ϵM Reconstruction

Reading for this lecture:

Lecture Notes.

The Learning Channel ...

ϵ-Machine of a Process: Intrinsic representation!

Predictive (or causal) equivalence relation:

$$\overleftarrow{s}' \sim \overleftarrow{s}'' \iff \Pr(\overrightarrow{S} \mid \overleftarrow{S} = \overleftarrow{s}') = \Pr(\overrightarrow{S} \mid \overleftarrow{S} = \overleftarrow{s}'')$$

$$\overleftarrow{s}', \overleftarrow{s}'' \in \overleftarrow{\mathbf{S}}$$

Causal State:

Set of pasts with same morph $\Pr(\overrightarrow{S} \mid \overleftarrow{s})$: $\mathcal{S} = \{\overleftarrow{s}' : \overleftarrow{s}' \sim \overleftarrow{s}\}$

Set of causal states: $\boldsymbol{\mathcal{S}} = \overleftarrow{\mathbf{S}} / \sim \;\; = \{\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2, \ldots\}$

Causal state map: $\epsilon : \overleftarrow{\mathbf{S}} \to \boldsymbol{\mathcal{S}}$ $\quad \epsilon(\overleftarrow{s}) = \{\overleftarrow{s}' : \overleftarrow{s}' \sim \overleftarrow{s}\}$

Causal state morph: $\Pr\left(\overrightarrow{S}^L | \mathcal{S}\right)$

The Learning Channel ...

Causal State Dynamic:

State-to-State Transitions:

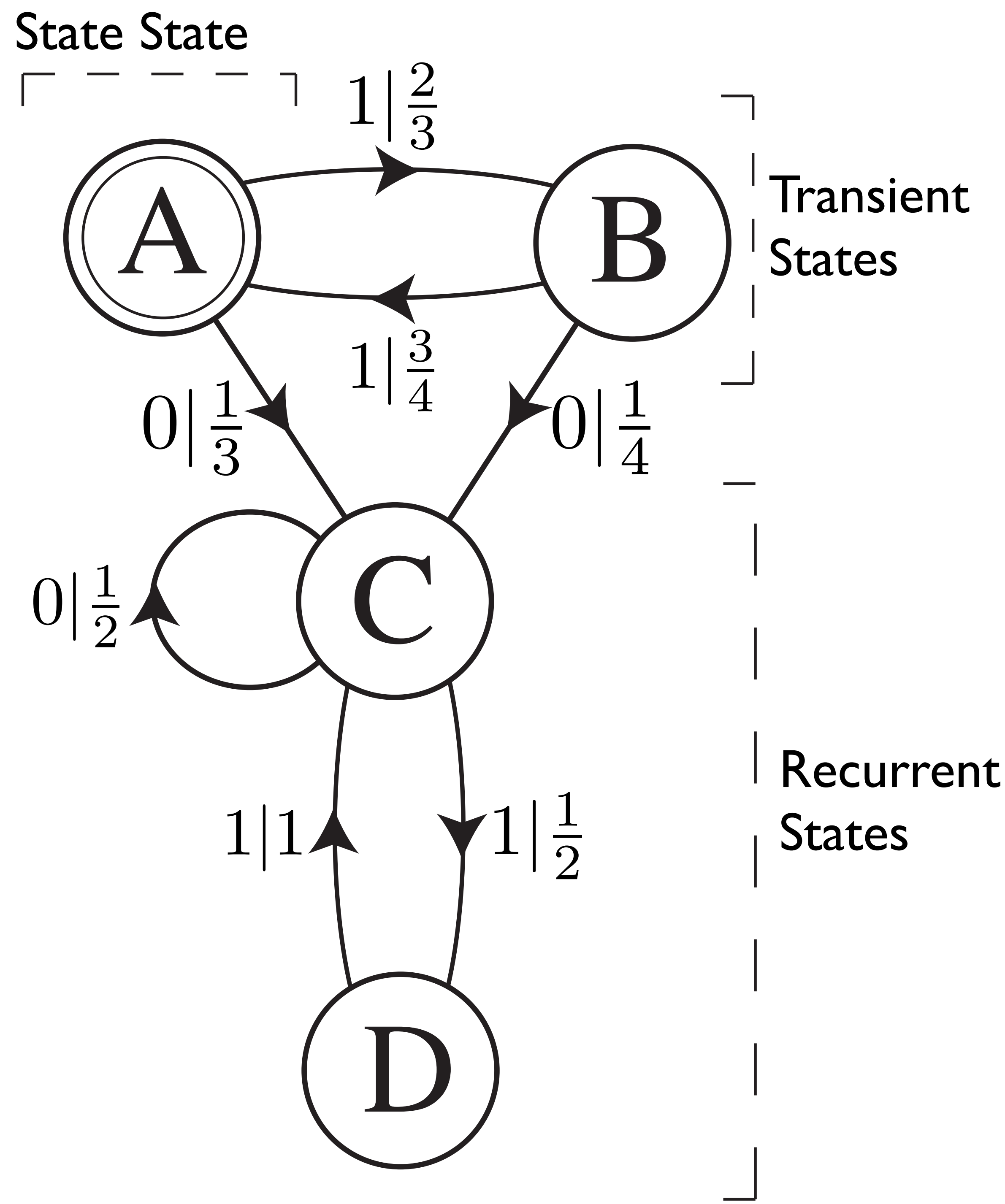
$$\{T_{ij}^{(s)} : s \in \mathcal{A},\ i, j = 0, 1, \ldots, |\boldsymbol{\mathcal{S}}|\}$$

$$\begin{aligned} T_{ij}^{(s)} &= \Pr(\mathcal{S}_j, s | \mathcal{S}_i) \\ &= \Pr\left(\mathcal{S}_j = \epsilon(\overleftarrow{s}\, s) | \mathcal{S}_i = \epsilon(\overleftarrow{s})\right) \end{aligned}$$

The Learning Channel ...

The ϵ-Machine of a Process ...

$$\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$



Unique Start State:

No measurements made:

$$\overleftarrow{s} = \lambda$$

Start state:

$$\mathcal{S}_0 = [\lambda]$$

Start state distribution:

$$\Pr(\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2, \ldots) = (1, 0, 0, \ldots)$$

Machine Reconstruction ...

ϵM Reconstruction:

Any method to go from process $\mathcal{P} \sim \Pr(\overleftrightarrow{S})$ to its ϵM

(1) Analytical: Given model, equations of motion, description, ...

(2) Statistical inference: Given samples of $\mathcal{P}$

(i) Subtree Reconstruction: Time or spacetime data to ϵM

(ii) State-splitting (CSSR): Time or spacetime data to ϵM

(iii) Spectral (eMSR): Power spectra to ϵM

(iv) Optimal Causal Inference: Time or spacetime data to ϵM

(v) Enumerative Bayesian Inference

Machine Reconstruction ...

How to reconstruct an $\epsilon\mathrm{M}$: Subtree algorithm

Given: Word distributions $\Pr(s^D), \; D = 1, 2, 3, \ldots$

Steps:

(1) Form depth-D parse tree.

(2) Calculate node-to-node transition probabilities.

(3) Causal states: Find morphs $\Pr(\overrightarrow{s}^L \mid \overleftarrow{s}^K)$ as subtrees.

(4) Label tree nodes with morph (causal state) names.

(5) Extract state-to-state transitions from parse tree.

(6) Assemble into $\epsilon\mathrm{M}$: $\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$.

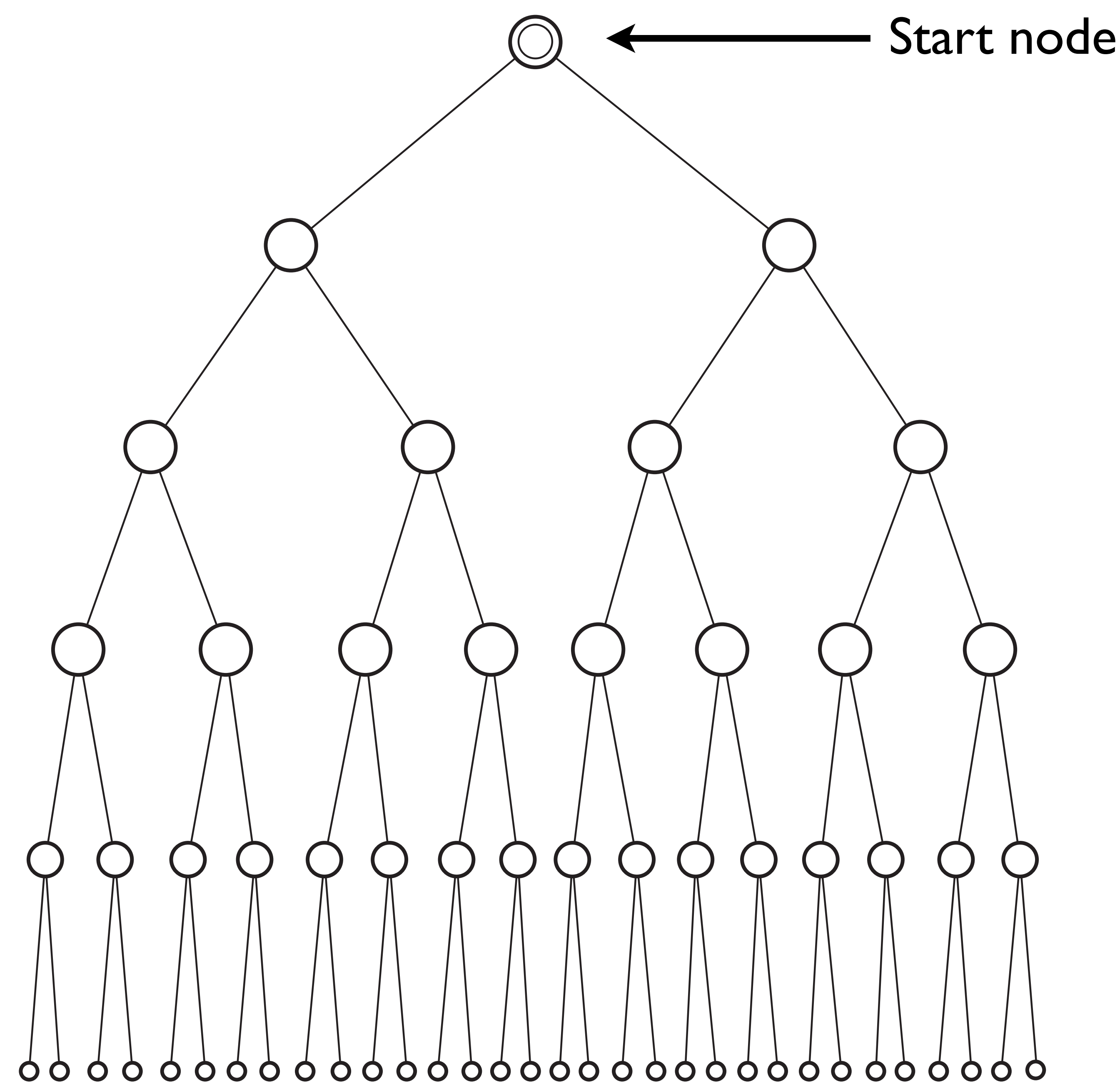
Algorithm parameters: D, L, K

Machine Reconstruction ...

How to reconstruct an $\epsilon\mathrm{M}$...

Form parse tree estimate of $\Pr(s^D)$.

Data stream: $s^M = \ldots 0101011110101001010101010101011$



Parse tree of depth $D = 5$

Number of samples: $M - D$

History length:

$K = 0, 1, 2, 3, \ldots$

Machine Reconstruction ...

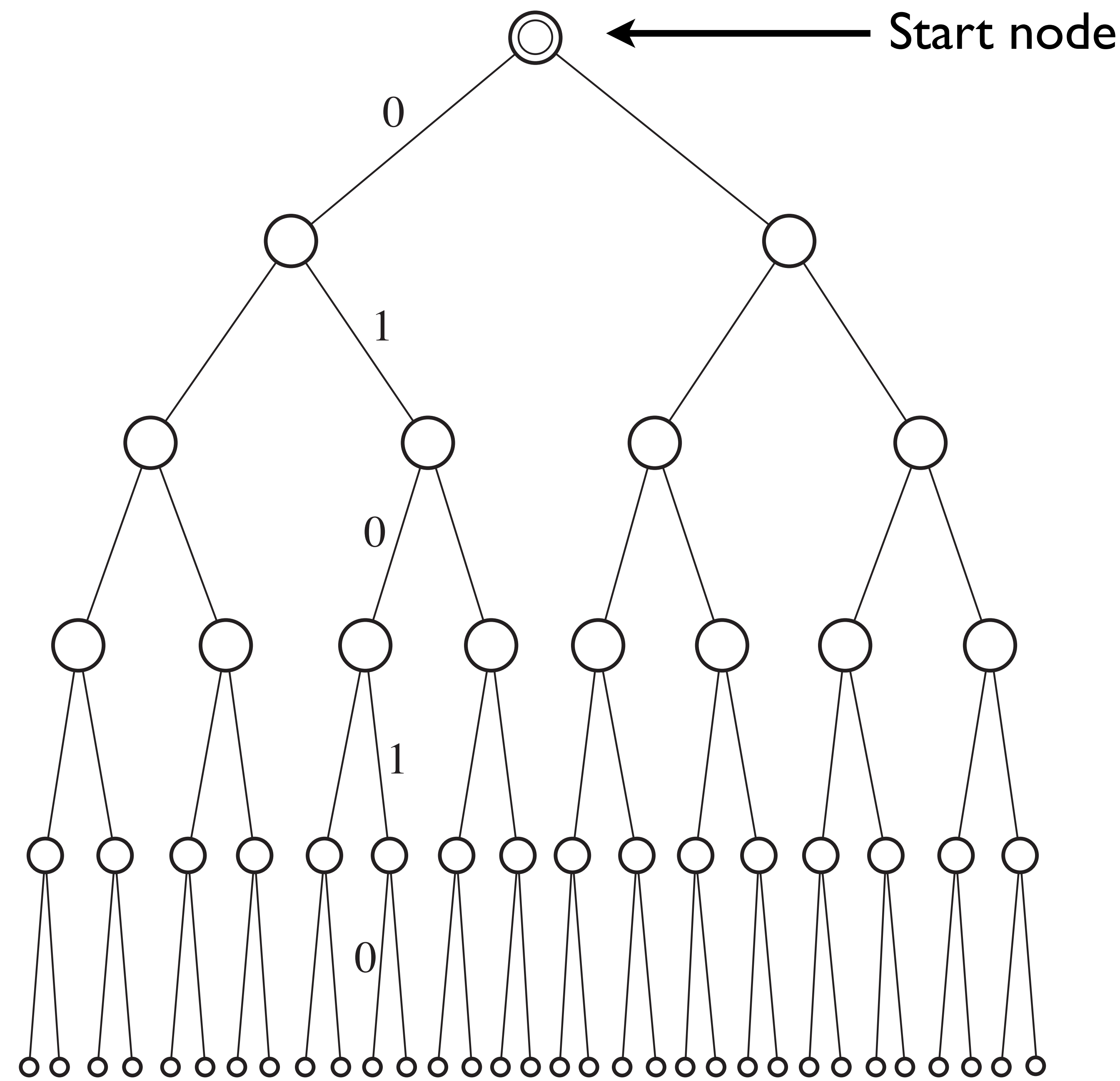
How to reconstruct an ϵM ...

Form parse tree estimate of $\Pr(s^D)$.

Data stream: $s^M = \ldots 010101111010100101010101011$

Parse tree of depth $D = 5$

$\ldots \underline{01010}1111010100101010101011$



Machine Reconstruction ...

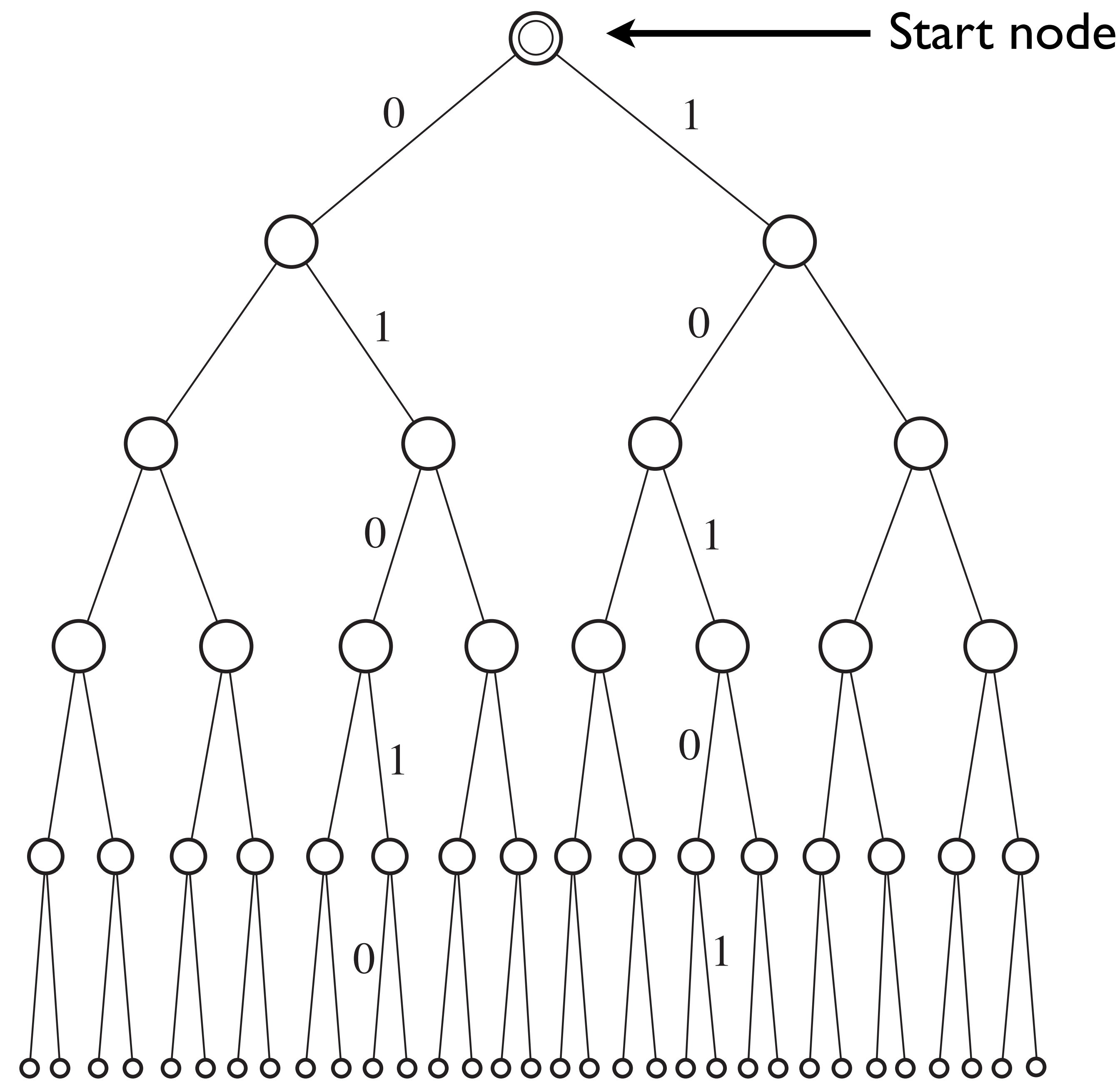
How to reconstruct an $\epsilon\mathrm{M}$...

