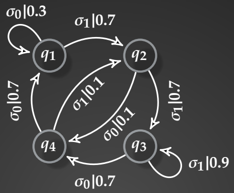
Probabilistic Finite State Automata

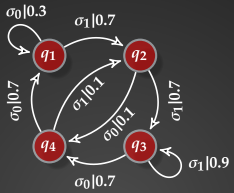
Models For Quantized Stationary Ergodic Stochastic Processes



State Set	Q	$q_1, \cdots, q_4$
Alphabet	Σ	σ_0, σ_1
Morph probabilities	$\widetilde{\pi} : Q \times \Sigma^\star \rightarrow [0,1]$	$\begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.7 & 0.3 \end{pmatrix}$
Stationary distribution	$\wp^\star$	$\wp^\star P = \wp^\star$

Probabilistic Finite State Automata

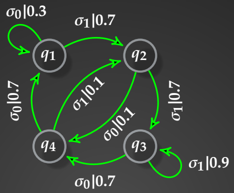
Models For Quantized Stationary Ergodic Stochastic Processes



State Set	Q	$q_1, \cdots, q_4$
Alphabet	Σ	σ_0, σ_1
Morph probabilities	$\widetilde{\pi} : Q \times \Sigma^\star \to [0,1]$	$\begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.7 & 0.3 \end{pmatrix}$
Stationary distribution	$\wp^\star$	$\wp^\star P = \wp^\star$

Probabilistic Finite State Automata

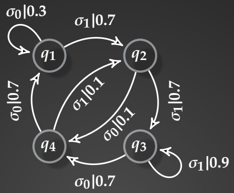
Models For Quantized Stationary Ergodic Stochastic Processes



State Set	Q	$q_1, \cdots, q_4$
Alphabet	Σ	σ_0, σ_1
Morph probabilities	$\widetilde{\pi} : Q \times \Sigma^\star \rightarrow [0,1]$	$\begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.7 & 0.3 \end{pmatrix}$
Stationary distribution	$\wp^\star$	$\wp^\star P = \wp^\star$

Probabilistic Finite State Automata

Models For Quantized Stationary Ergodic Stochastic Processes

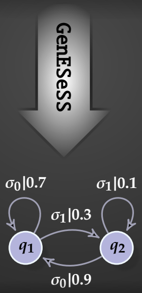


State Set	Q	$q_1, \cdots, q_4$
Alphabet	Σ	σ_0, σ_1
Morph probabilities	$\widetilde{\pi} : Q \times \Sigma^\star \rightarrow [0,1]$	$\begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.7 & 0.3 \end{pmatrix}$
Stationary distribution	$\wp^\star$	$\wp^\star P = \wp^\star$

Explicit Construction

Algorithm **GenESeSS**

0100100010010000100000010000100001010000100001010000101001000000001000001000000100100100000000010
0010100010100010001000100000010001010110000101000101000110010000000100000001000000101010000000010001
0010100100011000110010100010000000010001000010010000110010001000011001000010010000001010000000000101
0100010000000000011110000000010000101000100010010000000001000001000001010001010101010010010100100010 1
0010000000100001100010000000010010100000000010000001000000000100101000000100001000 1000 ···

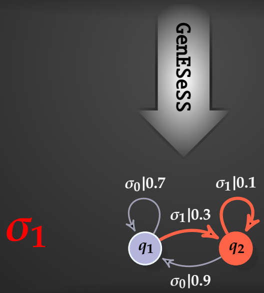


generator Extraction from Self-Similar Semantics

Explicit Construction

Algorithm **GenESeSS**

010010001001000010000000100001000010100001000010100001010010000000001000000100000010010010000000000010
001010001010001000100010000000100010101100001010001010001100100000001000000010000000101010000000010001
001010010001100011001010001000000000100010000100100001100100010000110010000100100000001010000000000101
010001000000000000111100000000100001010001000100100000000010000001000000101000101010101001001010010001 01
0010000000010000110001000000000100101000000000010000001000000000100101000000100001000100 0 $\cdots$

