Information flow on trees

Elchanan Mossel

Microsoft Research

Broadcasting on the tree

Given:

- A Markov-matrix: $M_{i,j} = \mathbf{P}[i \to j] (i, j = 1, \ldots, m)$, We assume that M is ergodic (irreducible and aperiodic).

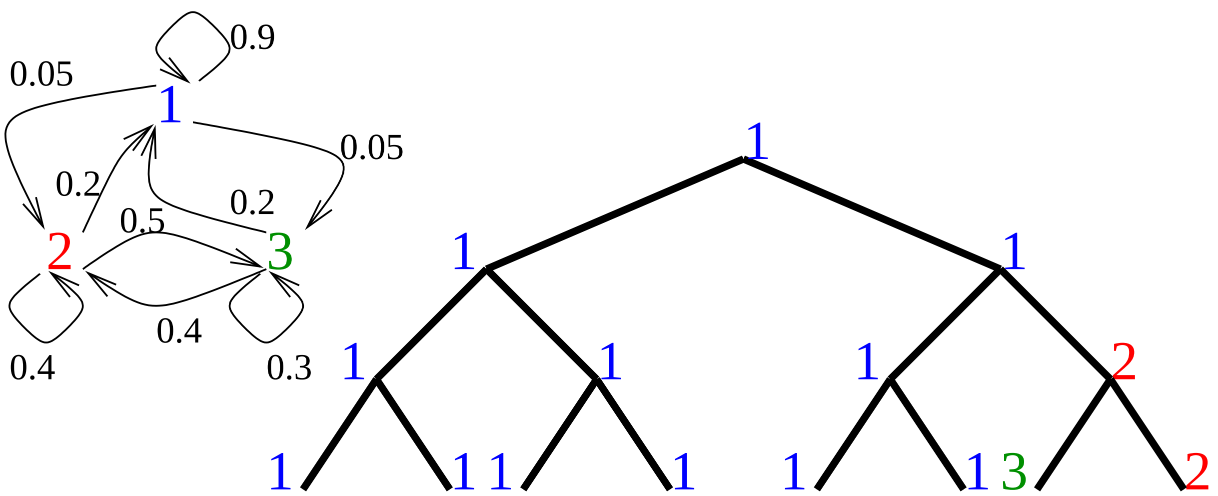
- An initial distribution $\pi_i (i = 1, \ldots, m)$,

- A tree T rooted at r.

For each $i = 1, \ldots, m$, we look at $\{\sigma_v\}_{v \in T}$, where:

- $\mathbf{P}[\sigma_r = i] = \pi_i$ for the root r.

- $\mathbf{P}[\sigma_v = j | \sigma_{v'} = i] = M_{i,j}$ independently for all pairs (v', v) of a parent and a child.



Interpretations

- Broadcasting on the tree. Each edge is an independent copy of the channel M.

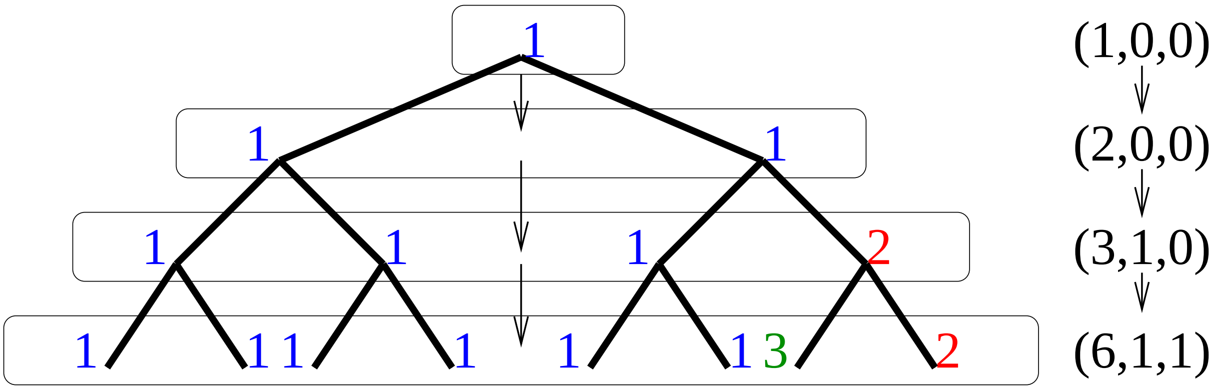
- Propagation of a gene

- The Ising and Potts models on the tree (with external field).