Form parse tree estimate of $\Pr(s^D)$.

Data stream: $s^M = \ldots 01010111101010010101010101011$

Parse tree of depth $D = 5$

$\ldots 0\underline{10101}11101010010101010101011$

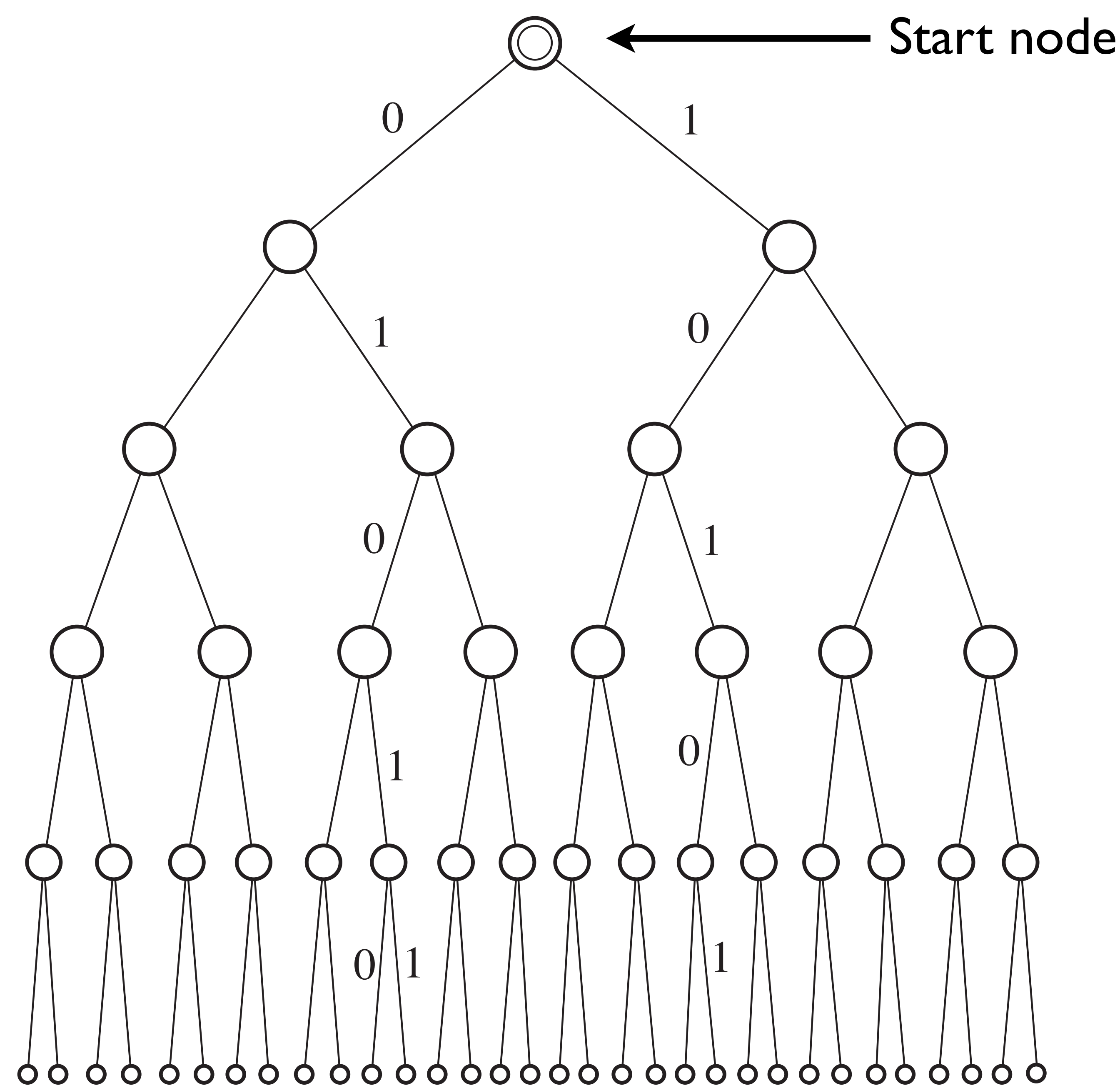


Machine Reconstruction ...

How to reconstruct an $\epsilon\mathrm{M}$...

Form parse tree estimate of $\Pr(s^D)$.

Data stream: $s^M = \ldots 01010111101010010101010101011$



Parse tree of depth $D = 5$

$\ldots 01\underline{01011}11101010010101010101011$

Machine Reconstruction ...

How to reconstruct an $\epsilon\mathrm{M}$...

Form parse tree estimate of $\Pr(s^D)$.

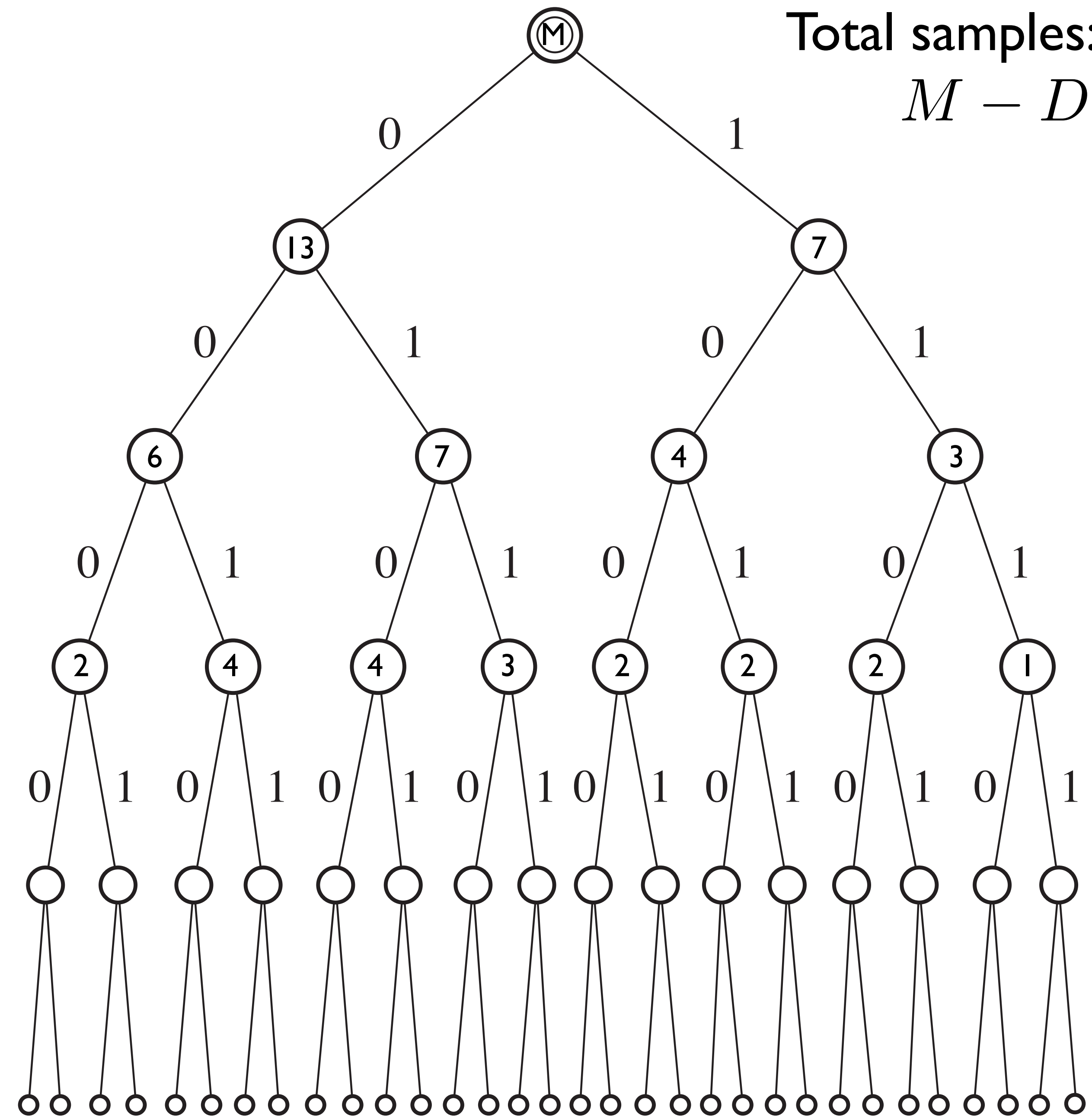
Data stream: $s^M = \ldots 01010111101010010101010101011$

Parse tree of depth $D = 5$

Store word counts at nodes.

Probability of node
= Probability of word
leading to node:

$$\Pr(w) = \frac{\text{node count}}{M}$$

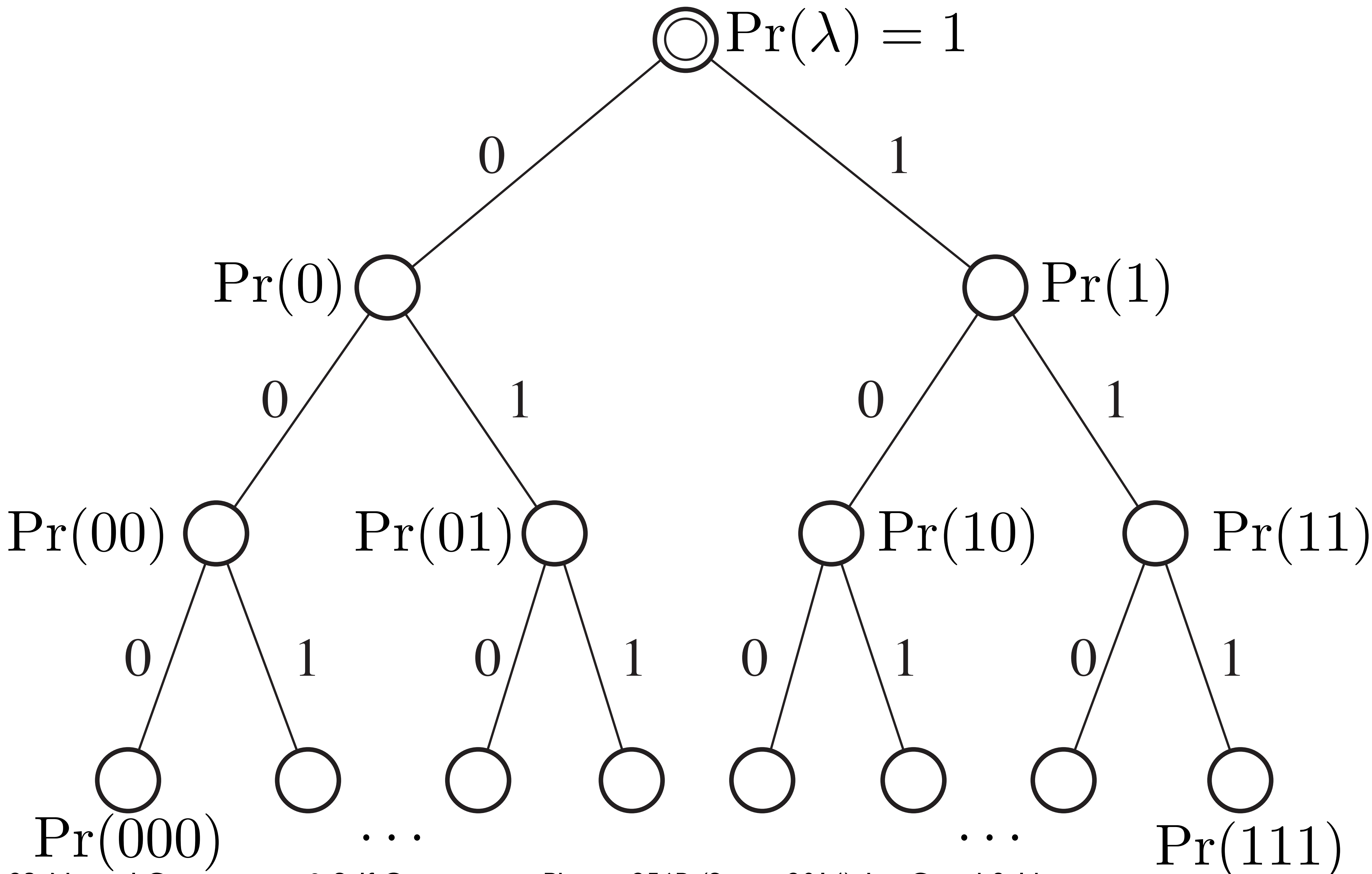


Total samples:
$M - D \approx M$

Machine Reconstruction ...

How to reconstruct an ϵM ...

Assume we have correct word distribution: $\Pr(s^D)$



Machine Reconstruction ...

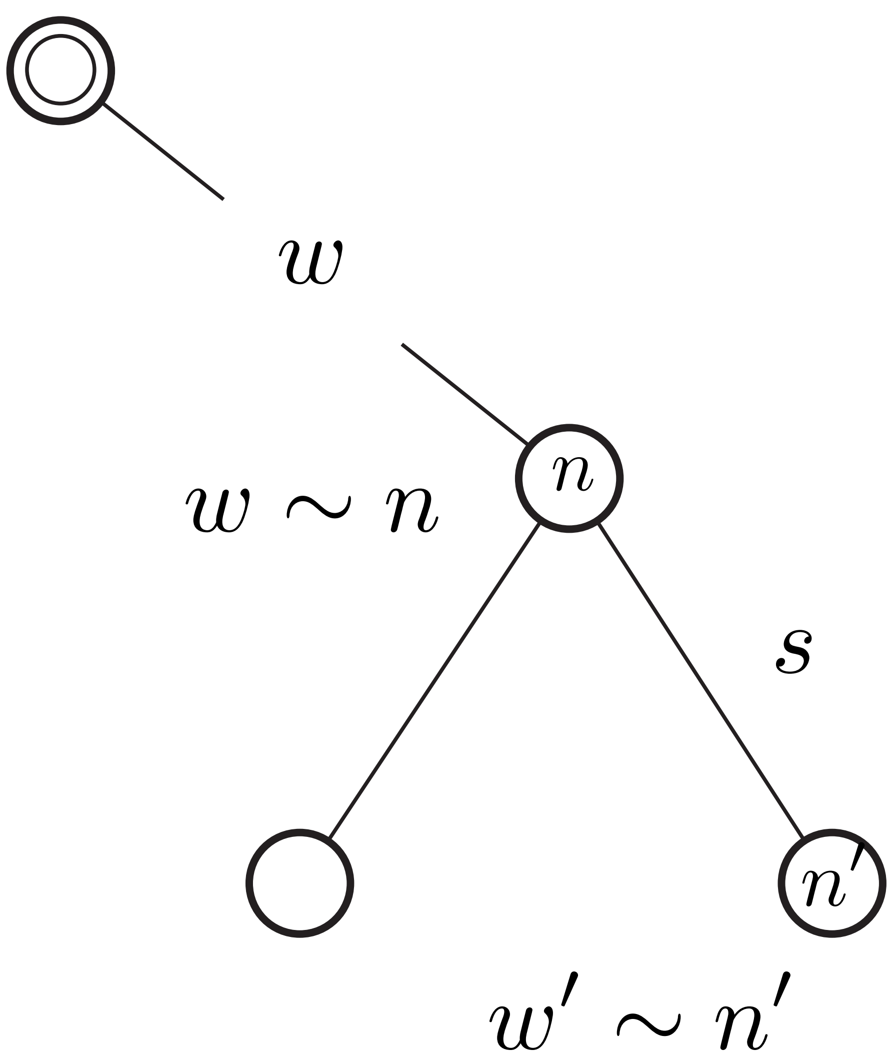
How to reconstruct an ϵM ...

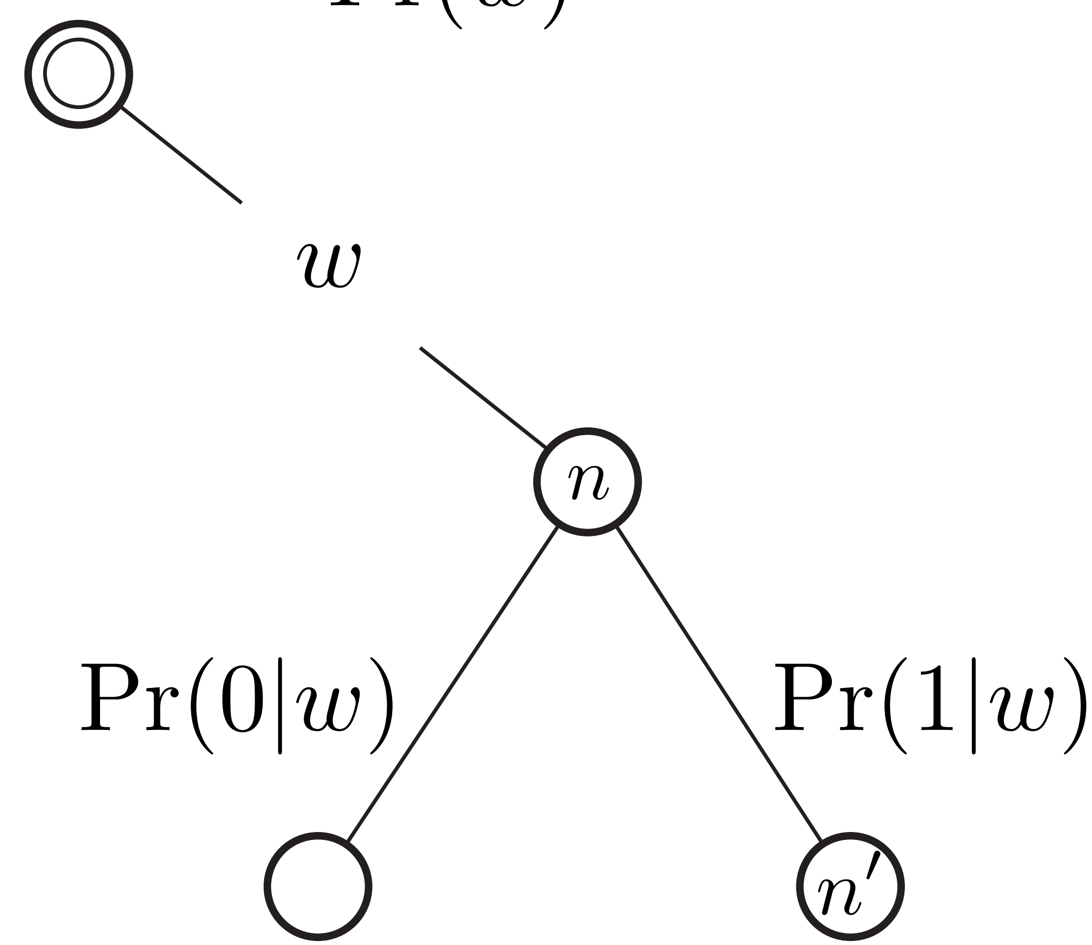
Node-to-node transition probability:

$$w' = ws$$

$$\Pr(n'|n,s) = \Pr(n \rightarrow n') = \frac{\Pr(n')}{\Pr(n)}$$

$$= \frac{\Pr(w')}{\Pr(w)} = \Pr(s|w)$$





Machine Reconstruction ...

