Metropolis Procedural Modeling

Johannes Agricola

December 4, 2012

- Control the model itself

- Control the model itself
- Grammar as model generator

- Control the model itself
- Grammar as model generator
- Any high level input specification

- Control the model itself
- Grammar as model generator
- Any high level input specification
- control via Markov chain Monte Carlo

1 Introduction

2 Parametric, Stochastic, Context-free Grammars

3 Markov Chain Monte Carlo

4 Reversible Jump Markov Chain Monte Carlo

5 RJMCMC applied to Grammars

6 Implementation

7 Results
- Architecture
- Trees
- Mondrian Art

- Remember: Grammars

$$G = (N, T, P, S)$$

  - N: Nonterminals
  - T: Terminals
  - P: Productions
  - S: Axiom

- Remember: Grammars

$$G = (N, T, P, S)$$

  - N: Nonterminals
  - T: Terminals
  - P: Productions
  - S: Axiom

- Example: Modeling branches: $N = \{X\}, T = \{F\}, S = X$

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Example

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Example Derivation:

$$\underline{X}$$

Example

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Example Derivation:

$$\underline{X} \rightarrow F[-\underline{X}][+X]$$

Example

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Example Derivation:

$$\begin{aligned} \underline{X} &\rightarrow F[-\underline{X}][+X] \\ &\rightarrow F[-F[-\underline{X}][+\underline{X}]][+\underline{X}] \end{aligned}$$

Example

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Example Derivation:

$$\begin{aligned} \underline{X} &\rightarrow F[-\underline{X}][+X] \\ &\rightarrow F[-F[-\underline{X}][+\underline{X}]][+\underline{X}] \\ &\rightarrow F[-F[-F][+F]][+F] \end{aligned}$$

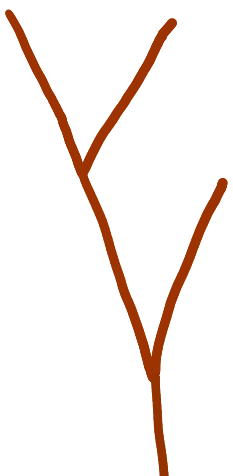
Example

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Example Derivation:

$$\begin{aligned} \underline{X} &\rightarrow F[-\underline{X}][+X] \\ &\rightarrow F[-F[-\underline{X}][+\underline{X}]][+\underline{X}] \\ &\rightarrow F[-F[-F][+F]][+F] \end{aligned}$$

Branch represented by this Example



$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Insufficiencies:

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Insufficiencies:

- No probability definitions for inference.

$$P = \begin{Bmatrix} X & \rightarrow & F[-X][+X] \\ X & \rightarrow & F \end{Bmatrix}$$

Insufficiencies:

- No probability definitions for inference.

$\rightarrow$ Specify inference probabilities:

$$P = \begin{Bmatrix} X & \overset{2/3}{\rightarrow} & F[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F \end{Bmatrix}$$

$$P = \left\{ \begin{array}{ccc} X & \overset{2/3}{\rightarrow} & F[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F \end{array} \right\}$$

Insufficiencies:

$$P = \left\{ \begin{array}{ccc} X & \overset{2/3}{\rightarrow} & F[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F \end{array} \right\}$$

Insufficiencies:

- No inherent fine structures

$$P = \left\{ \begin{matrix} X & \overset{2/3}{\rightarrow} & F[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F \end{matrix} \right\}$$

Insufficiencies:

- No inherent fine structures

$\rightarrow$ Add stochastic parameters:

$$P = \left\{ \begin{matrix} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{matrix} \right\}$$

Example

$$P = \left\{ \begin{array}{lll} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$\underline{X}$

Example

$$P = \left\{ \begin{array}{lll} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\underline{X} \rightarrow F(.31)[-\underline{X}][+X]$$

Example

$$P = \left\{ \begin{array}{lll} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\begin{aligned} \underline{X} &\rightarrow F(.31)[-\underline{X}][+X] \\ &\rightarrow F(.31)[-F(.55)[-\underline{X}][+\underline{X}]][+\underline{X}] \end{aligned}$$

