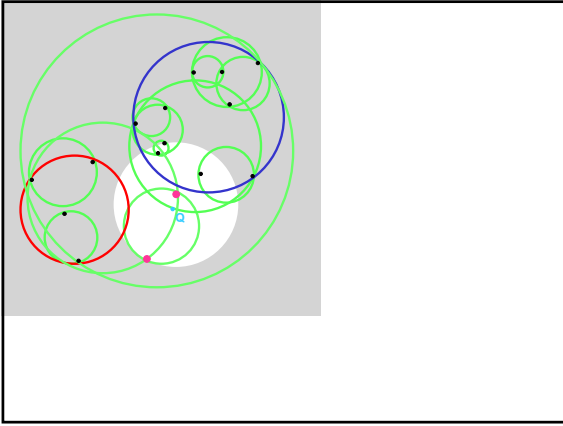
$$|x - Q| \leq |Q - B.\text{center}| + B.\text{radius}$$

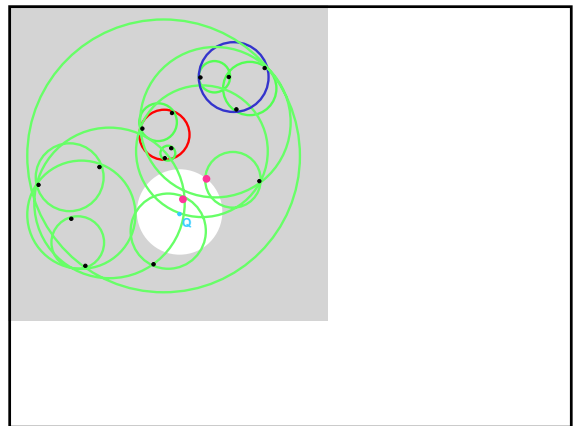
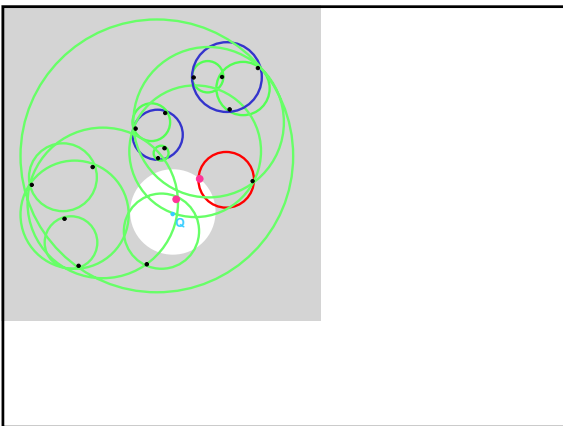
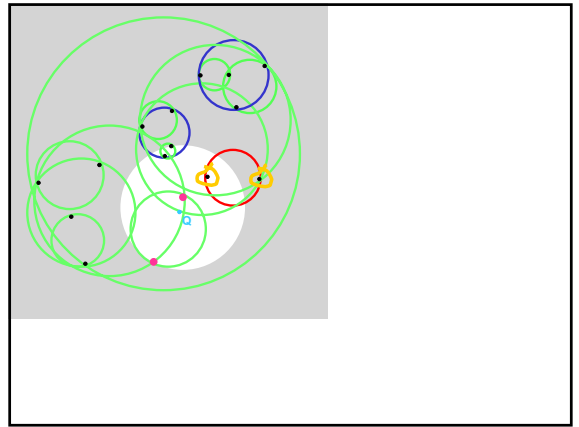
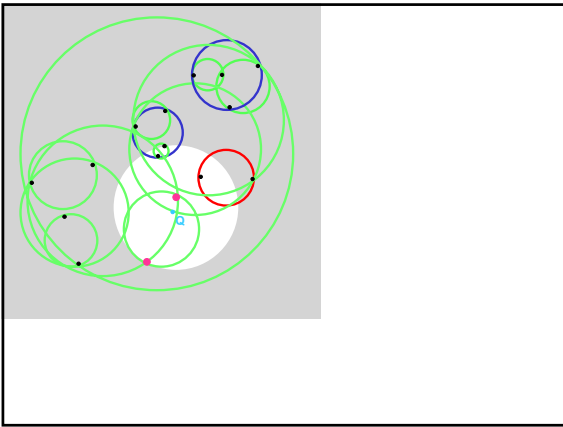
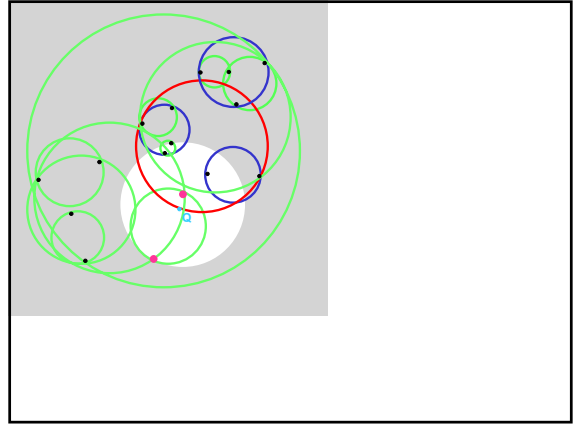
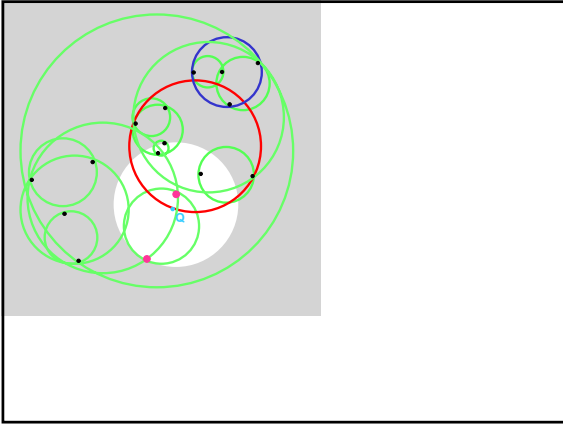
Goal: Find out the 2-nearest neighbors of Q .