How to reconstruct an ϵM ...

Find morphs $\Pr(\overrightarrow{s}^L \mid \overleftarrow{s}^K)$ as subtrees

Future: $L = 2$

Past: $K = 1$

Morph: $\Pr(\overrightarrow{S}^2 \mid \overleftarrow{S}^1 = 0)$

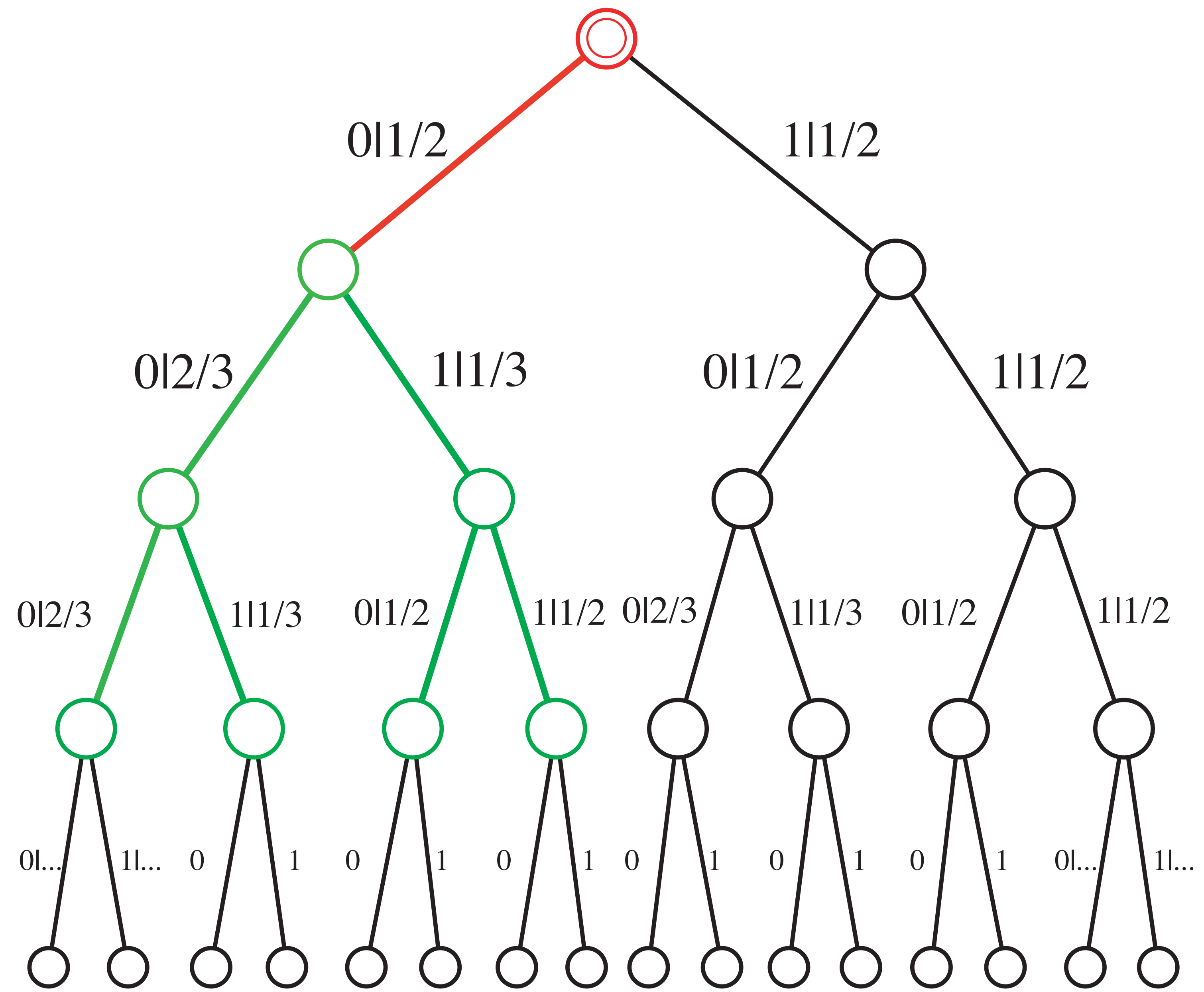
Given: $\overleftarrow{s}^K = 0$

$$\Pr(00) = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$$

$$\Pr(01) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$$

$$\Pr(10) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$\Pr(11) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$



Machine Reconstruction ...

How to reconstruct an ϵM ...

Find morphs $\Pr(\overrightarrow{s}^{L} | \overleftarrow{s}^{K})$ as subtrees

Future: $L = 2$

Past: $K = 1$

Morph: $\Pr(\overrightarrow{S}^{2} \mid \overleftarrow{S}^{1} = 1)$

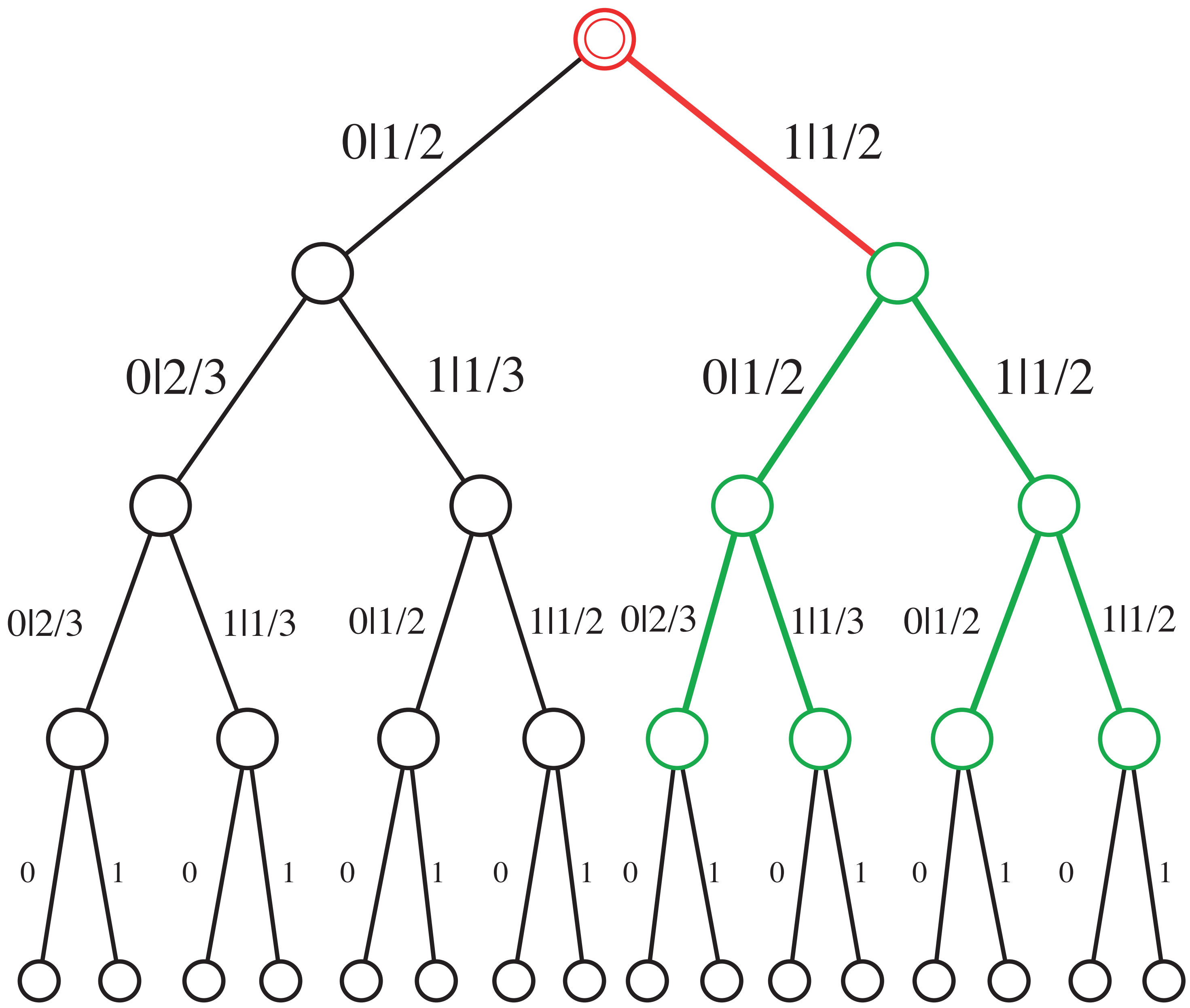
Given: $\overleftarrow{s}^{K} = 1$

$$\Pr(00) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$$

$$\Pr(01) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$

$$\Pr(10) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\Pr(11) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

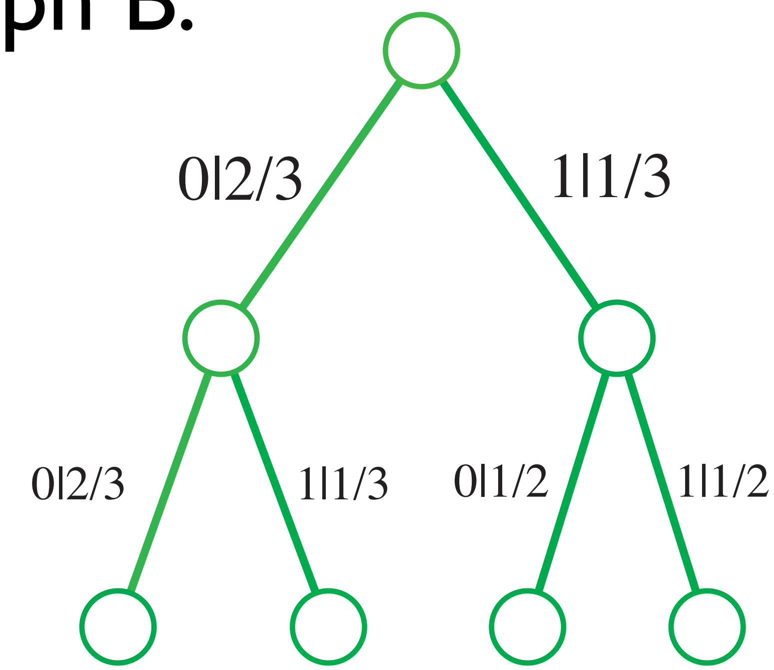


Machine Reconstruction ...

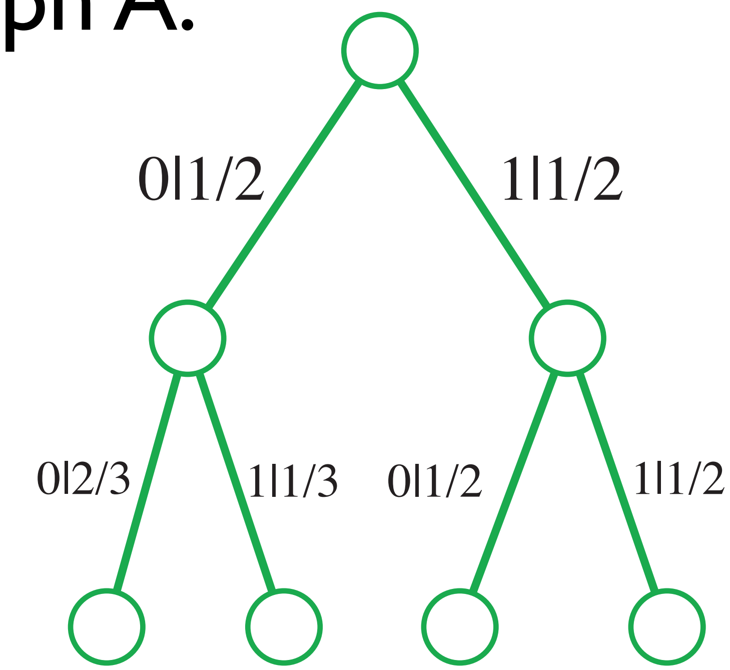
How to reconstruct an ϵM ...

Set of distinct morphs:

Morph B:



Morph A:



Set of causal states = Set of distinct morphs.

$$\mathcal{S} = \{A, B\}$$

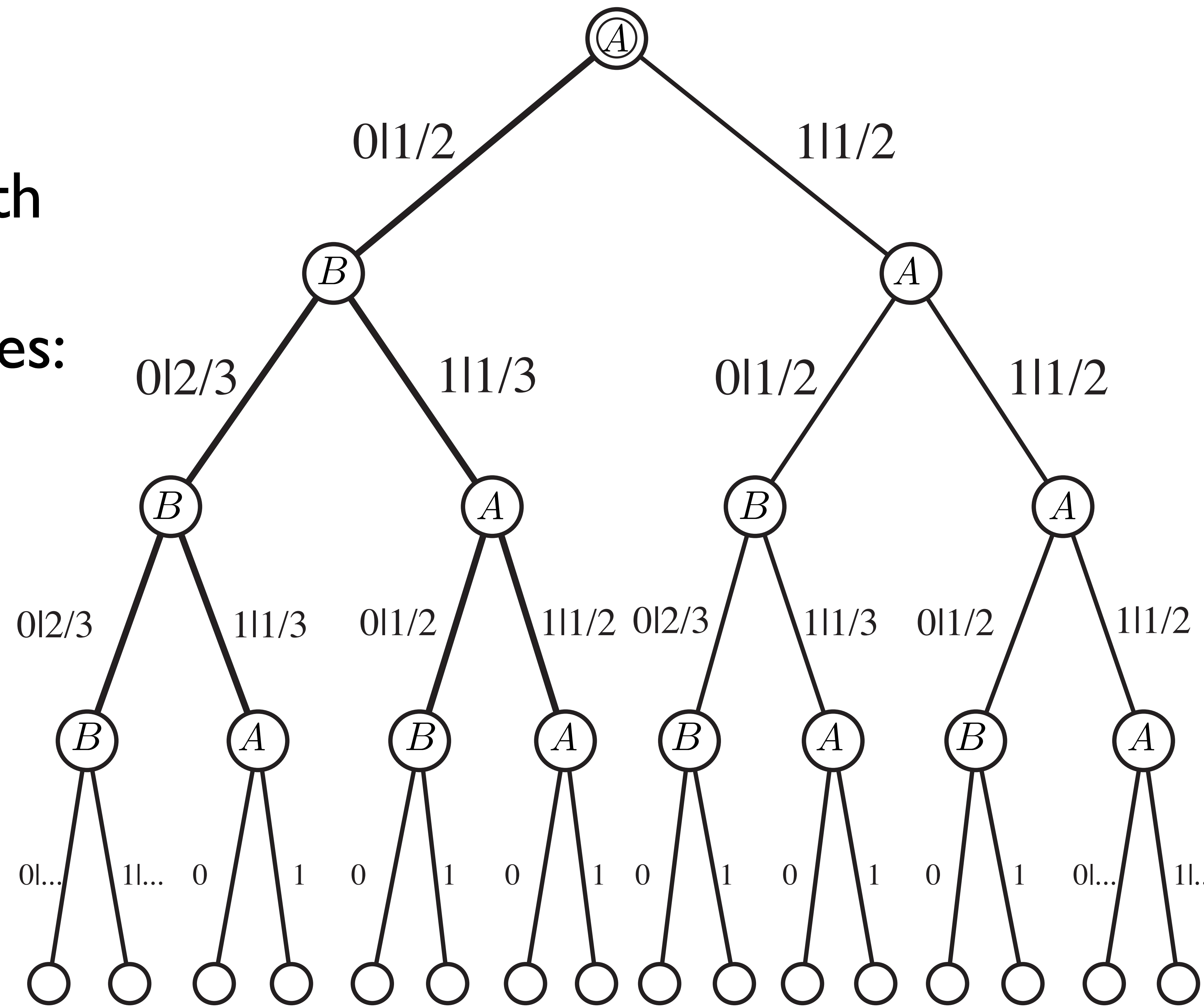
Machine Reconstruction ...

How to reconstruct an ϵM ...

Causal state transitions?

Label tree nodes with their morph (causal state) names:

$$\mathcal{S} = \{A, B\}$$



Machine Reconstruction ...

How to reconstruct an $\epsilon\mathrm{M}$ ...

Form $\epsilon\mathrm{M}$:

Causal states: $\mathcal{S} = \{A, B\}$

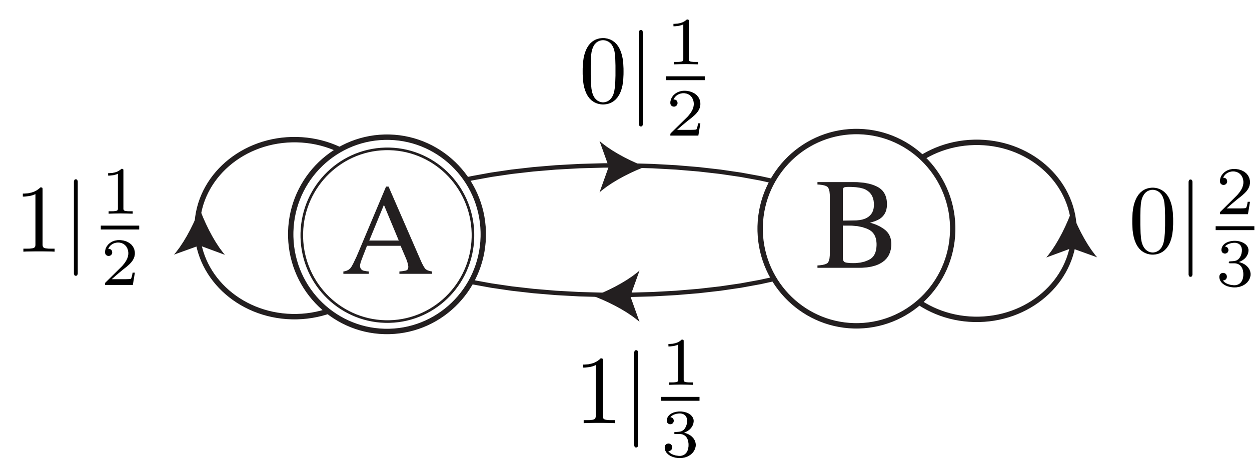
Start state ~ top tree node

Machine Reconstruction ...