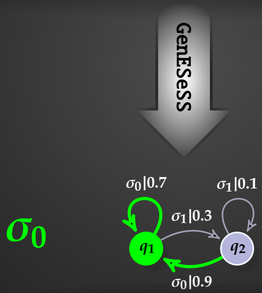


generator Extraction from Self-Similar Semantics

Explicit Construction

Algorithm **GenESeSS**

0100100010010000100000001000010000101000010000101000010100100000000010000001000000100100100000000000010
0010100010100010001000100000001000101011000010100010100011001000000010000000100000010101000000010001
0010100100011000110010100010000000010001000010010000110010001000011001000010010000001010000000000101
0100010000000000011110000000010000101000100010010000000001000000100000101000101010101001001010010000101
0010000000100001100010000000010010100000000010000001000000001001010000001000010001000 ···

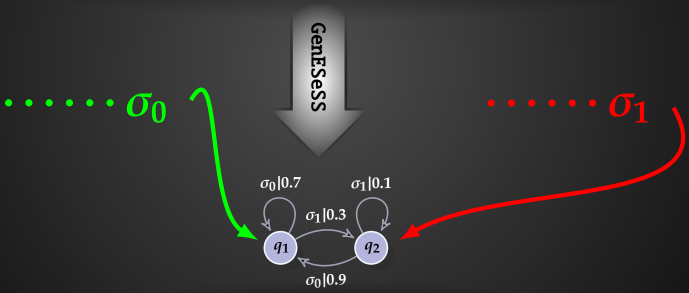


genErator Extraction from Self-Similar Semantics

Explicit Construction

Algorithm **GenESeSS**

0100100010010000100000010000100001010000100001010000101001000000000100000010000001001001000000000010
0010100010100010001000100000010001010110000101000101000110010000001000000010000001010100000000010001
0010100100011000110010100010000000010001000010010000110010001000011001000010010000001010000000000101
0100010000000000011110000000010000101000100010010000000001000000100000101000101010101001001010010000101
0010000000010000110001000000000100101000000000100000010000000010010100000010000100001000 ···

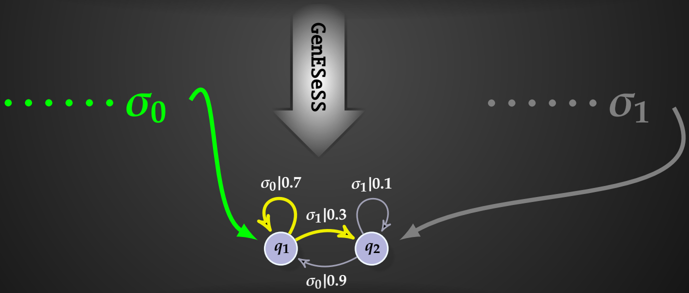


generator **E**xtraction from **Se**lf-**S**imilar **S**emantics

Explicit Construction

Algorithm **GenESeSS**

0100100010010000100000010000100001010000100001010000101001000000000100000010000000100100100000000000010
0010100010100010001000100000010001010110000101000101000110010000000100000001000000101010000000010001
0010100100011000110010100010000000001000100001001000011001000100001100100001001000000101000000000101
0100010000000000111100000000100001010001000100100000000010000010000010100010101010100100101001000101
001000000001000011000100000000100101000000000100000010000000010010100000100001000100 $\cdots$

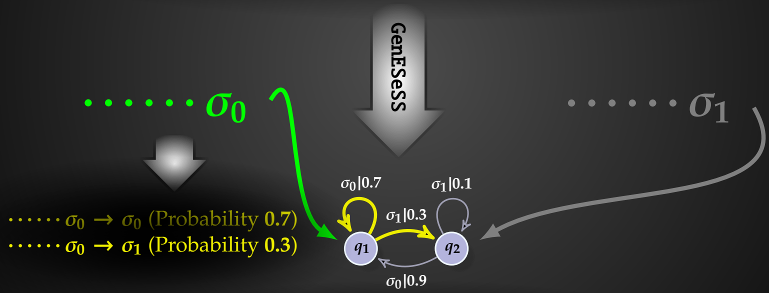


generator Extraction from Self-Similar Semantics

Explicit Construction

Algorithm **GenESeSS**

0100100010010000100000001000010000101000010000101000010100100000000100000010000000100100100000000000010
0010100010100010001000100000001000101011000010100010100011001000000010000000100000001010100000000010001
0010100100011000110010100010000000001000100001001000011001000100001100100001001000000101000000000101
0100010000000000011110000000010000101000100010010000000001000000100000101000101010101001001010010001000101
0010000000010000110001000000000100101000000000010000001000000000100101000000100001000100 0 ···