Two Markov chains

Let L_n be the n'th level of tree. We look at the following Markov-chains:

- $X_n = \sigma_{L_n} = (\sigma_v)_{v \in L_n}$, and,
- $X_n = c_{L_n}(i) = \#\{v \in L_n : \sigma_v = i\}$.



The basic question: Does the tail of the chain gives information on X_0?

Reconstruction Solvability

We say that the reconstruction problem is solvable if one of the following equivalent conditions hold:

- $\exists \pi$ s.t. ($\forall$ non-degenerate π) :
  $\lim_{n\to\infty} I(X_0, X_n) > 0$.

- $\exists i, j$ s.t. $\lim_{n\to\infty} \|\mathbf{P}_n^i - \mathbf{P}_n^j\| > 0$, where $\mathbf{P}_n^l$ denotes the conditional distribution on X_n given that $X_0 = l$.

- If π is uniform on X_0, then, $\liminf_{n\to\infty} \Delta_n > 1/m$, where Δ_n is the probability of correct reconstruction of X_0 given X_n.

- $\{X_n\}$ has a non-trivial tail.

Example: The 1-ary tree

Here it is impossible to reconstruct as:

$$\max_{i,j} \|\mathbf{P}_n^i - \mathbf{P}_n^j\| \sim C\lambda_2^n(M)$$

where $\lambda_2(M)$ is the second eigen-value value of M.

The Ising Model on the tree

Theorem 1: If M is the matrix of the binary-symmetric channel:

$$\begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix}$$

then reconstruction problem is solvable for σ_{L_n} (c_{L_n}) **iff**

$$d\lambda_2^2(M) = d(1-2\epsilon)^2 > 1.$$

- Kamae (1977): Lower bound for regular trees.
- Bleher, Ruiz and Zagrebnov (1995): Upper bound for regular trees.
- Evans, Kenyon, Peres and Schulman (1995,1998) : general trees.
- Ioffe (1996): An alternative proof.
- Kenyon, M, and Peres (2000): An alternative proof.

$d\lambda^2 > 1 \Longrightarrow$ reconstruction

- Take the variable

$$S_n = \#_n(+) - \#_n(-).$$

- We have

$$\mathbf{E}_+[S_n] = \sum_{v \in L_n} \mathbf{E}_+[\sigma_v] = d^n \lambda^n.$$

and $\mathbf{E}_-[S_n] = -d^n \lambda^n$.

- If $d^n \lambda^n > \sqrt{d^n}$ reconstruction is expected.

- We see the bound $d\lambda^2 > 1$.

Proof of no-reconstruction when $d\lambda^2 < 1$ [Kenyon-M-Peres 2000]

The proof uses a general principle relating spectral gap of Glauber dynamics for particle systems to decay of mutual information.

- **General principle.**
  $\lambda = \Omega(1) \implies I(X_0, X_n) = e^{-\Omega(n)}$.
  (Bounded spectral gap $\implies$ exponential decay of mutual information between local functions and boundary conditions.)

- By using "path coupling", "block dynamics" and strange metric on the state space we prove that when $d\lambda^2 < 1$, Glauber dynamics has bounded spectral gap for the Ising model on the tree.

The sequence c_{L_n}

Theorem 2[M-Peres 2000]: The reconstruction problem is solvable for c_{L_n} if $d\lambda_2^2(M) > 1$, and is unsolvable if $d\lambda_2^2(M) < 1$.

Remarks:

- When $d\lambda_2^2(M) > 1$, let v_2 is the left eigen-vector of M which corresponds to λ_2. Reconstruct using the statistics

$$< v_2, c_{L_n} > .$$

- Kesten and Stigum (1966): if $d\lambda_2^2(M) < 1$ then c_{L_n} satisfies the same CLT for all root values.

- Use the last level in order to get LCLT.