How to reconstruct an ϵM ...

Form ϵM:

Causal-state transitions from node-to-node transitions:



Machine Reconstruction ...

How to reconstruct an $\epsilon\mathrm{M}$: Subtree algorithm

Given: Word distributions $\Pr(s^D),\ D = 1, 2, 3, \ldots$

Steps:

(1) Form depth-D parse tree.

(2) Calculate node-to-node transition probabilities.

(3) Causal states: Find morphs $\Pr(\overrightarrow{s}^L \mid \overleftarrow{s}^K)$ as subtrees.

(4) Label tree nodes with morph (causal state) names.

(5) Extract state-to-state transitions from parse tree.

(6) Assemble into $\epsilon\mathrm{M}$: $\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$.

Algorithm parameters: D, L, K

Machine Reconstruction ...

How to reconstruct an ϵM...

Example Processes:

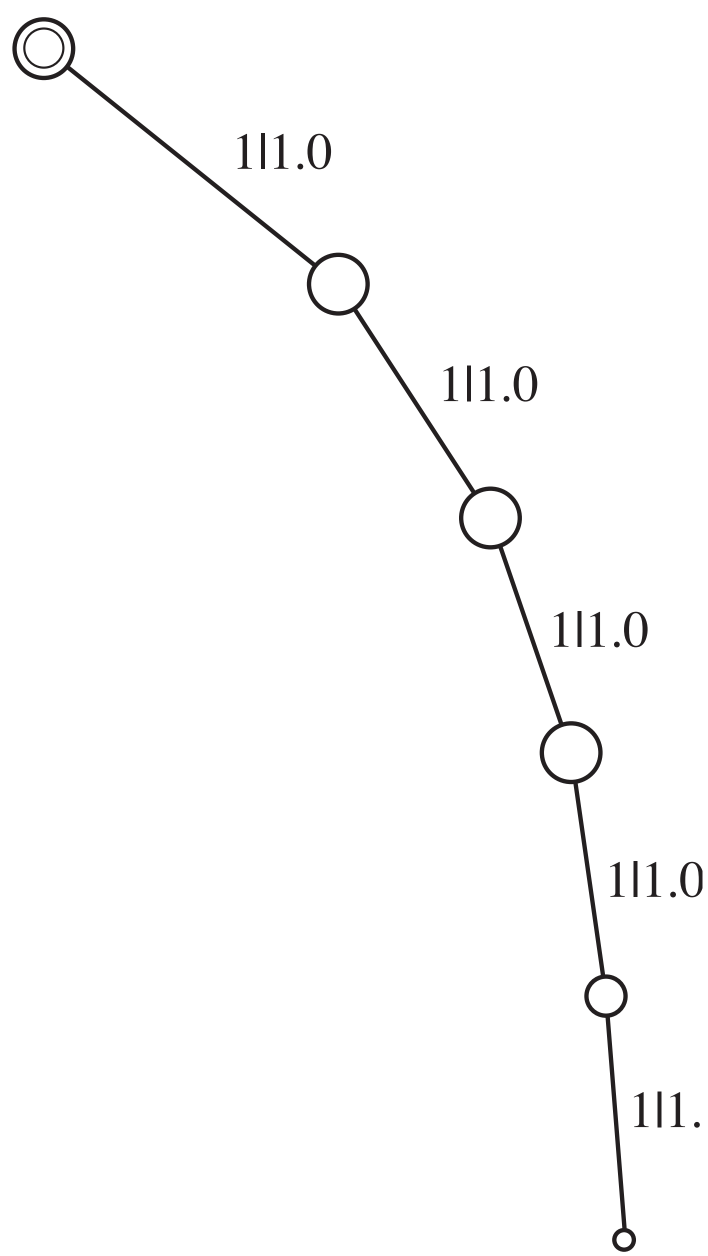
1. Period-1
2. Fair Coin
3. Biased Coin
4. Period-2
5. Golden Mean Process
6. Even Process

Machine Reconstruction ...

Examples (back to the Prediction Game):

Period-1: $\ldots 11111111111111$

Parse Tree D = 5

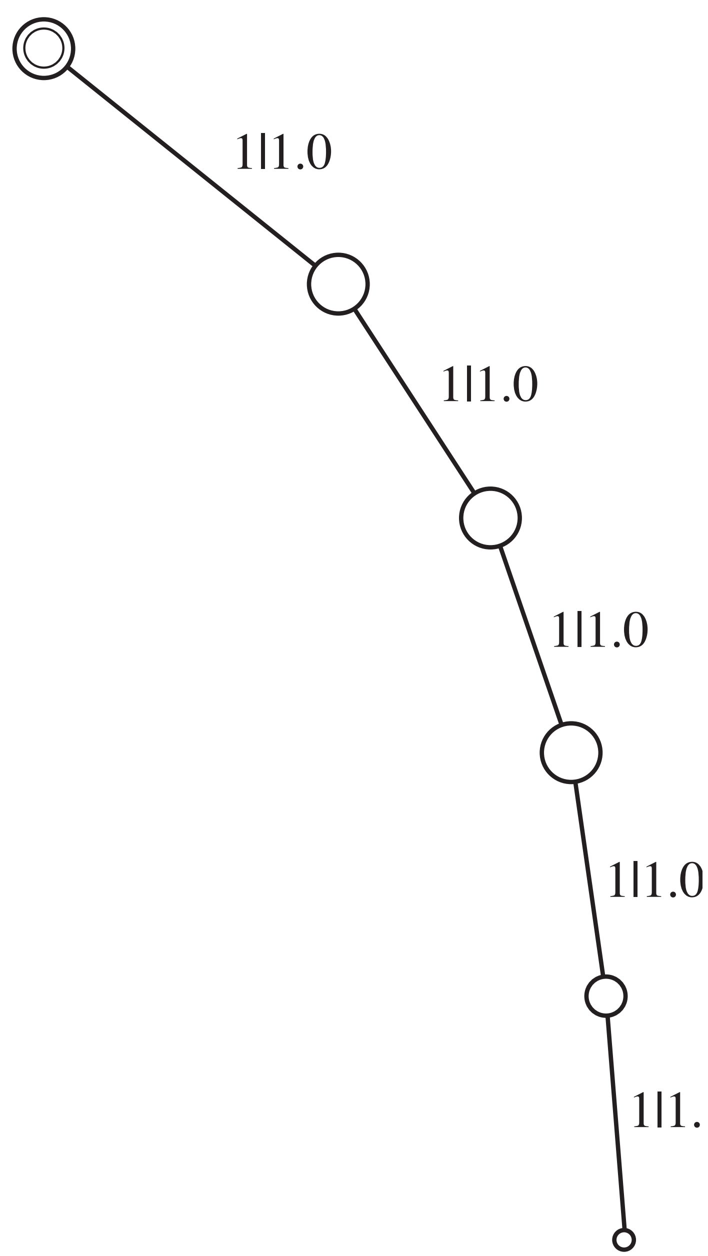


Machine Reconstruction ...

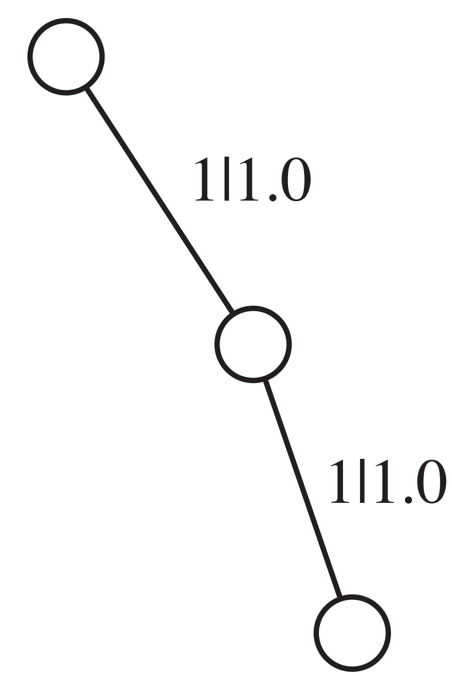
Examples (back to the Prediction Game):

Period-1: $\ldots 11111111111111$

Parse Tree D = 5



Morph L = 2



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-1: $\ldots 11111111111111$

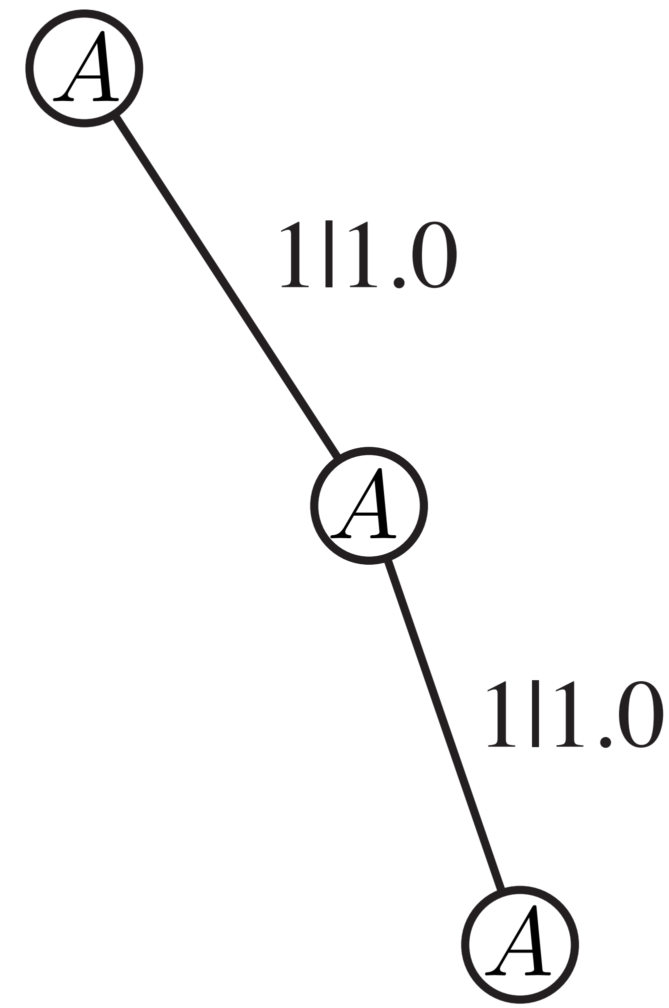
Space of histories: A single point.

One future morph:

Support: $\{1^+\}$

Distribution:

$$\Pr(\overrightarrow{S}^L = 1^L | \overleftarrow{s} = 1^K) = 1$$



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

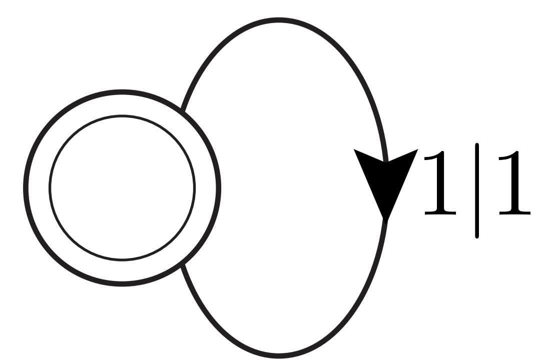
Period-1 ...

$$\epsilon\text{M:}\ \ \mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

$$\boldsymbol{\mathcal{S}} = \{\mathcal{S}_0 = \{\ldots 111111\}\}$$

$$T^{(0)} = (0)$$

$$T^{(1)} = (1)$$

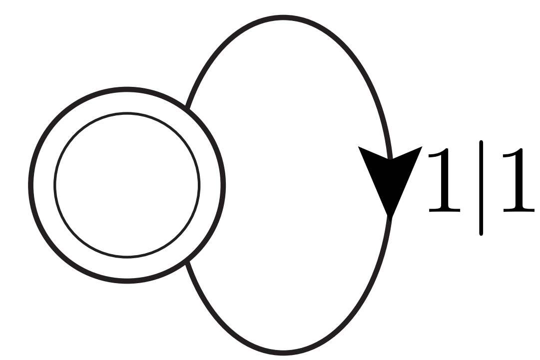


Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-1 ...

Causal state distribution: $p_{\mathcal{S}} = (1)$



Entropy Rate: $h_\mu = 0$ bits per symbol

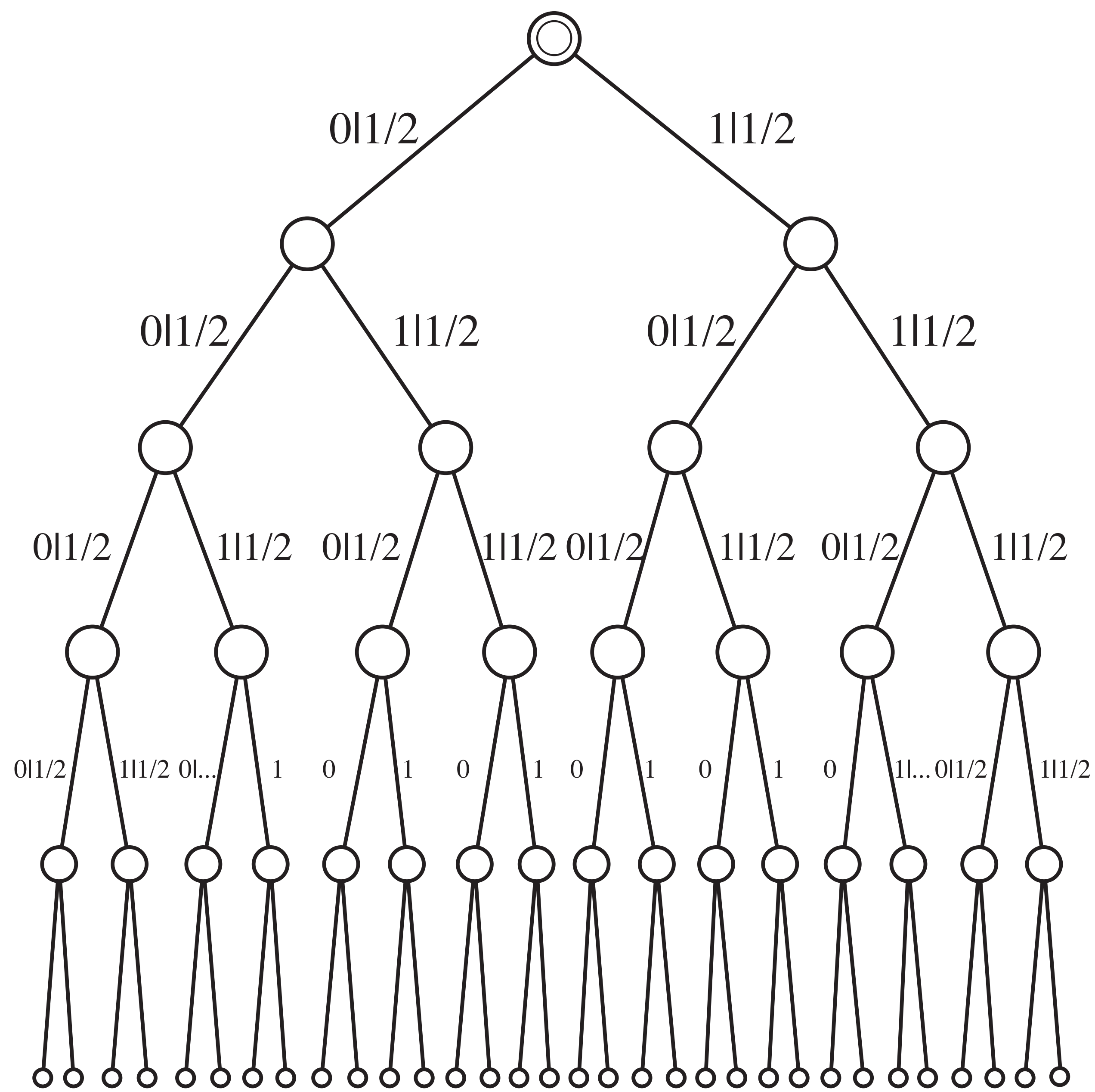
Statistical Complexity: $C_\mu = 0$ bits

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Fair Coin: $\ldots 0101001110001101$

Parse Tree D = 5

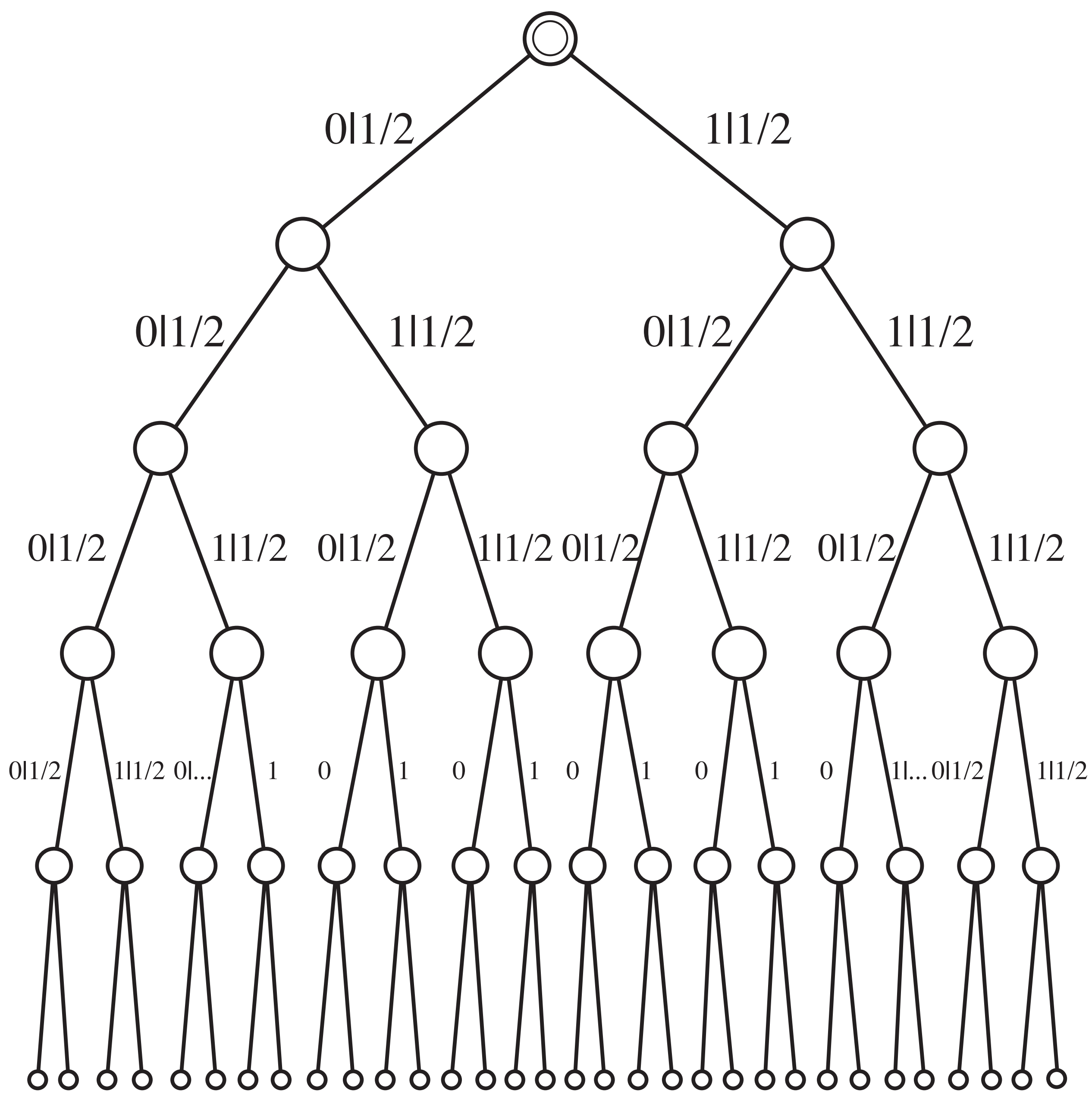


Machine Reconstruction ...

