A Big(ger) Picture

High dimensional problems:

- nearest neighbor: l_1/l_2, l_∞, Hausdorff
- proximity: by reduction to NN
- clustering: only few subquadratic algorithms known
- regression: almost nothing known
- graph problems: MST, matching
- ...

Different tools

- random projections
- embeddings (l_∞, Hamming cube)

  (although sometimes can use existing reductions)

Low vs High Dimensional CG

	Low dimensional	High dimensional
Solutions	Exact and Approximate	Approximate
Tools	Combinatorial Geometry	Functional Analysis
Randomization	Las Vegas	Monte Carlo

Low vs High ctd.

Not all CG problems seem useful when dimension is high (e.g. convex hull).

However, there is additional variation:

- norms/metrics:
  - shape metrics (Hausdorff, Matchings, EMD, Frechet); transformations
  - string metrics (Levenstein etc)
- input form
  - substrings of a string: Approximate Pattern Matching
  - lattice: Closest Vector Problem
- high dimension is an additional source of intractability

Concrete open problems

- lower bounds:
  - unconditional ?
  - conditional ? One can prove:

    If a subset query problem can be done efficiently, then we can find a k-clique in a n-node graph in poly(n) time, provided we have exp(k) (not exp(n)) advice.

    What does this mean ??

- upper bounds:
  - NN in normed spaces (unification of norms)
  - other metrics (edit distance ?)