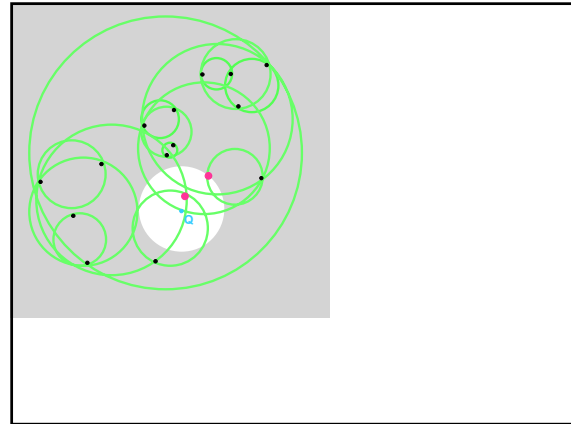
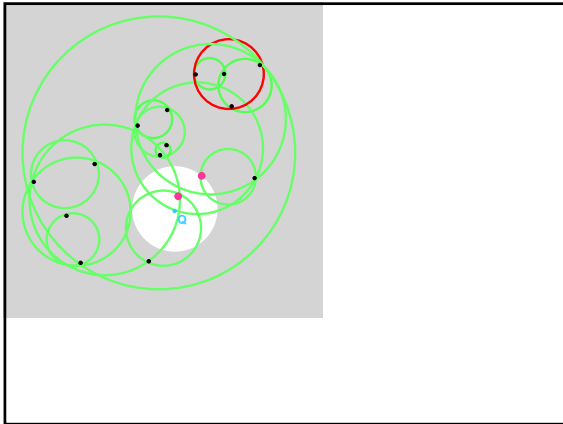
•J. Uhlmann, 1991
•S. Omohundro, NIPS 1991











What's a metric tree got that a kd-tree hasn't?

- Resistance to higher dimensions
- Can work with non-Euclidian metric (you only need to respect the triangle inequality)
- Many non-metric similarity measures can be bounded by metric quantities.
- Downside: in lower dimensions often slower (the balls aren't as small as the hyper-rectangles)

Generalizations

Kernel Regression [Deng and Moore, 1995]
 Kernel Density Estimation [Deng and Moore, 1995]
 Kernel-based Bayes Classifiers [Moore and Schneider, 97]
 Locally weighted regression [Moore, Schneider, Deng, 97]
 K-Means [Pelleg and Moore, 99], [Moore 2000]
 Gaussian Mixtures [Moore, 1999]
 Kernel-density-based clustering [Wong and Moore, 2002]
 2-point function computation [Gray and Moore 2001]
 N-point function computation [Gray and Moore 2001]
 Categorical Data [Moore and Lee, 1998]
 KNS2, KNS3 [Ting, Moore, Gray, NIPS 2003]
 Also work by Priebe, Ramakrishnan, Schaal, D'Souza, Elkan. . .
 Papers (and software): www.autonlab.org