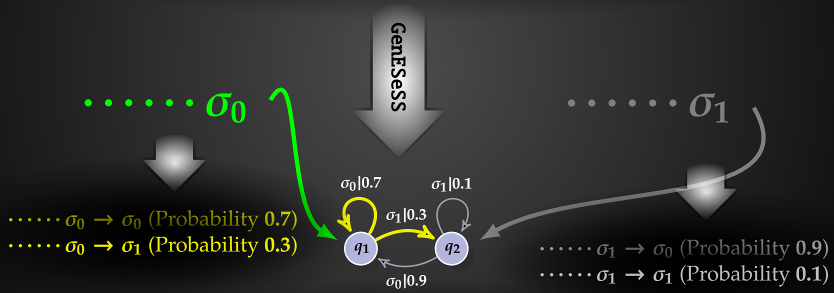


generator Extraction from Self-Similar Semantics

Explicit Construction

Algorithm **GenESeSS**

0100100010010000100000010000100001010000100001010000101001000000000100000010000000100100100000000000010
0010100010100010001000100000001000101011000010100010100011001000000010000000100000010101000000010001
0010100100011000110010100010000000001000100001001000011001000100001100100001001000000101000000000101
0100010000000000011110000000010000101000100010010000000001000000100000101000101010101001001010010001011
00100000000100001100010000000001001010000000000100000010000000001001010000001000010001000 $\cdots$

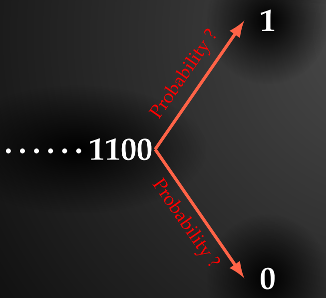


generator Extraction from Self-Similar Semantics

Symbolic Derivatives

Estimating Average Immediate Future

0100100010010000100000010000100001010000100001010000101001000000000100000100000010010010000000000010
0010100010100010001000100000010001010110000101000101000110010000000100000001000000101010000000010001
0010100100011000110010100010000000001000100001001000011001000100001100100001001000000101000000000101
0100010000000000011110000000010000101000100010010000000010000010000010100010101010100100101001000101
00100000000100001100010000000001001010000000000100000010000000001001010000001000010001000 ···

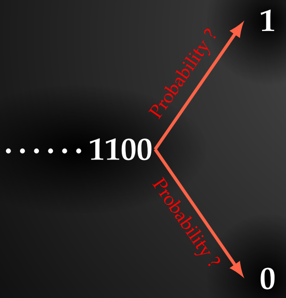


$$\phi^s_\omega = \begin{pmatrix} Pr(\sigma_0) \\ Pr(\sigma_1) \end{pmatrix}$$

where $\omega = \sigma_1\sigma_1\sigma_0\sigma_0$

Symbolic Derivatives

Estimating Average Immediate Future



$$\phi^s_\omega = \begin{pmatrix} Pr(\sigma_0) \\ Pr(\sigma_1) \end{pmatrix}$$

where $\omega = \sigma_1\sigma_1\sigma_0\sigma_0$

Symbolic Derivatives

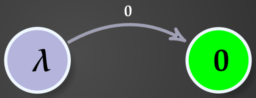
Estimating Average Immediate Future

λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮

Symbolic Derivatives

Estimating Average Immediate Future

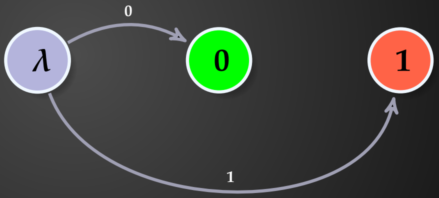
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

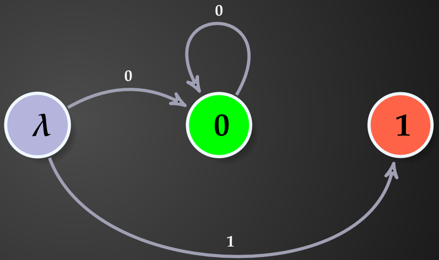
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