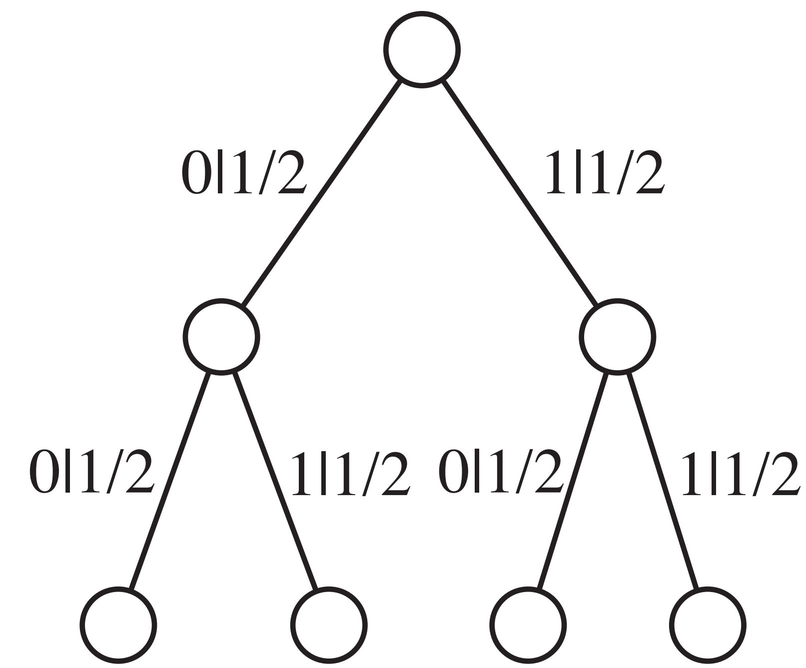
Examples (back to the Prediction Game) ...

Fair Coin: $\ldots 0101001110001101$

Parse Tree D = 5



Future Morph L = 2



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Fair Coin ...

Space of histories: $\overleftarrow{\mathbf{S}}^{K} = \mathcal{A}^{K}$

One future morph:

Support: $\mathcal{A}^{L}$

Distribution: $\Pr(\overrightarrow{S}^{L} | \overleftarrow{s}) = 2^{-L}$

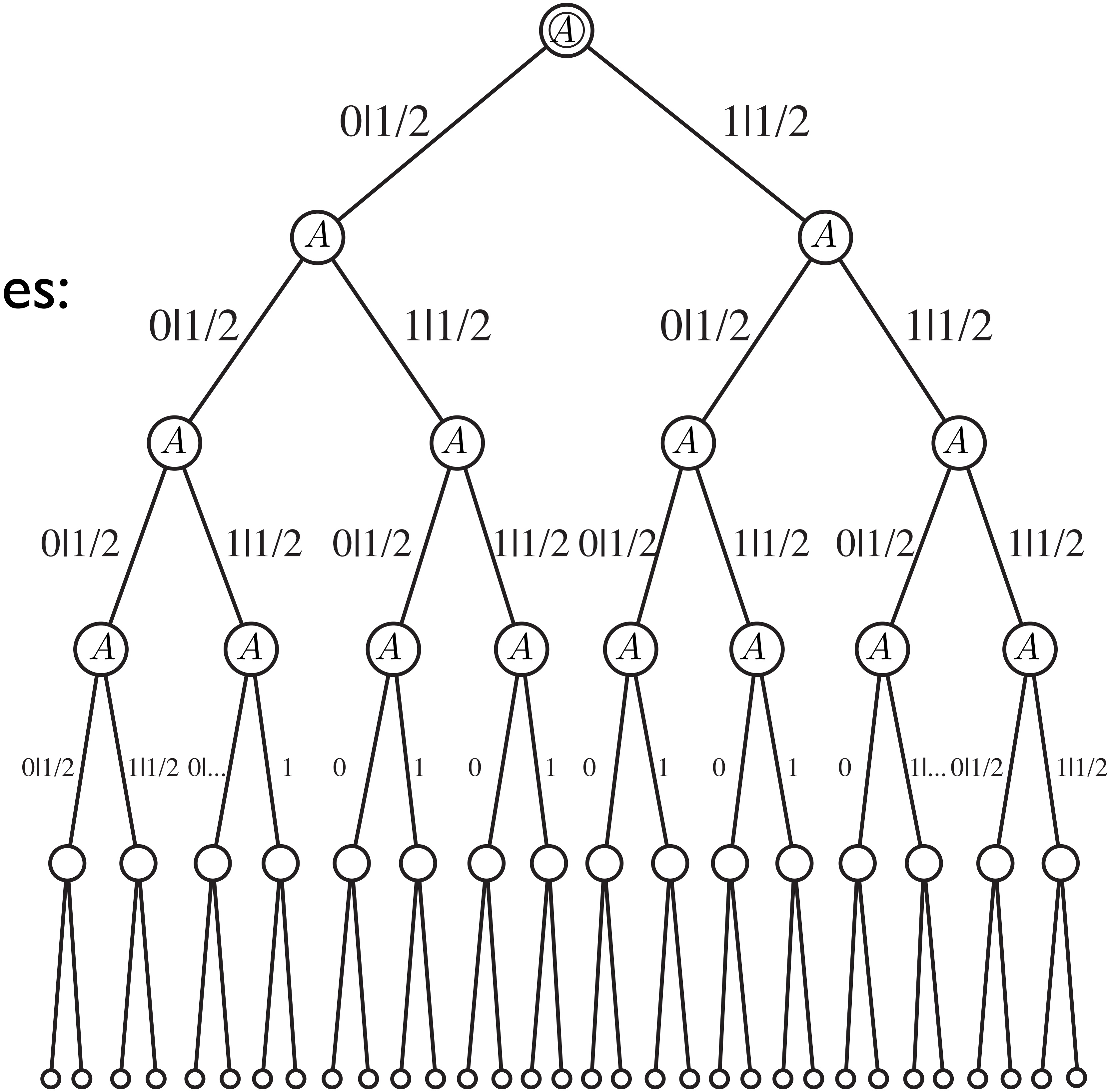
Call it state "A".

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Fair Coin ...

Label tree nodes
with state names:



Machine Reconstruction ...

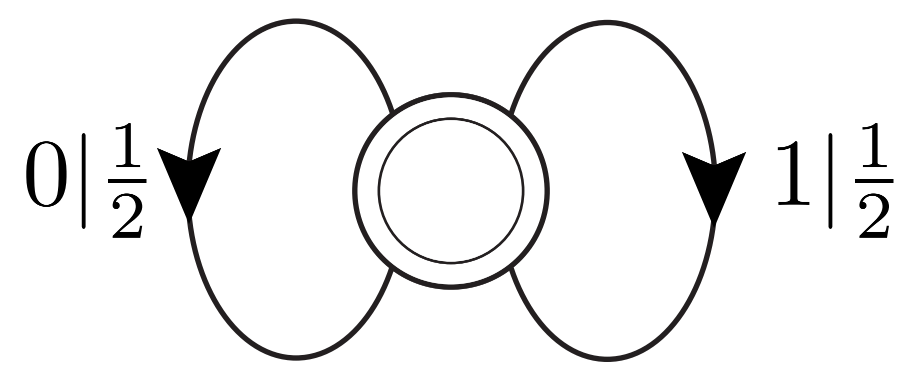
Examples (back to the Prediction Game) ...

Fair Coin ...

$$\epsilon\mathrm{M}\colon \mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

States:

$$\boldsymbol{\mathcal{S}} = \{\mathcal{S}_0 = \mathcal{A}^L\}$$



Transitions:

$$T^{(0)} = \left(\tfrac{1}{2}\right)$$

$$T^{(1)} = \left(\tfrac{1}{2}\right)$$

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Fair Coin ...

Causal state distribution: $p_{\mathcal{S}} = (1)$

Entropy Rate: $h_\mu = 1 \text{ bit per symbol}$

Statistical Complexity: $C_\mu = 0 \text{ bits}$

Machine Reconstruction ...

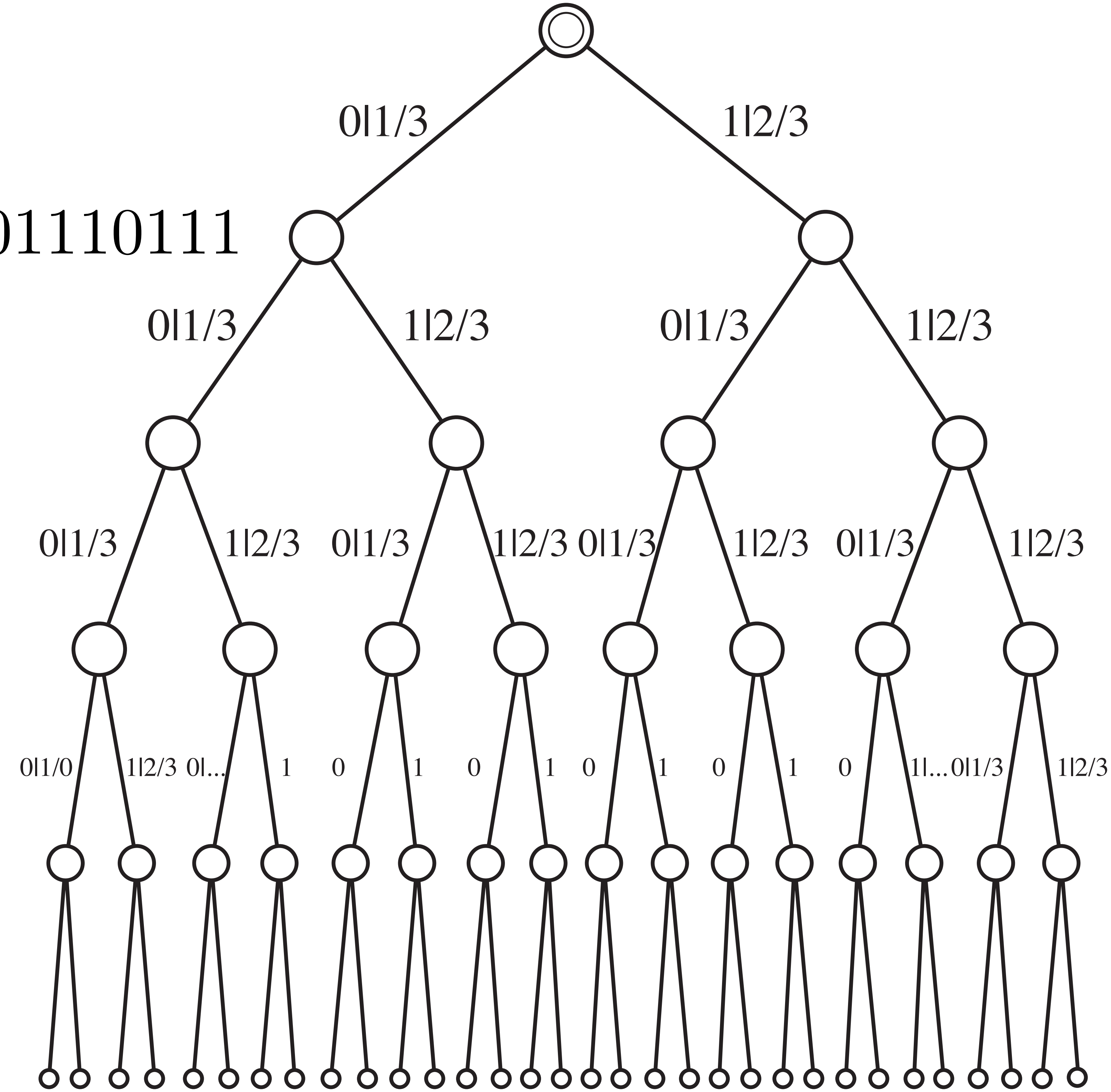
Examples ...

Biased Coin:

$$p \equiv \Pr(1) = \frac{2}{3}$$

$$\ldots 1011011101100111011101111$$

Parse Tree D = 5

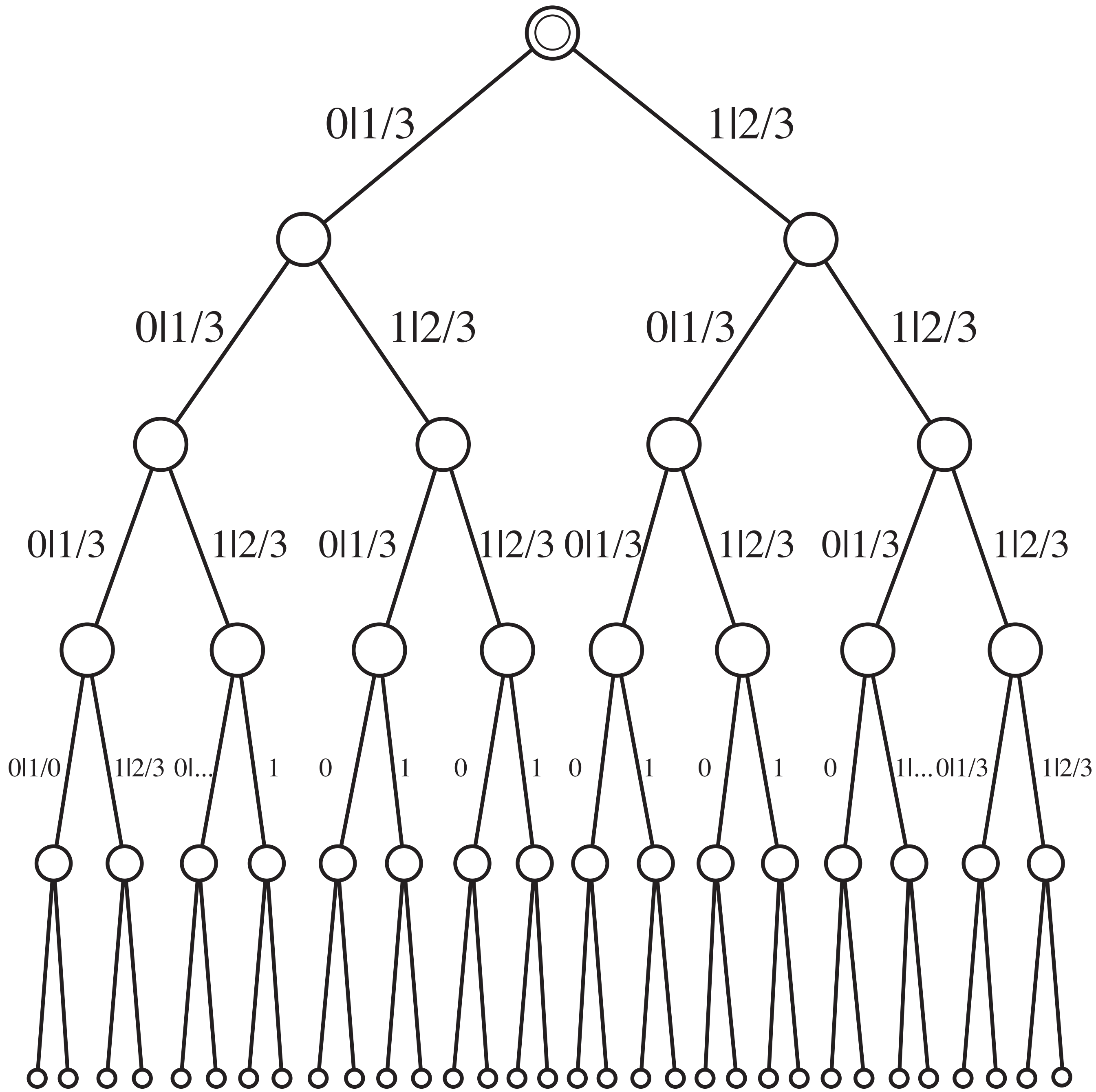


Machine Reconstruction ...

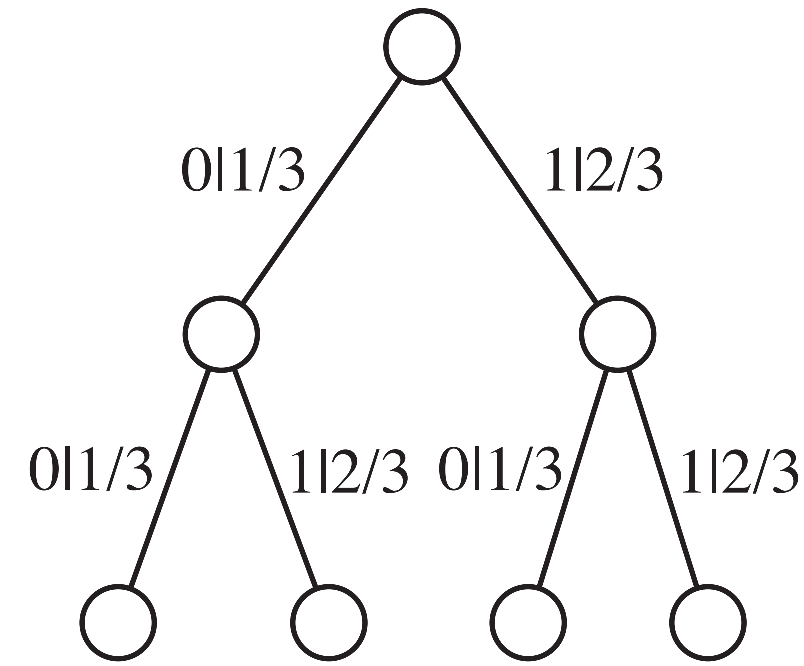
Examples (back to the Prediction Game) ...

Biased Coin:

Parse Tree D = 5



Future Morph L = 2



Examples (back to the Prediction Game) ...

Biased Coin ...