Example

$$P = \left\{ \begin{array}{lll} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

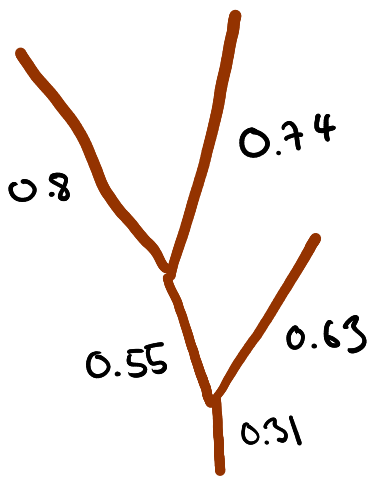
$$\begin{aligned} \underline{X} &\rightarrow F(.31)[-\underline{X}][+X] \\ &\rightarrow F(.31)[-F(.55)[-\underline{X}][+\underline{X}]][+\underline{X}] \\ &\rightarrow F(.31)[-F(.55)[-F(.8)][+F(.74)]][+F(.63)] \end{aligned}$$

Example

$$P = \left\{ \begin{array}{lll} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\begin{aligned} \underline{X} &\rightarrow F(.31)[-\underline{X}][+X] \\ &\rightarrow F(.31)[-F(.55)[-\underline{X}][+\underline{X}]][+\underline{X}] \\ &\rightarrow F(.31)[-F(.55)[-F(.8)][+F(.74)]][+F(.63)] \end{aligned}$$



$$P = \left\{ \begin{array}{lcc} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Insufficiencies:

$$P = \left\{ \begin{array}{lcc} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Insufficiencies:

- No inherent symmetry

$$P = \left\{ \begin{matrix} X & \overset{2/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X][+X] \\ X & \overset{1/3}{\rightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{matrix} \right\}$$

Insufficiencies:

- No inherent symmetry

$\rightarrow$ Add algebraic parameters:

$$P = \left\{ \begin{matrix} X(l) & \overset{l/3}{\longrightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X(l-1)][+X(l-1)] \\ X(l) & \overset{1-l/3}{\longrightarrow} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{matrix} \right\}$$

with $S = X(3)$.

Example

$$P = \left\{ \begin{array}{lll} X(l) & \xrightarrow{l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X(l-1)][+X(l-1)] \\ X(l) & \xrightarrow{1-l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\underline{X(3)}$$

Example

$$P = \left\{ \begin{array}{lll} X(l) & \xrightarrow{l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X(l-1)][+X(l-1)] \\ X(l) & \xrightarrow{1-l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\underline{X(3)} \rightarrow F(.31)[-\underline{X(2)}][+X(2)]$$

Example

$$P = \left\{ \begin{array}{lll} X(l) & \xrightarrow{l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X(l-1)][+X(l-1)] \\ X(l) & \xrightarrow{1-l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\begin{aligned} \underline{X(3)} &\rightarrow F(.31)[-\underline{X(2)}][+X(2)] \\ &\rightarrow F(.31)[-F(.55)[-\underline{X(1)}][+\underline{X(1)}]][+\underline{X(2)}] \end{aligned}$$

Example

$$P = \left\{ \begin{array}{lll} X(l) & \xrightarrow{l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X(l-1)][+X(l-1)] \\ X(l) & \xrightarrow{1-l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Example Derivation:

$$\begin{aligned} \underline{X(3)} &\rightarrow F(.31)[-\underline{X(2)}][+X(2)] \\ &\rightarrow F(.31)[-F(.55)[-\underline{X(1)}][+\underline{X(1)}]][+\underline{X(2)}] \\ &\rightarrow F(.31)[-F(.55)[-F(.8)][+F(.74)]][+F(.63)] \end{aligned}$$

Definition

- $\Delta(G)$: space of all possible derivations from G.
- $\delta = (\tau, \varphi) \in \Delta(G)$ consists of
  - derivation tree τ
  - list of descriptive parameters φ