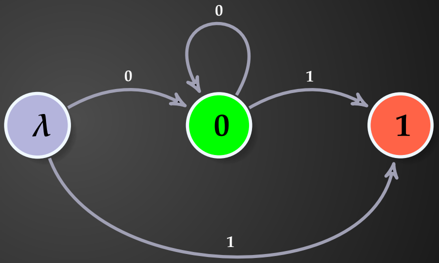
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

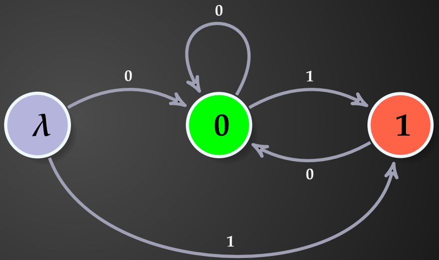
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

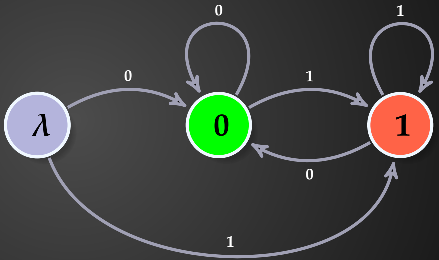
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

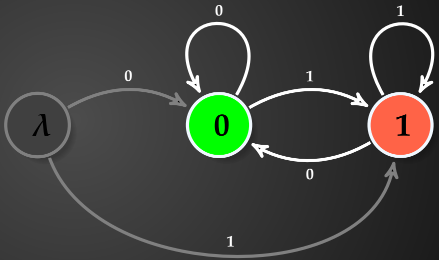
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

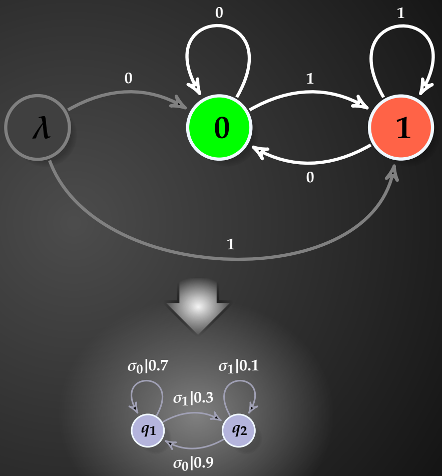
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

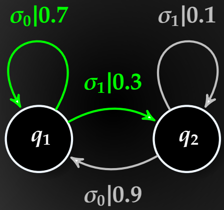
λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Symbolic Derivatives

Estimating Average Immediate Future

λ	0.750285	0.249715
0	0.700112	0.299888
1	0.901009	0.0989909
00	0.699844	0.300156
01	0.899111	0.100889
10	0.700711	0.299289
11	0.918285	0.0817152
000	0.699004	0.300996
001	0.898769	0.10123
010	0.701038	0.298962
011	0.917181	0.0828194
100	0.701763	0.298237
101	0.899911	0.100089
110	0.697797	0.302203
111	0.930693	0.0693069
0000	0.699284	0.300716
0001	0.902025	0.0979754
⋮	⋮	⋮



Merging of Symbolic Derivatives Under ϵ Resolution

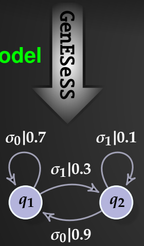
Minimizing Modeling Error

Combining **GenESeSS** with data smashing

0100100010010000100000001000010000101000010000101000010100100000000010000001000000010010010000000 0
00101000101000100010001000000010001010110000101000101000110010000000100000001000000010101000000000
00101001000110001100101000100000000010001000010010000110010001000011001000010010000000101000000
0100010000000000011110000000010000101000100010010000000001000001000001010001010101010010010100 10
0010000000010000110001000000000100101000000000100000010000000010010100000100001000 1000 ⋯

Infer Model

PFSA



Minimizing Modeling Error

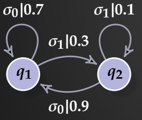
Combining **GenESeSS** with data smashing