Reconstruction algorithms

A **reconstruction algorithm** for T is a function

$$A : \{1, \ldots, m\}^{|\partial T|} \to \{1, \ldots, m\}.$$

the probability that A reconstructs the root is:

$$\mathbf{P}[A(\sigma_{\partial T}) = \sigma_r].$$

Some reconstruction algorithms

- **The trivial algorithm:**
  - $A \equiv j$, where j is independent from the data.
  - $\mathbf{P}[A(\sigma_{\partial T}) = \sigma_r] = 1/m$.

- **Maximum likelihood:**
  - Compute: $p_i(\sigma) = \mathbf{P}[\sigma_r = i | \sigma_{\partial T} = \sigma]$,
  - Let $A(\sigma) = j$, where $p_j(\sigma) = \max p_i(\sigma)$.
  - This scheme achieves the optimal success probability $\mathbf{P}[A(\sigma_{\partial T}) = \sigma_r]$ among all A.

More reconstruction algorithms

- **Linear estimators:**

  – Fix $v = (v_1, \ldots, v_m) \in \mathbf{R}^m$.

  – Reconstruct according to the value of

  $$< c_{L_n}, v > .$$

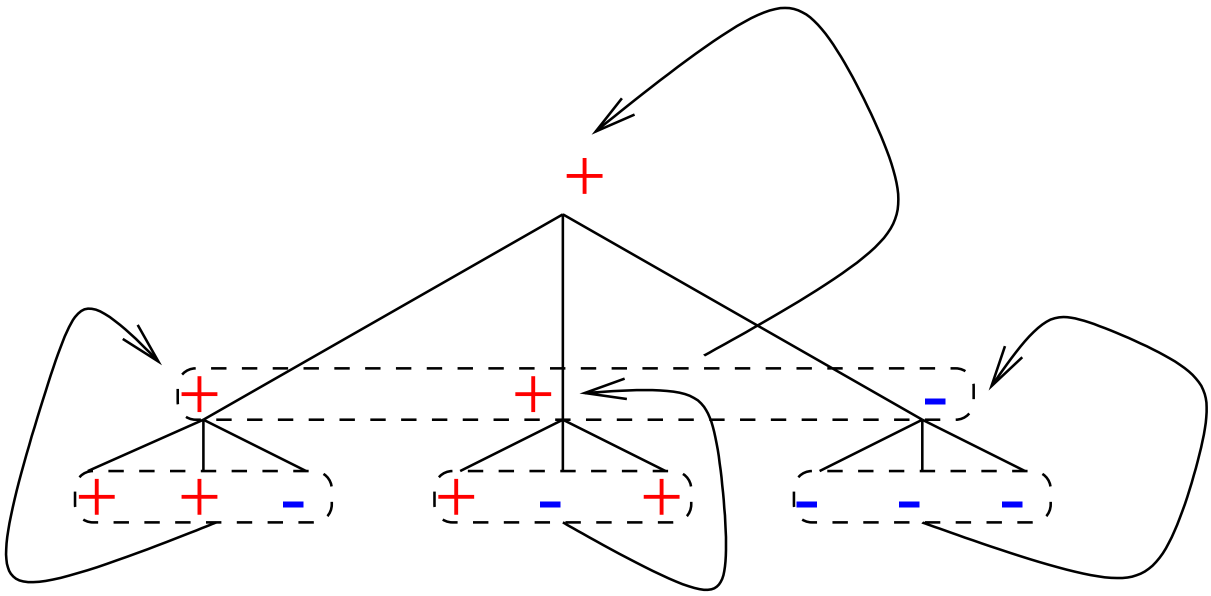
  – The location of the symbols on the boundary plays no role.

  – We know that when $d\lambda_2^2(M) > 1$ these estimators "work" for reconstructing σ_{L_n} as well as for c_{L_n}.

  – For the Ising model, whenever the reconstruction problem is solvable it is solvable using a Linear estimator.

