Data $x_1 \dots x_n$ | $y_1 \dots y_n$

Two Things:

1) Split Function

2) Prediction Function

=> issue = overfitting

3) Stopping Criteria

- Error Threshold - n elements in leafnode
- Variance Threshold - Δ Error
- never

Bootstrapping = subsample data, learn your tree, do it a "thousand times", then aggregate your trees

Human Connectome

=> network of elements and connections forming the human brain

=> there have been studies at macroscopic level (lobes, landmarks, etc.) and microscopic level but little done on connectivity patterns

Connectome = neural elements + connections

Two significant challenges:

1) complex organ w/ lots of structurally distinct heterogenious yet interconnected components

2) hard to define basic elements

Connection Scales

1) microscale: Single neurons + connections
- currently infeasible, size would be 10^{11} neurons + 10^{15} connections
- (= alterations of single synapses/cells has not been shown to be important) variable big + dynamic
- this scale is subject to rapid plastic changes

2) Macroscale: brain regions + neuronal popul.
- no universal parcellation scheme
- potentially very relevant (evidence from smaller mammal brains)
- many valid experimental methods exist

Definition: Segregated brain region
- all structural elements share similar

long range connectivity patterns

Most Promising Experimental Route:

- Correlated use of noninvasive structural & functional imaging methods (10^5 elem)

Mesoscale: cortical minicolums/ local neuronal populations

- currently impossible to trace all minicolums, might be able to do it region by region
- no way to coordinate across macro-regions (10^7-10^8 elem)

Variability & Development

⇒ above microscale there is little variability (stable forever)

⇒ variation between individuals (at all levels) and during growth

Steps Towards the Connectome:

1) DWI then probabalistic tractography of thalamocortical tracts & cortico-cortical interareal pathways

Result: voxel wise all to all structural connectivity matrix

2) correlation analysis of spatially registered resting activity/activation data (fMRI, MEG)

Result: voxel wise all to all functional connectivity matrix

3) Cluster analysis between results of step 1 and step 2

Result: Structure-function related regions identified

4) Compare step 3 results w/ macaque data

5) Validate strongest predictions generated by assembling combined structural-func. connectivity matrix

Potential Impact:

⇒ help map structure to function

⇒ brain activity data + Connectome will enable a mapping

- will require additional research & experimentation
- will enable comparison & integration of all the currently existant models

Affinity Propagation

Paper 1 - Toronto.edu

Affinity propagation - simultaneously considers all data points as exemplars

Two messages passed between data points: responsibilities ($r(i,k)$)
availabilities ($a(i,k)$)

Pseudo-Code:

Input = set of pairwise similarities → how to embed graph data to produce pairwise similarities?

$\{s(i,k)\}$ = how well suited data point k is as an exemplar for i

Initialize: all availabilities = 0

Repeat:

$$\forall i,k \quad r(i,k) = s(i,k) - \max_{k' \neq k} [s(i,k') + a(i,k')]$$

$$\forall i,k \quad a(i,k) = \begin{cases} \sum_{i' \neq i} \max[0, r(i',k)] & \text{for } k = i \\ \min\left[0, r(k,k) + \sum_{i' \notin \{i,k\}} \max[0, r(i',k)]\right] & \text{for } k \neq i \end{cases}$$

Output = $\hat{C} = (\hat{c}_1, \hat{c}_2, \ldots)$

where $\hat{c}$ = exemplar of each data point

Preference (p): global & shared, typically $s(i,i) = p$

⇒ lower p = penalize more clusters

⇒ higher p = enable more clusters

Note matlab implementation @ psi.toronto.edu

⇒ messages must be damped to prevent oscillation

⇒ successful damping factor $\lambda = .9$

Examples w/ Affinity propagation:

1) Olivetti Dataset: $s(i,k)$ calculated as similarity of image i to image k by neg sum of squared pixel differences

2) mushroom Dataset: $s(i,k)$ calculated to be # of matching attributes

3) UPS digits: similarity = negative sum of squared pixel difference

4) Netflix movies + customer ratings : sparse matrix w/ similarity det to similarity between user movie rankings

Paper 2:

Markov Clustering vs Affinity Prop.
for partitioning protein interaction graphs

=> proteins represented as networks
=> Goal: find clustered areas w/ in protein
=> evaluated on weighted and unweighted graphs
=> Conclusion: MCL better than AP, more Robust, especially better on unweighted graphs

MCL
=> considers "connectivity properties" of the underlying network
=> high degree of noise tolerance

AP
=> outperforms vertex substitution heuristic on large problems

- MCL simulates random walks on the interaction network using expansion and inflation
  1) Loops added to graph => weight = max (all edges)
  2) Graph w/ Loops translated to Markov matrix
  => represents prob of random walk of length n between any two nodes by raising it to the n power } inflation
  3) Repeat expansion + inflation until graph is in subsets

- AP chooses exemplars
  1) all nodes considered exemplars according to preference values
  2) repeatedly calculates responsibility and availability
  => damping factor controls oscillations

On unweighted network => AP clusters were more fragmented
On weighted graphs => less noticeable diff but AP also performs worse (lower Acc)