Space of histories: $\overleftarrow{\mathbf{S}}^{K} = \mathcal{A}^{K}$

Single future morph:

Support: $\mathcal{A}^{L}$

Distribution: $\Pr(\overrightarrow{S}^{L} \mid \overleftarrow{s}) = p^{n}(1-p)^{L-n}$ $(n \equiv \#1 \in \overrightarrow{s}^{L})$

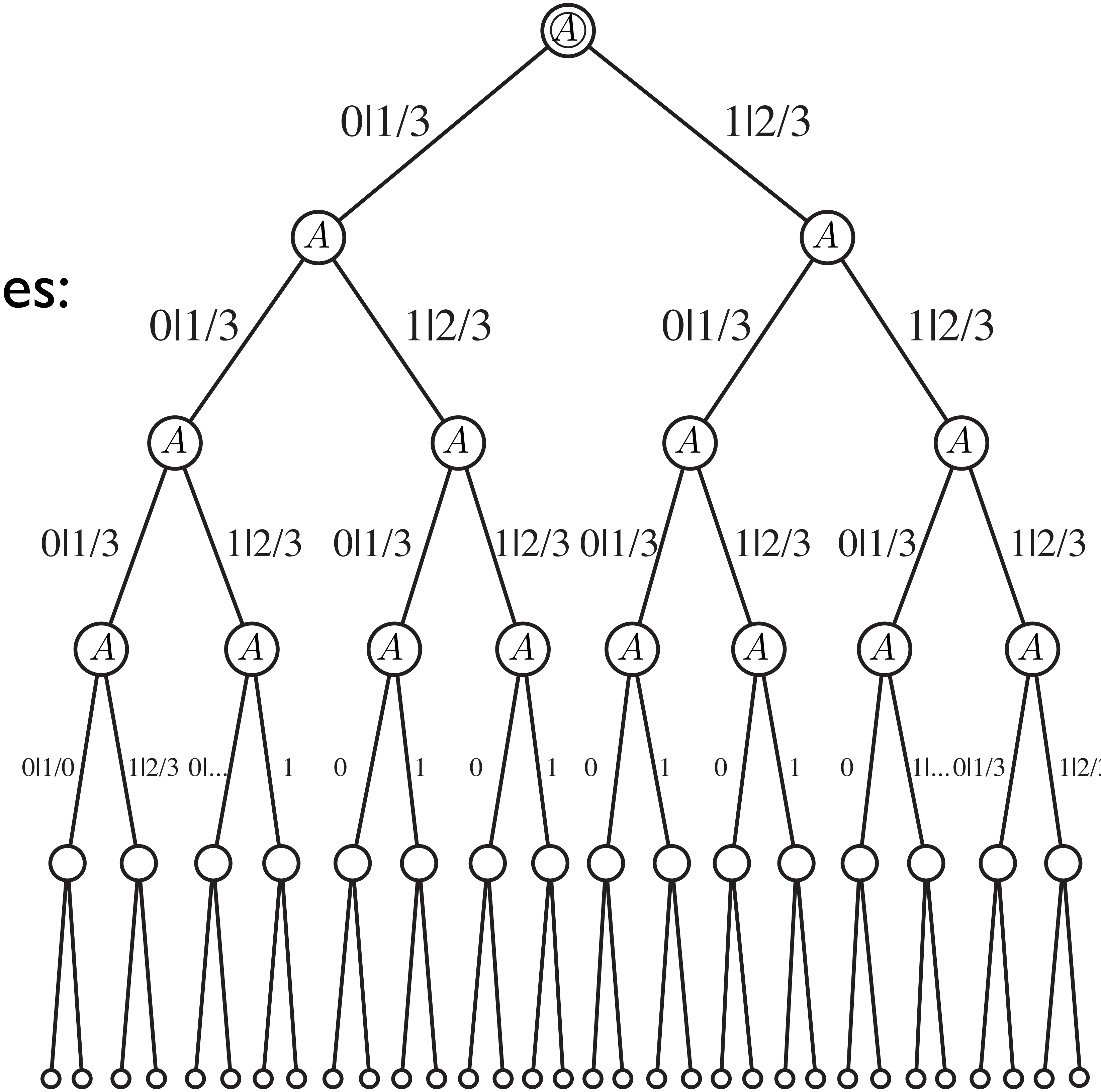
Call it state "A".

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Biased Coin ...

Label tree nodes
with state names:



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Biased Coin ...

$$\epsilon\text{M}\colon \mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

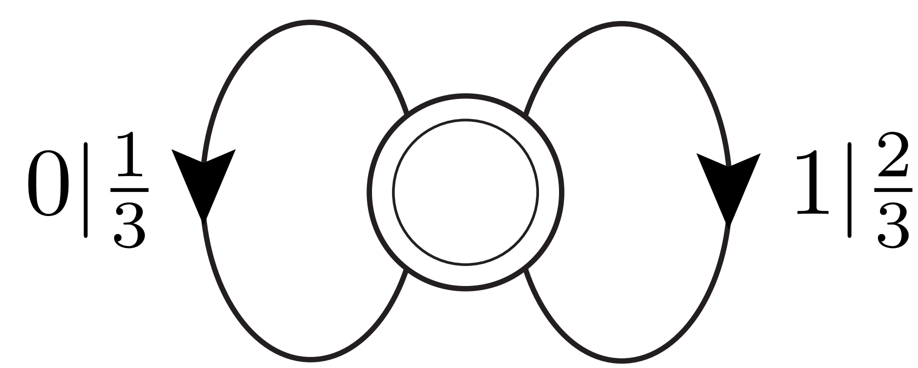
States:

$$\boldsymbol{\mathcal{S}} = \{\mathcal{S}_0 = \mathcal{A}^L\}$$

Transitions:

$$T^{(0)} = \left(\frac{1}{3}\right)$$

$$T^{(1)} = \left(\frac{2}{3}\right)$$



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Biased Coin ...

Causal state distribution: $p_{\mathcal{S}} = (1)$

Entropy Rate: $h_\mu = H\left(\frac{2}{3}\right) \approx 0.9183 \text{ bits per symbol}$

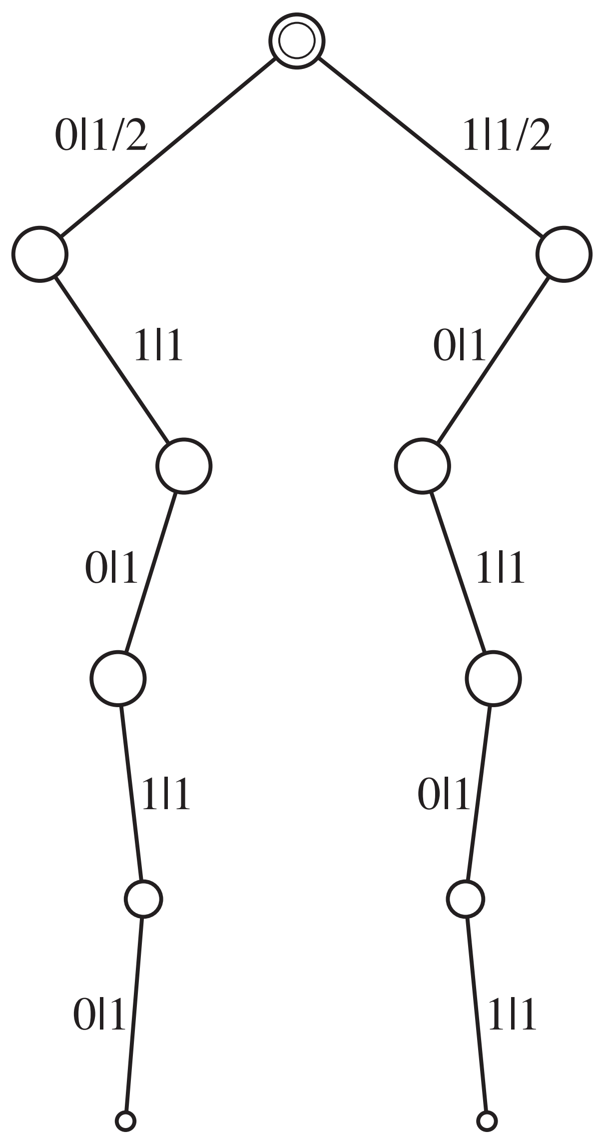
Statistical Complexity: $C_\mu = 0 \text{ bits}$

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-2 Process: $\ldots 01010101010101$

Parse Tree D = 5

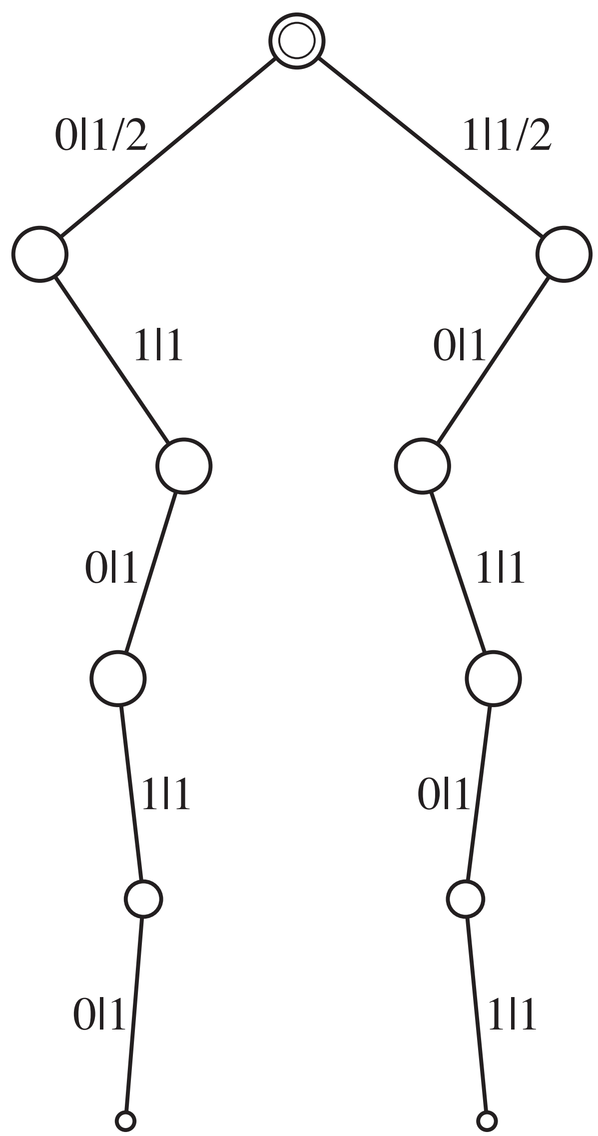


Machine Reconstruction ...

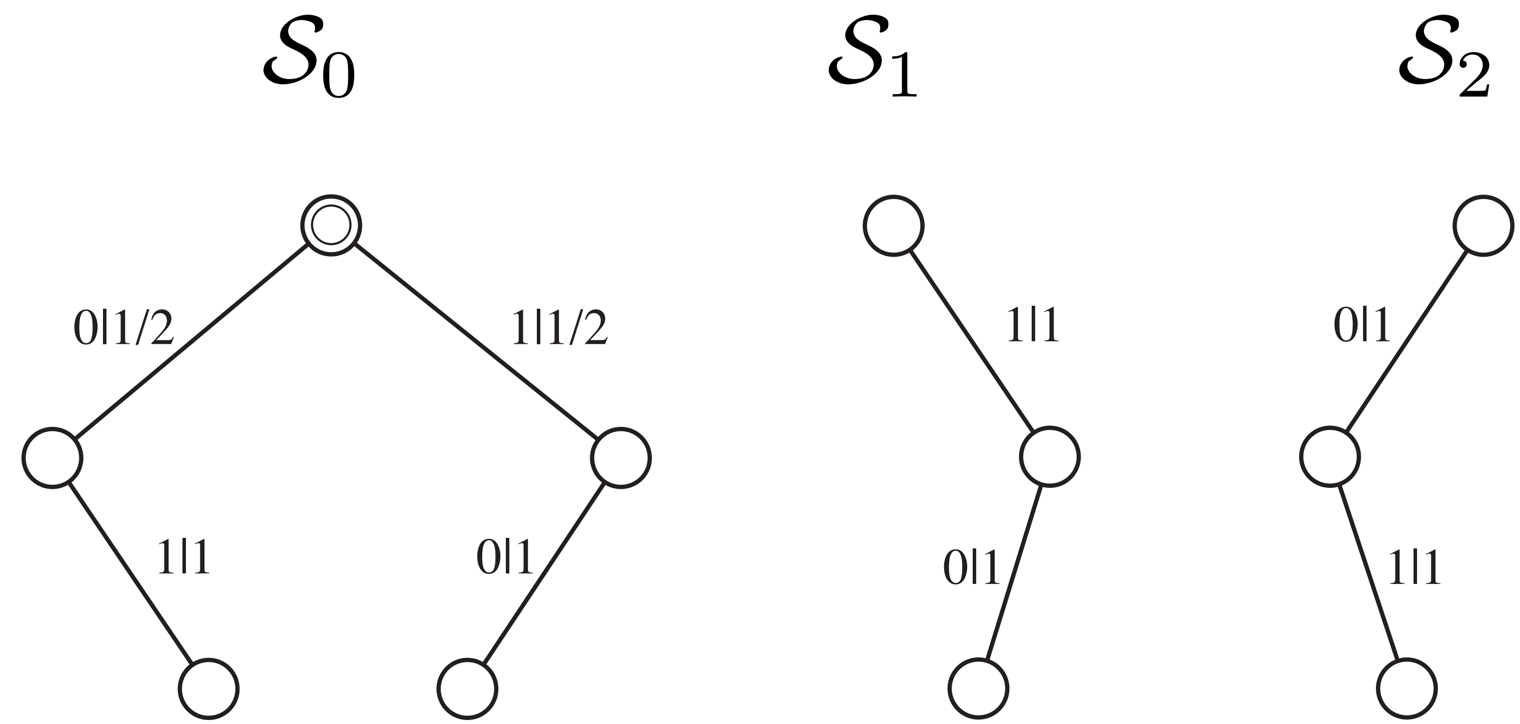
Examples (back to the Prediction Game) ...

Period-2 Process: $\ldots 01010101010101$

Parse Tree D = 5



Future Morphs at L = 2



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-2 Process: $\ldots 01010101010101$

Space of histories: $\overleftarrow{\mathbf{S}} = \{\overleftarrow{s}_0 = \ldots 101010, \overleftarrow{s}_1 = \ldots 010101\}$

Future morphs:

$$\{\overrightarrow{S}^1 | \lambda\} = \{101010\ldots, 010101\ldots\}$$
$$\{\overrightarrow{S}^1 | 0\} = \{101010\ldots\}$$
$$\{\overrightarrow{S}^1 | 1\} = \{010101\ldots\}$$
$$\{\overrightarrow{S}^1 | 10\} = \{101010\ldots\}$$
$$\{\overrightarrow{S}^1 | 01\} = \{010101\ldots\}$$

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-2 Process: $\ldots 01010101010101$

Morph distributions:

$$\Pr(0|\lambda) = \frac{1}{2}$$
$$\Pr(1|\lambda) = \frac{1}{2}$$

$$\Pr(1|0) = 1$$
$$\Pr(0|0) = 0$$

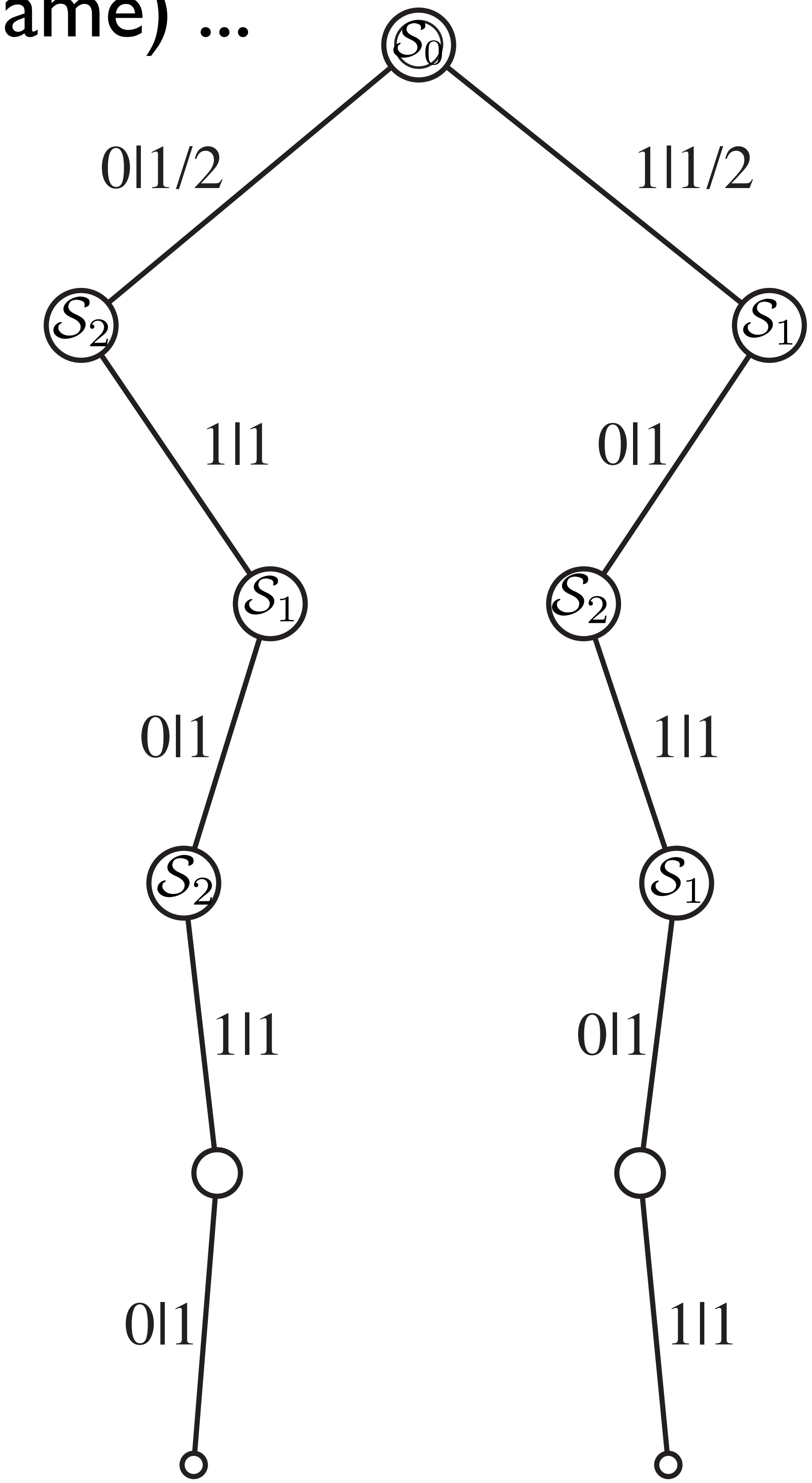
$$\Pr(1|1) = 0$$
$$\Pr(0|1) = 1$$

Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-2 Process ...