More reconstruction algorithms

- **Recursive methods:**



For $A : \{1, \dots, m\}^d \to \{1, \dots, m\}$ the iterates of A are defined:

- $A^1 = A$,
- $A^n : \{1, \dots, m\}^{d^n} \to \{1, \dots, m\}$ by:

$$A^k(\sigma_{1,\dots,d^k}) =$$

$$A\Big(A^{k-1}(\sigma_{1,\dots,d^{k-1}}), \dots, A^{k-1}(\sigma_{(d-1)d^{k-1}+1,\dots,d^k})\Big)$$

Th [M1998]: Let M be the binary symmetric channel, and suppose we are looking at recursive algorithms $A : \{0,1\}^{d^h} \to \{0,1\}$, where A do "h levels at a time". Recursive reconstruction is possible **iff** $\epsilon < \epsilon_r$ where

$$(\frac{1}{2})^{d-1} \lceil \frac{d}{2} \rceil \binom{d}{\lceil \frac{d}{2} \rceil} (1 - 2\epsilon_r)^h = 1.$$

The algorithms $\mathbf{M} : \{0,1\}^d \to \{0,1\}$ which recursively solve the reconstruction problem for all $\epsilon < \epsilon_r$ are majority algorithms.

Corollary: No such recursive algorithm achieve the optimal reconstruction parameter.

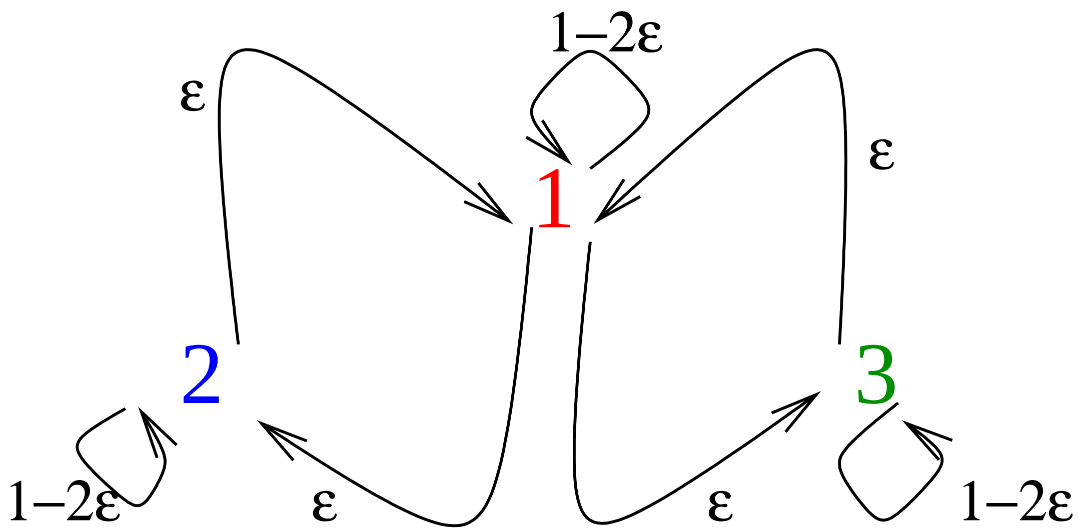
The Potts Model

This is a q-state model where

$$M = \begin{pmatrix} 1-(q-1)\epsilon & \epsilon & \dots & \epsilon \\ \epsilon & 1-(q-1)\epsilon & \epsilon & \dots \\ \vdots & \dots & \ddots & \vdots \\ \epsilon & \dots & \epsilon & 1-(q-1)\epsilon \end{pmatrix}$$

and the initial distribution is: $\pi = (1/q, \dots, 1/q)$.

With probability $\lambda = 1 - q\epsilon$ a symbol is kept; Otherwise we choose uniformly one of the q symbols.



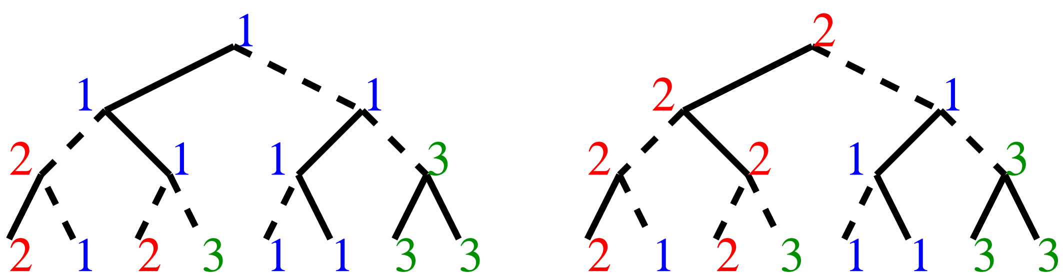
- $\lambda_2 = \dots = \lambda_q = \lambda = 1 - q\epsilon$.
- Reconstruction using a linear estimator is possible if and only if $d\lambda^2 = d(1-q\epsilon)^2 > 1$.