0100100010010000100000010000100001010000100001010000101001000000000100000100000010010010000000
0010100010100010001000100000010001010110000101000101000110010000000100000001000000101010000000
0010100100011000110010100010000000010001000010010000110010001000011001000010010000000101000000
0100010000000000111100000000100001010001000100100000000010000010000010100010101010100100101010010
0010000000100001100010000000001001010000000000100000010000000001001010000010000100001000 · · ·

Infer Model

GenESeSS

PFSA



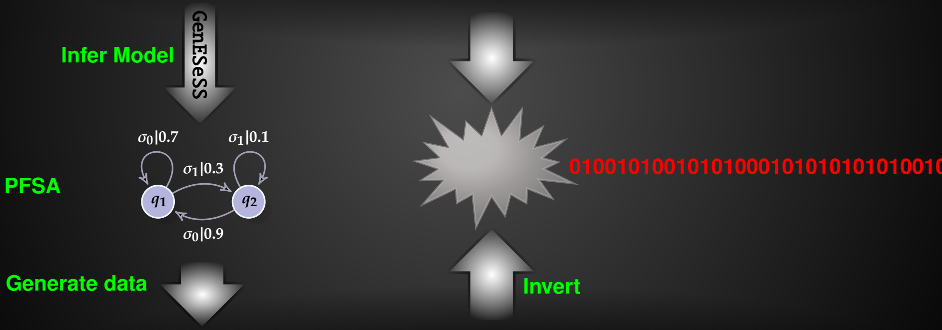
Generate data

000101100101010000000101000000100010100101001000001001000000000011010010010100101010010010001
100100010100000001000100001000000000101000000000100100010010010000100100010010100000000001000101
000000000000101001000010010100101001000100000101000100000001000000000000000010001100100100000101
000000100000000010100001011010000000001000000000101100000000111000000100100001100000000110000100010
1010010001010100100101000000000000100000001001 · · ·

Minimizing Modeling Error

Combining **GenESeSS** with data smashing

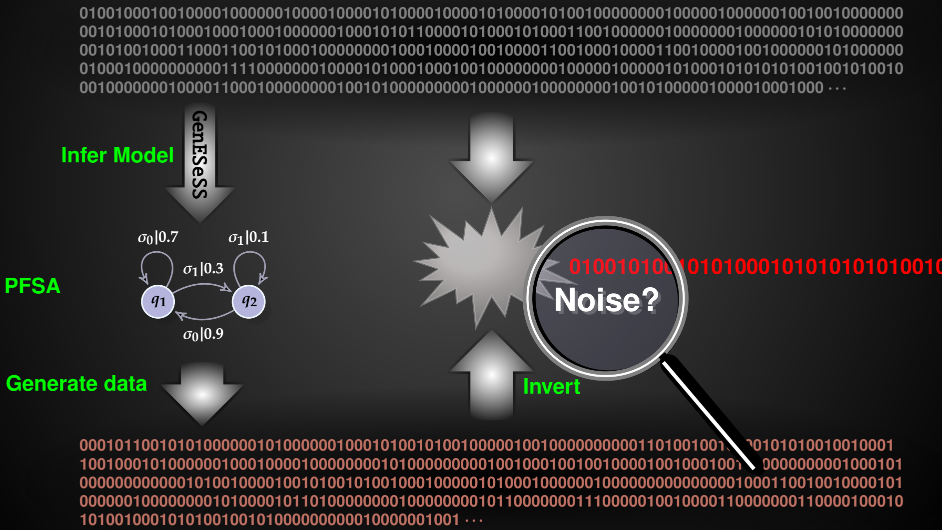
0100100010010000100000010000100001010000100001010000101001000000000100000010000001001001000000000
0010100010100010001000100000010001010110000101000101000110010000000100000001000000101010000000
0010100100011000110010100010000000001000100001001000011001000100001100100001001000000101000000
0100010000000000011110000000010000101000100010010000000001000001000001010001010101010010010100100
0010000000100001100010000000001001010000000000100000010000000010010100000100001000100 0 · · ·



0001011001010100000010100000001000101001010010000010010000000000011010010010100101010010010001
1001000101000000100010000100000000101000000000100100010010010000100100010010100000000001000101
0000000000001010010000100101001010010001000001010001000000100000000000000001000110010010000101
0000001000000000101000010110100000000010000000010110000000011100000100100001100000000110000100010
101001000101010010010100000000000010000001001 · · ·