Label tree nodes:



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-2 Process ...

$$\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

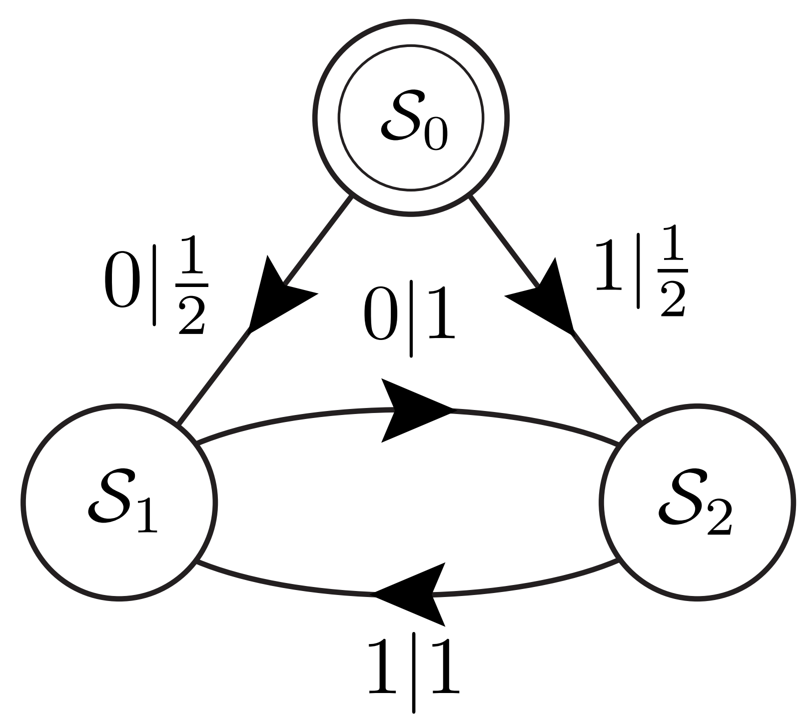
States:

$$\boldsymbol{\mathcal{S}} = \{\mathcal{S}_0 = \{\ldots 0101, \ldots 1010\}, \mathcal{S}_1 = \{\ldots 1010\}, \mathcal{S}_2 = \{\ldots 0101\}\}$$

Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

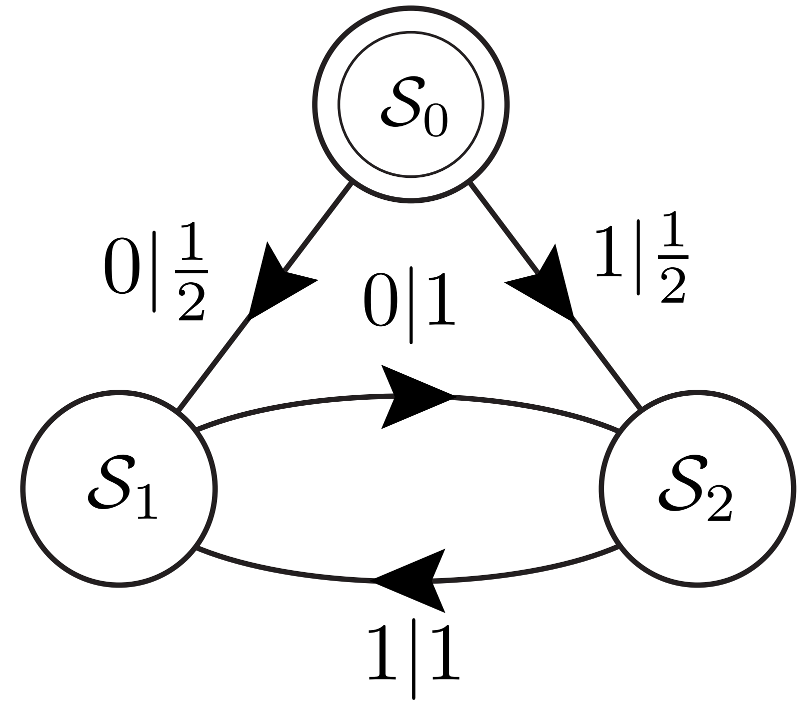
$$T^{(1)} = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$



Machine Reconstruction ...

Examples (back to the Prediction Game) ...

Period-2 Process ...



Causal State Distribution:

$$p_{\mathcal{S}} = \left(0, \tfrac{1}{2}, \tfrac{1}{2}\right)$$

Entropy rate: $h_\mu = 0$ bits per symbol

Statistical complexity: $C_\mu = 1$ bit

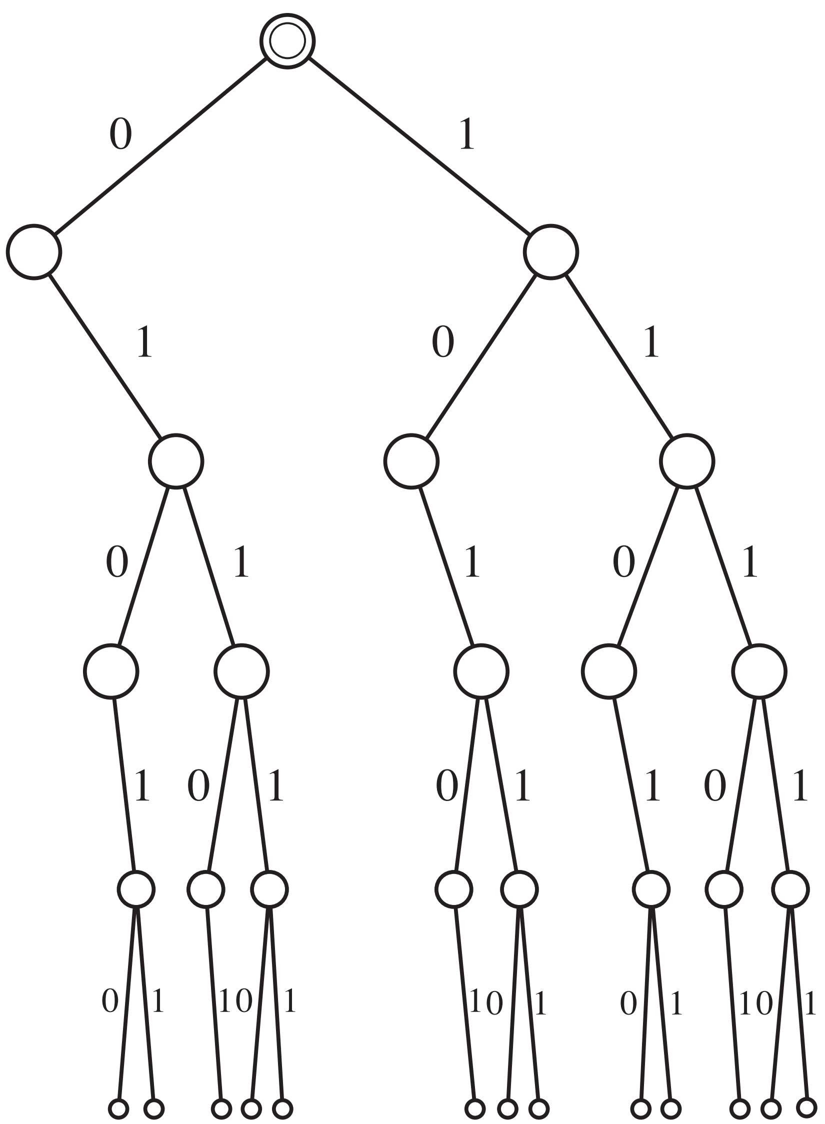
Machine Reconstruction ...

Examples ...

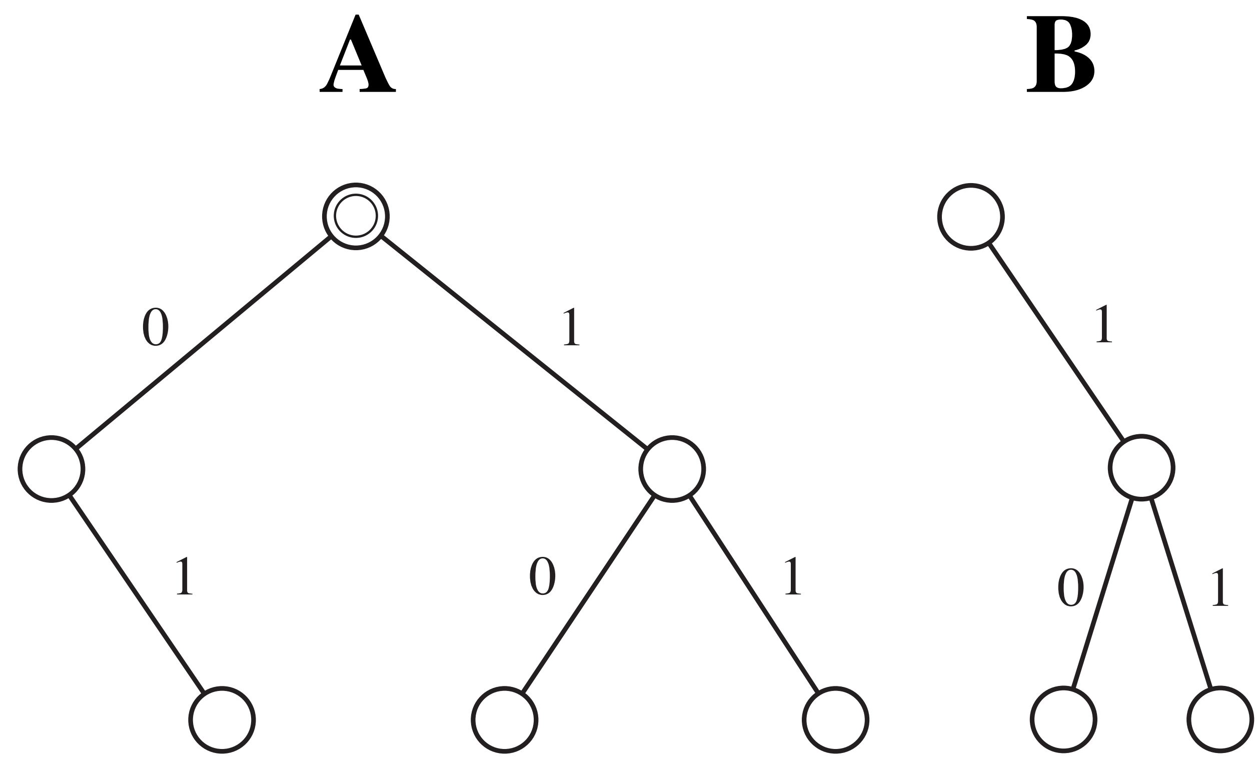
Golden Mean Process: Topological Reconstruction

(Only support of word distribution)

Parse Tree D = 5



Morphs at L = 2

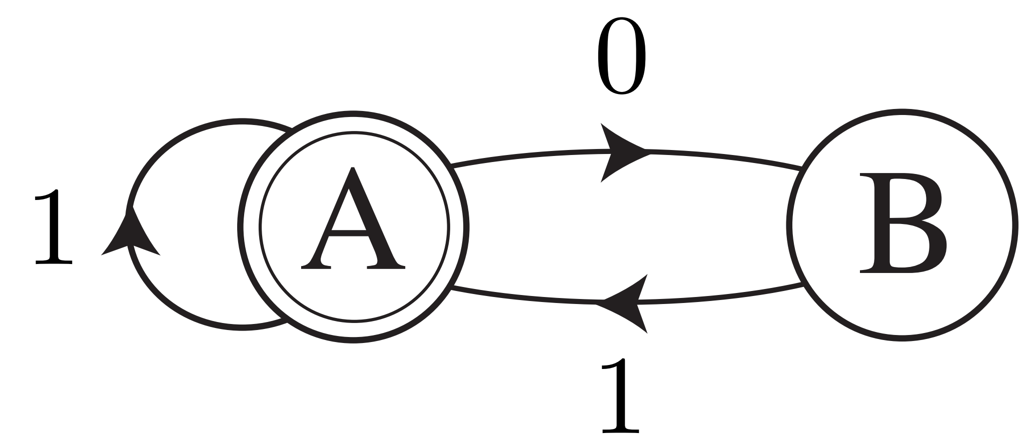


Machine Reconstruction ...

Examples ...

Golden Mean Process: Topological ϵ-machine

$$\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$



Topological Causal States:

$$\boldsymbol{\mathcal{S}} = \{A, B\}$$

Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

Machine Reconstruction ...

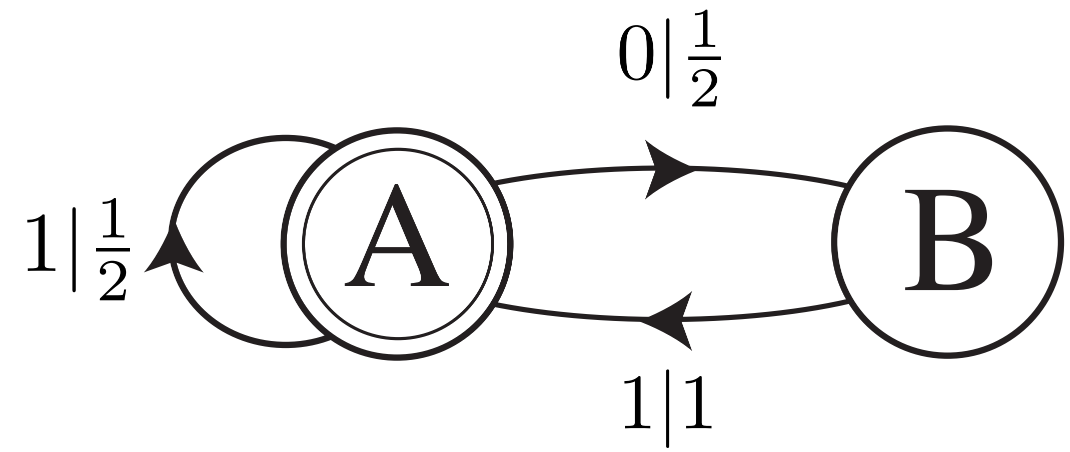
Examples ...

Golden Mean Process: Topological ϵ-machine

Probabilistic Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & 0 \end{pmatrix}$$



Causal State Distribution:

$$p_{\mathcal{S}} = \left(\tfrac{2}{3}, \tfrac{1}{3}\right)$$

Entropy rate: $h_\mu = \frac{2}{3}$ bits per symbol

Statistical complexity: $C_\mu = H(\frac{2}{3})$ bits

Machine Reconstruction ...

Examples ...

Golden Mean Process: Probabilistic reconstruction

(Capture full word distribution)

$$\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

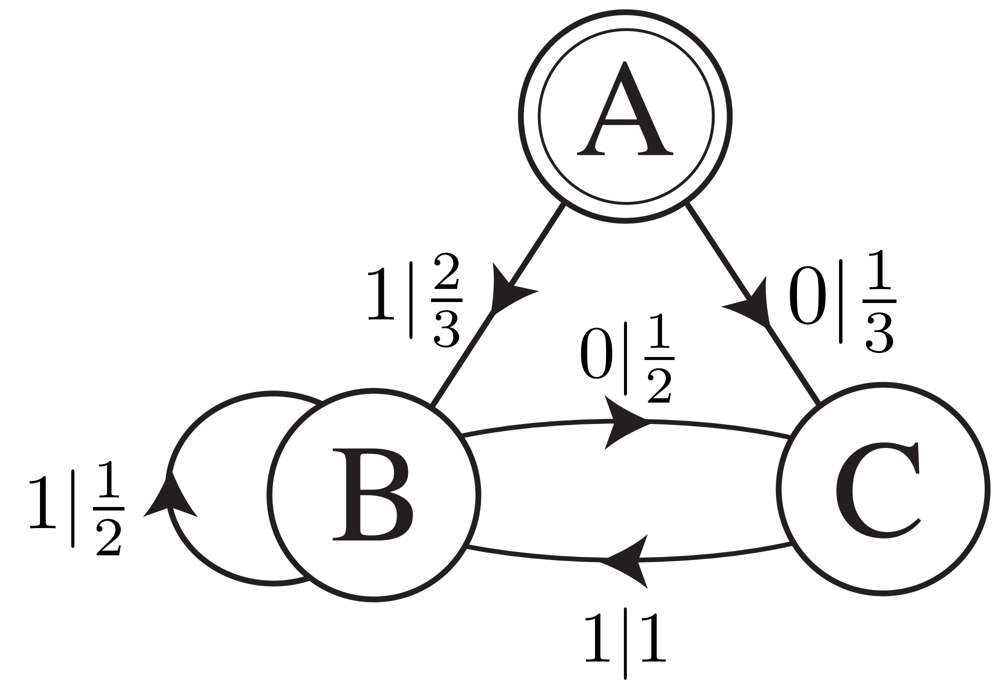
Causal States:

$$\boldsymbol{\mathcal{S}} = \{A, B, C\}$$

Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & 0 & 1/3 \\ 0 & 0 & 1/2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 0 & 2/3 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

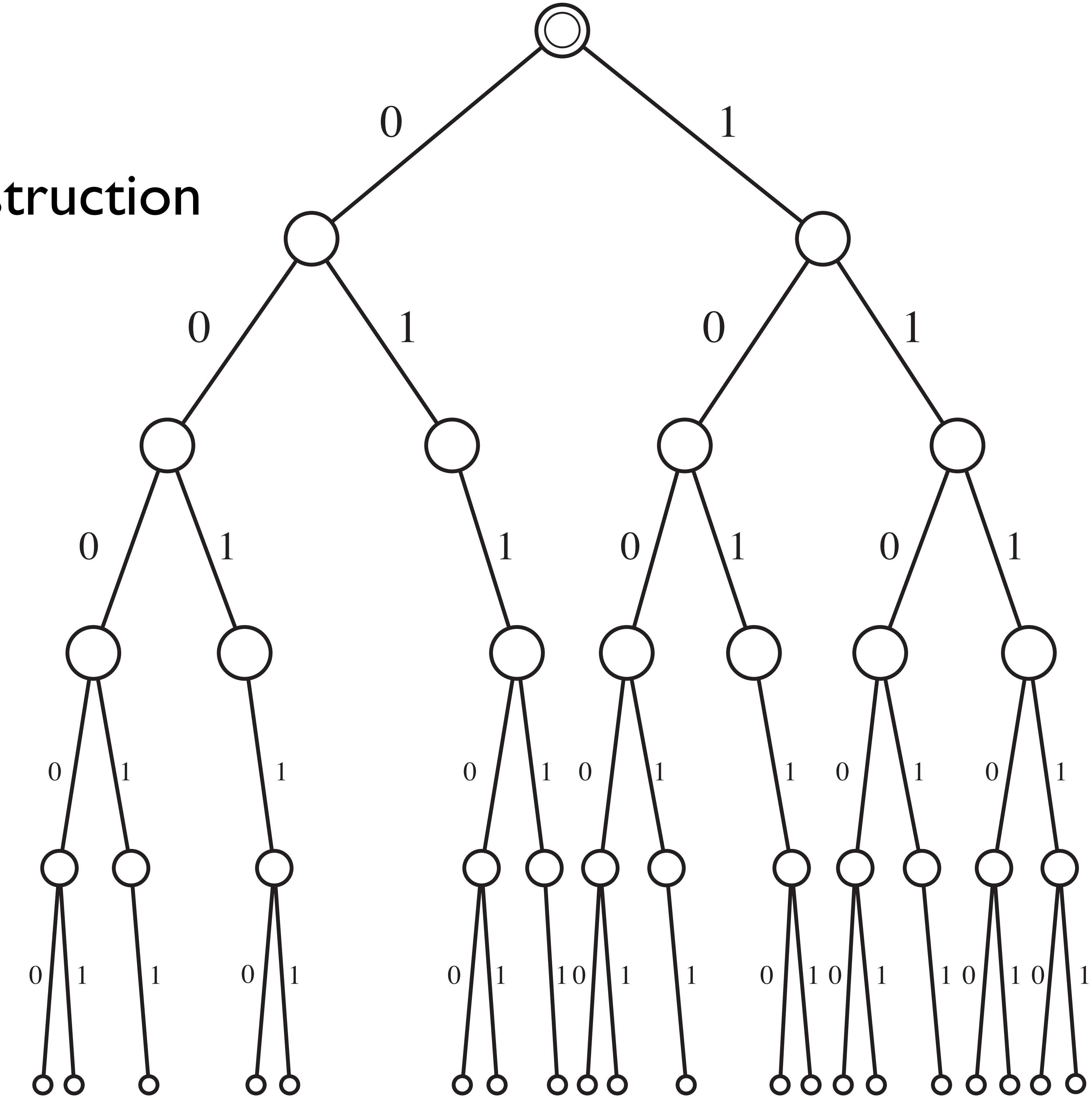


Machine Reconstruction ...