- **Reconstruction is impossible if $d\lambda \leq 1$.**

  Proof sketch (Random cluster representation):

  – For each symbol, we keep it with probability $\lambda = 1 - q\epsilon$; Otherwise, we choose uniformly one of the q symbols.

  – Thus if the root cluster does not intersect level n then, the configuration at level n is independent of the root.



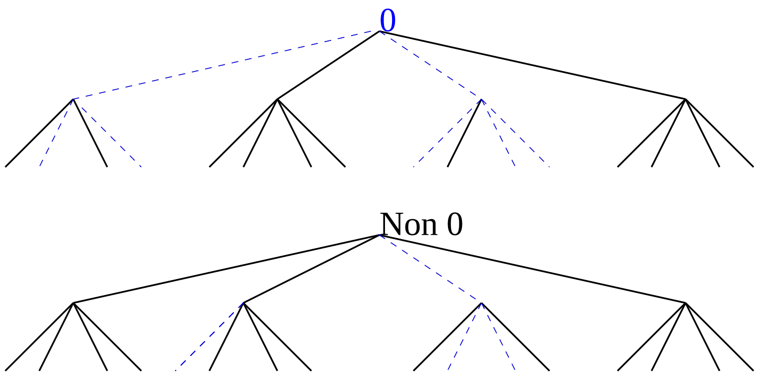
  – When λ-percolation dies we obtain $\|\mathbf{P}_n^i - \mathbf{P}_n^j\| \to 0$.

- **Theorem [M2000]:** If $d\lambda > 1$ and q is sufficiently large, then the reconstruction problem is solvable.

- **Open problem:** For $q = 3$ and the binary tree what is the critical value for reconstruction?

Proof Sketch.

- – When $d\lambda > 1$, percolation is super-critical.
  – We can show that with probability $\geq \delta > 0$ the root λ-cluster contains an open l-diluted binary tree rooted at ρ.
  – Assuming that the root value is 0 this is true also for the 0 cluster.



- – For ϵ small (i.e. q is large), we show that with probability $\geq 1-\delta/2$ the $(1-\epsilon)$-cluster contains (l^d-1)-ary l-diluted tree rooted at ρ.
  – If $\sigma_\rho \neq 0$, this is true for the nonzero cluster.
  – If we look at the 0's at level nl, then with probability at least $1-\delta/2$ they cannot contain the nl level of an l-diluted binary tree.

Improved bounds for non-reconstruction [M-Peres 2000]

We look at the $q = 3$ Potts model. We know:

$d(1-3\epsilon)^2 > 1 \Longrightarrow$ reconstruction,

$d(1-3\epsilon) \leq 1 \Longrightarrow$ no reconstruction.

Thm[M-Peres]: $\frac{d(1-3\epsilon)^2}{(1-\epsilon)} \leq 1 \implies$ no reconstruction.

Proof: The proof follows from the results for the Ising model on general trees [EKPS].

- We will show to it is asymptoticly impossible to distinct the root values 0 and 1.
- We do it by showing more: Suppose we are given all the node that are labeled by 2, then it is asymptoticly impossible to distinct the root values 0 and 1.

- We look at the $\{0,1\} \to \{0,1\}$ process.

  This process has

$$M = \frac{1}{1-\epsilon} \begin{pmatrix} 1-2\epsilon & \epsilon \\ \epsilon & 1-2\epsilon \end{pmatrix}$$

  Thus $\lambda = (1-3\epsilon)(1-\epsilon)^{-1}$.

- The tree has branching number (growth) $b = (1-\epsilon)$.