Minimizing Modeling Error

Combining **GenESeSS** with data smashing



Probably Approximately Correct

Time Complexity

Assuming $|s| > |\Sigma|$, the asymptotic time complexity of **GenESeSS** is:

$$\mathscr{T} = O\left(\frac{|s||\Sigma|}{\epsilon}\right)$$

PAC-Learnability

Ergodic, stationary quantized stochastic processes with finite number of causal states has the following property:

For $\epsilon, \eta > 0$, and for every sufficiently long sequence s generated by QSP $\mathcal{H}$, **GenESeSS** computes $\mathcal{P}'_{\mathcal{H}}$ as an estimate for $\mathcal{P}_{\mathcal{H}}$ with:

$$Pr(\Theta(\mathcal{P}_{\mathcal{H}}, \mathcal{P}'_{\mathcal{H}}) \leqq \epsilon) \geqq 1 - \eta$$

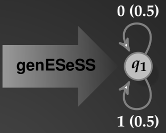
Asymptotic runtime is polynomial in $1/\epsilon, 1/\eta, |s|$, and sample complexity is:

$$|s| = O(C_0^{|Q|} \frac{1}{\epsilon} \log \frac{1}{\eta})$$

Learning Quantized Stochastic Processes

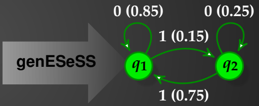
Algorithm GenESeSS

```
0011011110011111001100011000111010011111100101010011011000100100010011110001 0011
1101111011000100011011001111001001100001111011100011111010010101001110111011 0111
1000000001110111011100010000111011110000100011100010001110011001001010110000 0010
0111111111001100000000110110100010101100010000100010100110110010010011110000 0010
1100111110000110100111110001110001011101000101111101101011011000010010001111 0010
1110011101100011101101001101101111001101111110010011101111111101101100100100 0011
0101000011010010000101001110110101100000011000011111011010101100111101010001 0101
```



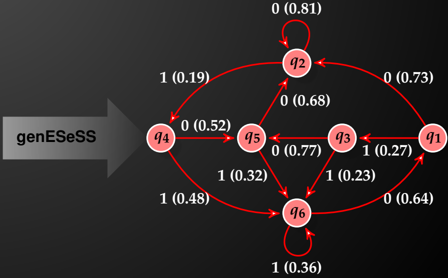
Flat White Noise (Fair Coin Toss Sequence)

```
0110000000000110000100110100011000000001100000011011000000000000000110001101 1011
0110000110110110000000011011000100101100001010000011000011000000000000000000 0000
0000000000001001000000000110000000011000011000000001100111010000000000000010 1000
0011001001000000111100001100110011000000111010000001100110010111000000000011 0110
0000001111011000000000111010000011000010100000000000110000000000101000000000 0000
0000000001100011101000000000011001100110000000001001110110000000000000000000 0000
0000000000000000000000000110000110110011000000000110000000000000011000000010 100
0001010000110011000000110010100000000110110000000000000110001100000000000000 0000
```



Stochastic process with 2 causal states

```
0000011000000000101000000000101000000000100110010000000000000000110110000000000011000
0000000000100111000001100000100100110110000000111001000000110011001100000000000000
0000010101011100011111110110000011000000101000000000011011000110000000011000000000
1100110101001100000000000000101101110000011000110000000101110110011001100110011001100
0000110000000000000000000001100000000110000000001100000000011000000000110011001
0000011011110000000001100000000000000000000000001100000010100000000000001010101010000
1110000000011000000000000011000000000011000000000110000000000000000000000000011000
0000001010000011110001100001100011000000000000000000000000000000001111000101010101000
```



Stochastic process with 6 causal states

Learning Quantized Stochastic Processes

Algorithm GenESeSS

```
0011011110011110011000110001110100111111001010100110110001001000100111100010011
1101111011000100011011001111001001100001111011100011111010010101001110111011
1000000001110111011100010000111011110000100011100010001110011001001010110000010
0111111111001100000000110110100010101100010000100010100110110010010011110000010
1100111110000110100111110001110001011101000101111101101010110110000100100011110010
1110011101100011101101001101101110011011111100100111011111111011011001001000011
0101000011010010000101001110110101100000011000011111011010101100111101010001010
```