Examples ...

Even Process:
Topological reconstruction

Parse Tree D = 5

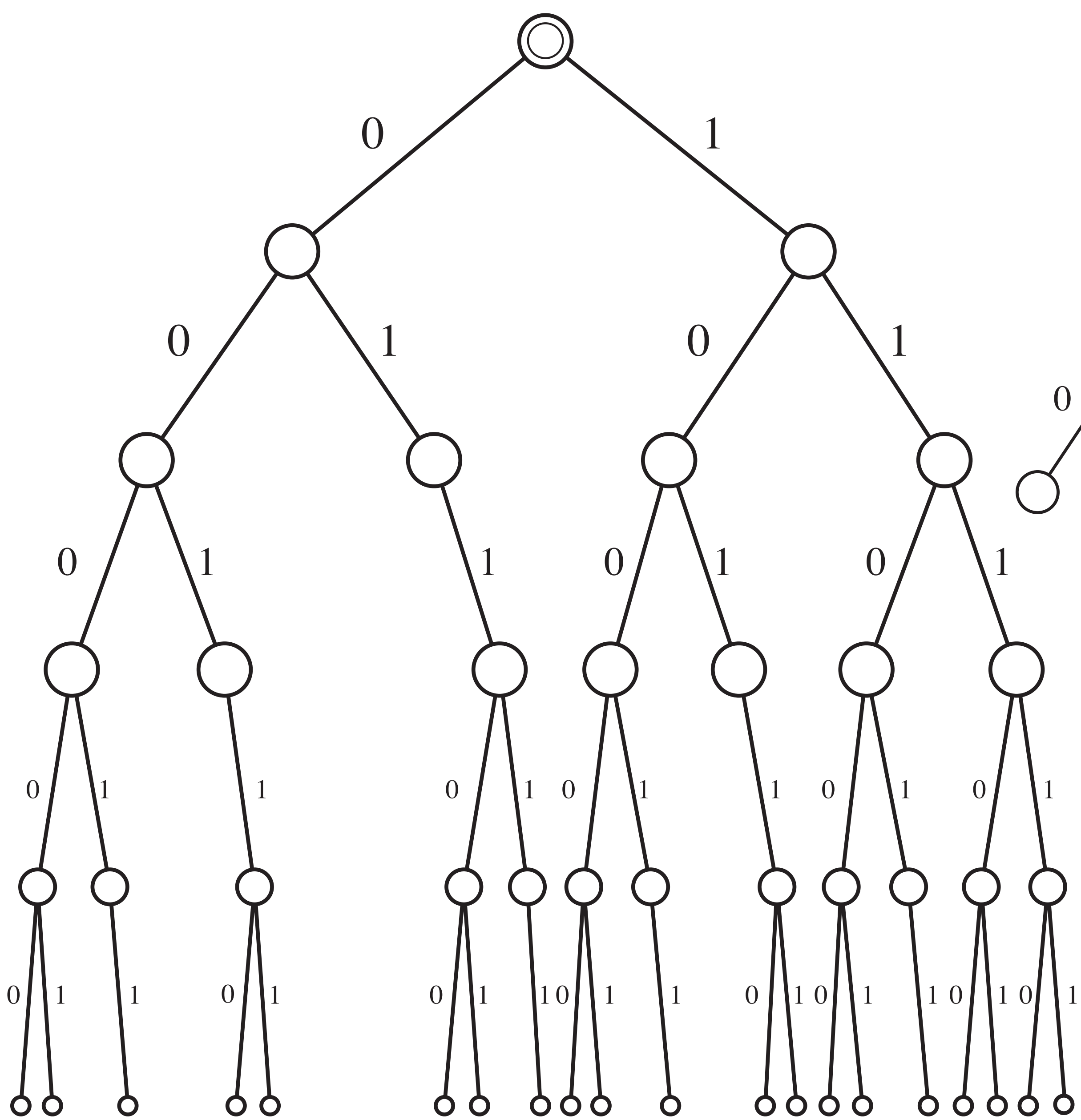


Machine Reconstruction ...

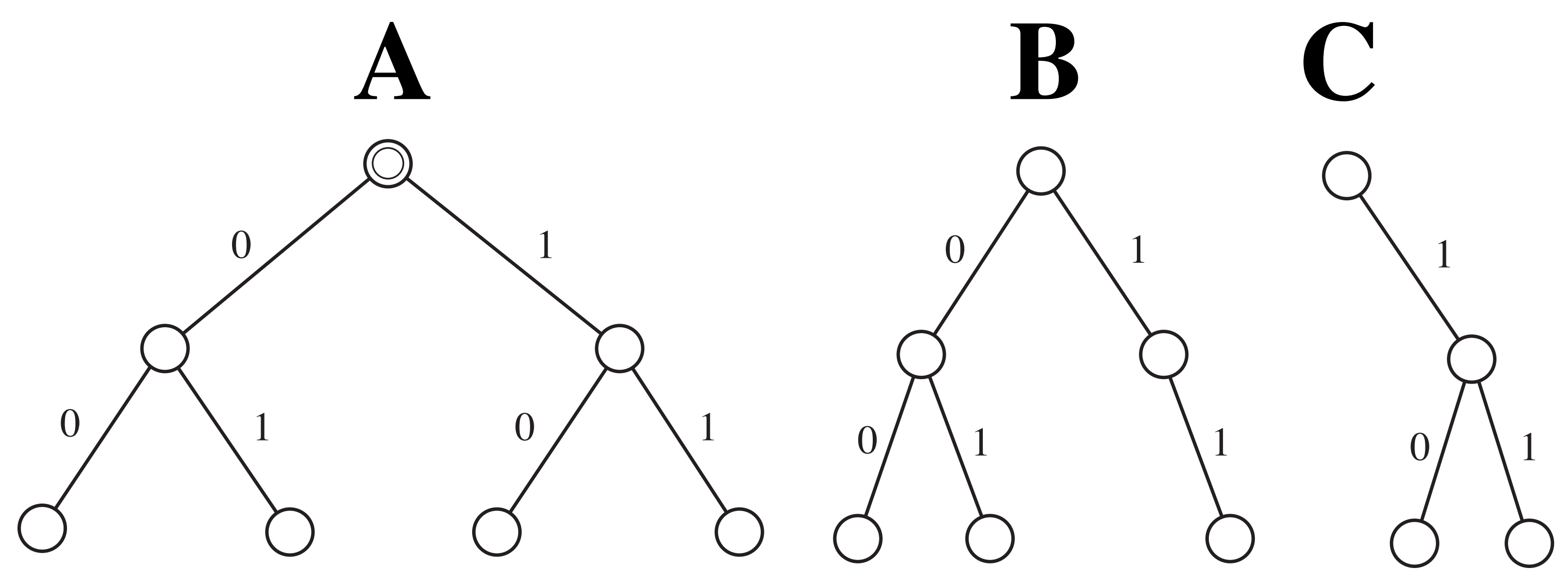
Examples ...

Even Process: Topological reconstruction

Parse Tree D = 5



Future Morphs at L = 2



Machine Reconstruction ...

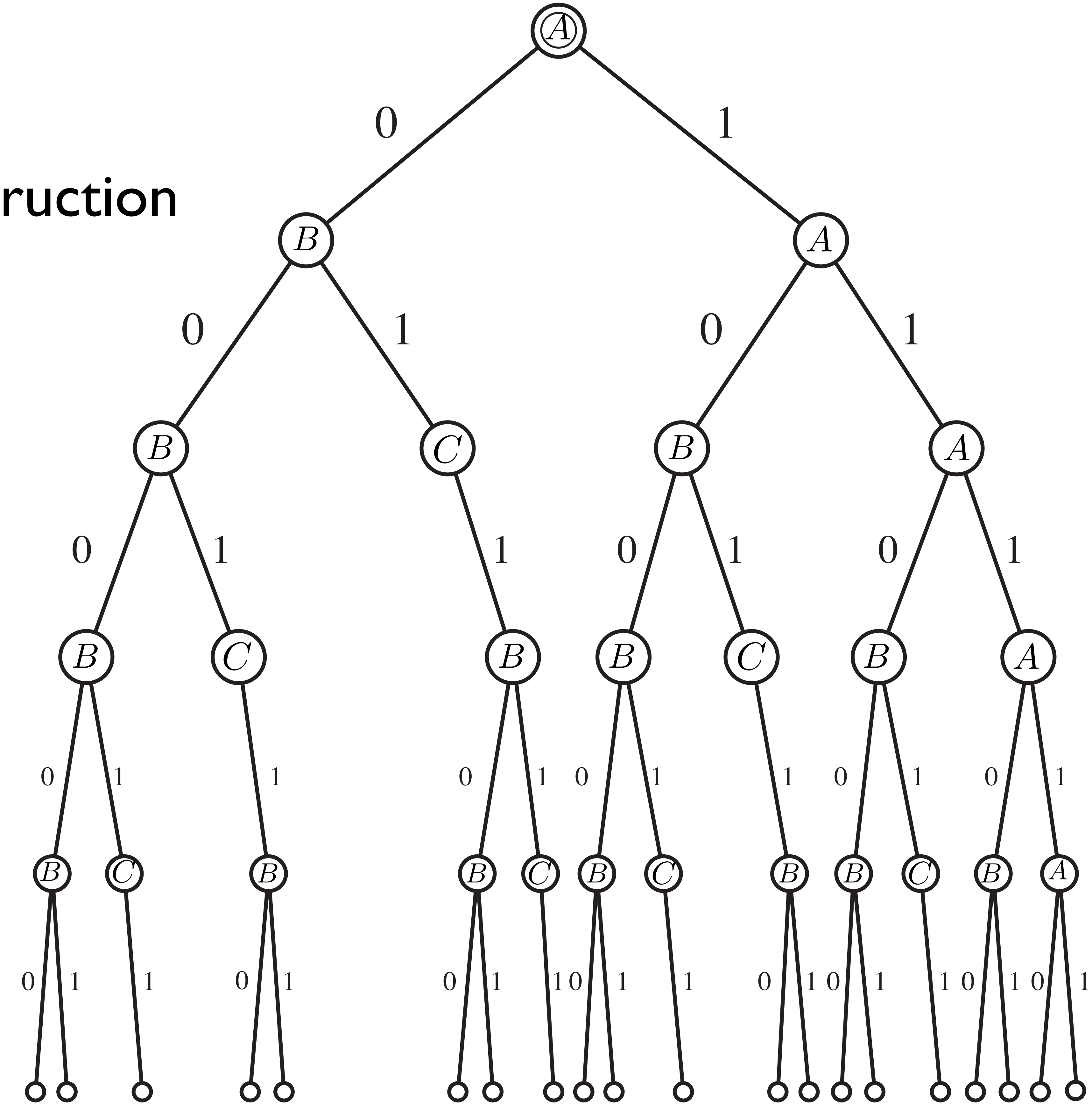
Examples ...

Even Process:

Topological reconstruction

Label tree nodes:

Parse Tree D = 5



Machine Reconstruction ...

Examples ...

Even Process: Topological reconstruction

Topological States:

$$\boldsymbol{\mathcal{S}} = \{A, B, C\}$$

Topological Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Machine Reconstruction ...

Examples ...

Even Process: Topological ε-machine

$$\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

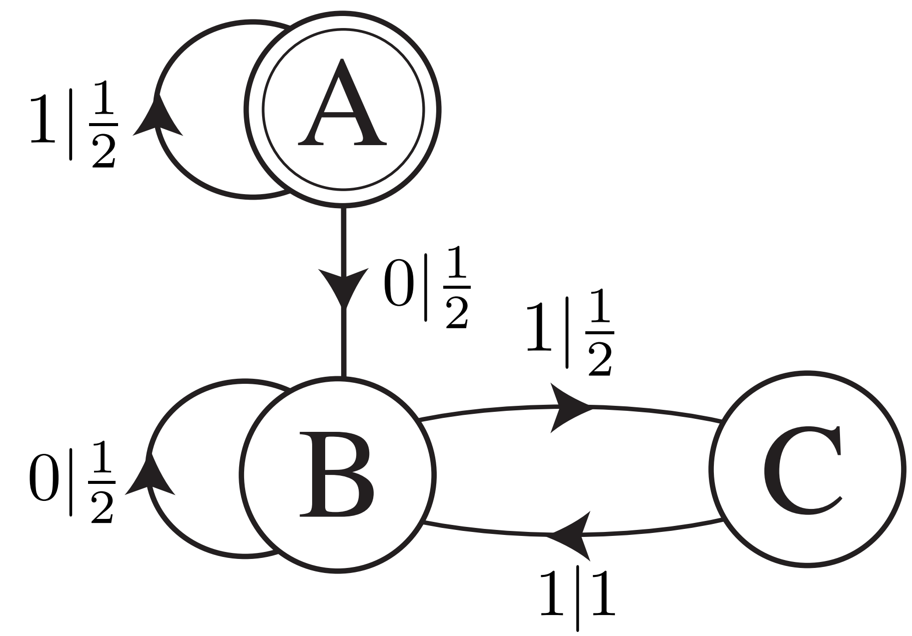
States:

$$\boldsymbol{\mathcal{S}} = \{A, B, C\}$$

Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$T^{(1)} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$



Machine Reconstruction ...

Examples ...

Even Process: Probabilistic reconstruction

$$\mathcal{M} = \left\{ \boldsymbol{\mathcal{S}}, \{T^{(s)}, s \in \mathcal{A}\} \right\}$$

States:

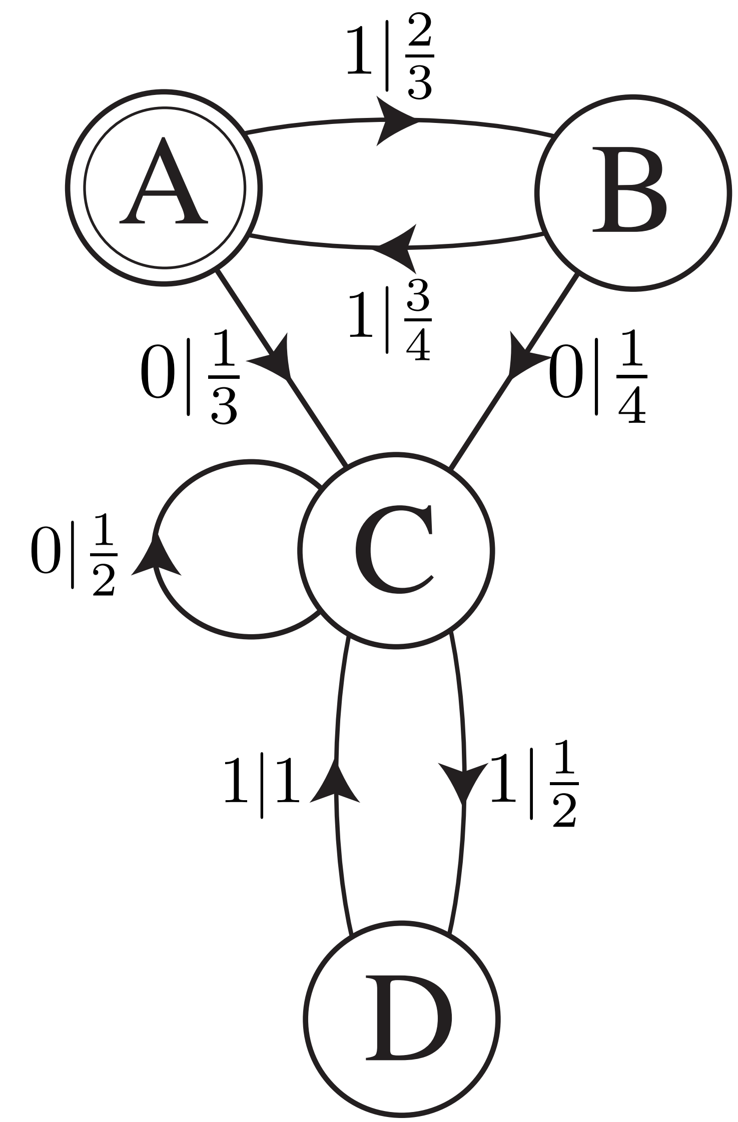
$$\boldsymbol{\mathcal{S}} = \{A, B, C, D\}$$

Transitions:

$$T^{(0)} = \begin{pmatrix} 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad T^{(1)} = \begin{pmatrix} 0 & \frac{2}{3} & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Entropy rate: $h_\mu = \frac{2}{3}$ bits per symbol

Statistical complexity: $C_\mu = H(\frac{2}{3})$ bits



Machine Reconstruction ...

Reading for next lecture:

CMR article *CMPPSS*

Homework: $\epsilon\mathrm{M}$ reconstruction for GMP, EP, & RRXOR

Helpful? Tree & morph paper at:

http://csc.ucdavis.edu/~cmg/compmech/tutorials.htm