- It is thus impossible to reconstruct if

$$1 \geq b\lambda^2 = \frac{d(1-3\epsilon)^2}{(1-\epsilon)}.$$

The asymmetric binary channel [M2000]

Consider $\mathbf{P}[\sigma_r = 0] = \mathbf{P}[\sigma_r = 1] = 1/2$, and

$$M = \begin{pmatrix} 1-\delta & \delta \\ 1-\theta & \theta \end{pmatrix}$$

- The second eigenvalue is $\lambda_2(M) = \theta - \delta$. Therefore, in order to reconstruct using a generalized majority scheme we need $(\theta - \delta)^2 > 1/d$.

- Using the random-cluster representation we see again that if $d\lambda_2(M) \leq 1$ it is impossible to reconstruct.

- We show that if $d\lambda_2(M) > 1$ and $\delta > 0$ is small, it is possible to reconstruct using a recursive scheme.

- In [M-Peres 2000] improved bounds for non-reconstruction.

Question: Does the bound $d\lambda_2(M) \leq 1$ implies in general non-reconstruction?

Examples

- **General principle:** If the for all $i \neq j$, $M_{i,*} \neq M_{j,*}$ and d is large the reconstruction problem is solvable.

- Take M to be the shift on $\{0,1\}^2$,

$$\mathbf{P}[M(x,y) = (y,0)] = \mathbf{P}[M(x,y) = (y,1)] = 0.5.$$

- After 2 steps the chain mixes, so $\lambda_2(M) = 0$.

- Reconstruction is impossible for all d: After one step $(0,0) \sim (1,0)$ and $(0,1) \sim (1,1)$. Now we have the chain:

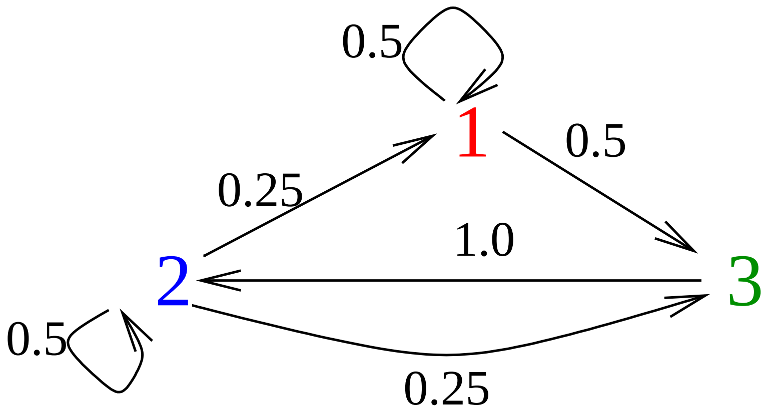
$$\mathbf{P}[0 \to 0] = \mathbf{P}[0 \to 1] = \mathbf{P}[1 \to 0] = \mathbf{P}[1 \to 1] = 0.5.$$

  So reconstruction is impossible.

- By lumping the states $(0,1)$ and $(1,0)$ we obtain again a chain with $\lambda_2(M) = 0$.

- Now

$$M = \begin{pmatrix} 0.5 & 0 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0 & 1 & 0 \end{pmatrix}$$



- Since the rows of the matrix are different, when d is large the reconstruction problem is solvable for the d-ary and this chain.

- This example works because M has the following Jordan form:

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

More strange Markov chains

- Above, if $\sigma = M(M(\tau))$ then $I(\sigma, \tau) = 0$.

- In [M-Peres2000] we construct Markov chains M such that if T is the 2 level d-ary tree, ρ is the root of T and $v_1, \ldots, v_{d^n}$ are the leaves of T, then

$$I\left(\sigma_\rho, (\sigma_{v_1}, \ldots, \sigma_{v_{d^n}})\right) = 0.$$

  On the other hand, the reconstruction problem for M and the $k >> d$ regular tree is solvable.

- – We let $F_p^d[t]$ be the set of all polynomials of degree $\leq d$ over the finite field F_p.
  – We let $\{x_1, \ldots, x_{d+1}\}$ be a set of non-zero elements.
  – $M : F_p^d[t] \to F_p^d[t]$ is the following Markov chain: $M(f) = g_{f(x_i)}$ where g_a is a random polynomial of degree d having a as a free coefficient and i is chosen uniformly in the range $1, \ldots, d+1$.