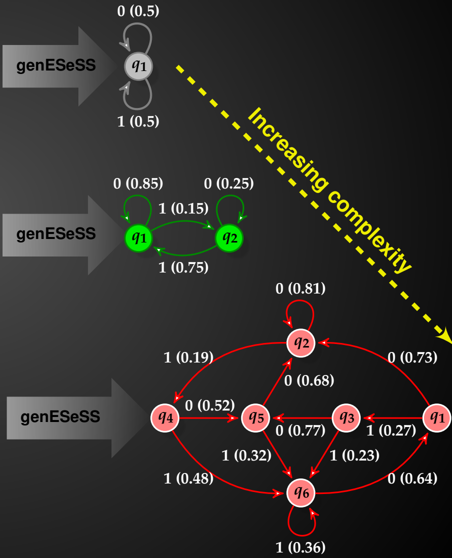
Flat White Noise (Fair Coin Toss Sequence)

```
0110000000000110000100110100011000000000110000001101100000000000000000110001101101
011000011011011000000000110110001001011000010100000110000110000000000000000000000
00000000000010010000000000110000000001100001100000000110011101000000000000001010000
00110010010000000111100001100110011000000111010000001100110010111000000000110110
0000001111011000000000111010000011000010100000000000110000000000101000000000000
000000000110001110100000000000110011001100000000010011101100000000000000000000000
0000000000000000000000000110000110110011000000000000100000000000000011000000010100
00010100001100110000000110010100000000001101100000000000000110001100000000000000000000
```

Stochastic process with 2 causal states

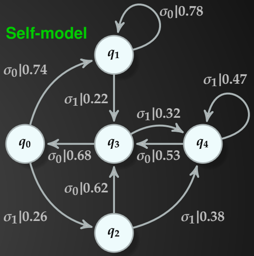
```
00000110000000001010000000000101000000000100110010000000000000000011011000000000011000
0000000000100111000000110000010010011011000000001110010000000110011001100000000000000
0000010101011100011111101100000110000001010000000000110110001100000000110000000000
11001101010011000000000000000010110111000000110001100000001011101100110011001100110
000011000000000000000000000001100000000011000000000011000000000011000000000110011001
0000011011110000000000110000000000000000000000000011000000101000000000000001010101010000
1110000000001100000000000000011000000000000110000000000011000000000000000000000000011000
00000010100000111100011000011000110000000000000000000000000000000011110001010101000
```

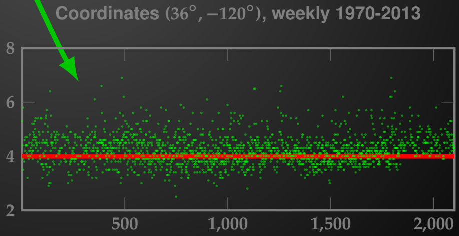
Stochastic process with 6 causal states



Predicting Seismic Events

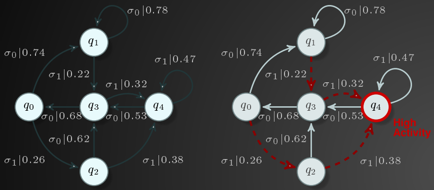
With Both Space & Time Quantization



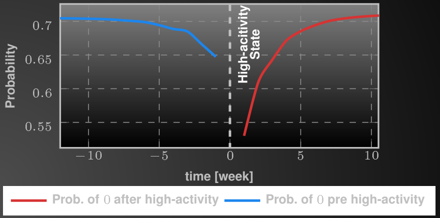


Validating The Omori-Utsu Law

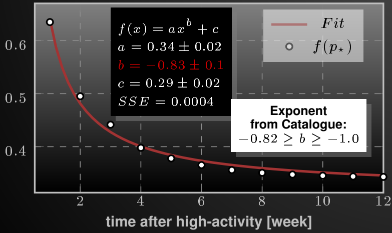
A. Inferred Self-Model for California
(400 mile radius around lat. 36° and long. -120°)



B Increasing probability of no-EQ before & after high-activity state



D. Indirect Validation of Omori-Utsu Law (aftershocks)



D. Indirect Validation of Omori-Utsu Law (foreshocks)