Example

$$P = \left\{ \begin{array}{lll} X(l) & \xrightarrow{l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right))[-X(l-1)][+X(l-1)] \\ X(l) & \xrightarrow{1-l/3} & F(\mathcal{N}\left(\mu, \sigma^2\right)) \end{array} \right\}$$

Revisiting the last example gives us:

$$\begin{aligned} \delta &= (\tau, \varphi) \\ \tau &= (\underline{X(3)} \\ &\quad \rightarrow F(.31)[-\underline{X(2)}][+X(2)] \\ &\quad \rightarrow F(.31)[-F(.55)[-\underline{X(1)}][+\underline{X(1)}]][+\underline{X(2)}] \\ &\quad \rightarrow F(.31)[-F(.55)[-F(.8)][+F(.74)]][+F(.63)]) \\ \varphi &= (0.31, 0.55, 0.8, 0.74, 0.63) \end{aligned}$$

Definition

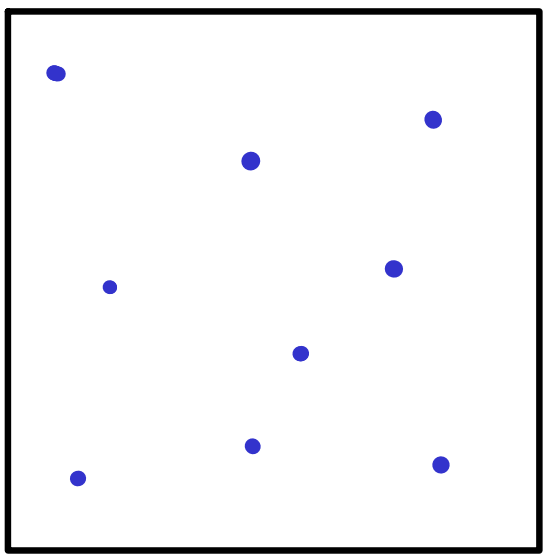
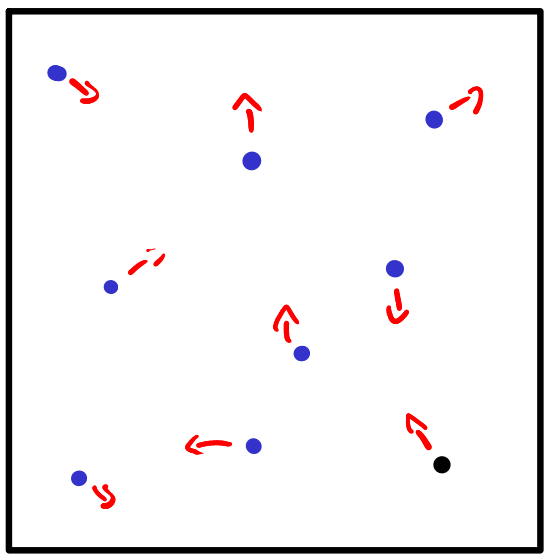
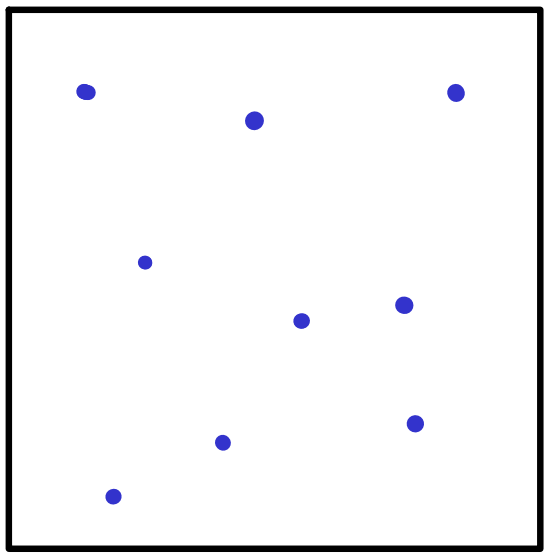
Use as posterior:

$$p\left(\delta|I\right) \propto L\left(I|\delta\right)\pi\left(\delta\right)$$

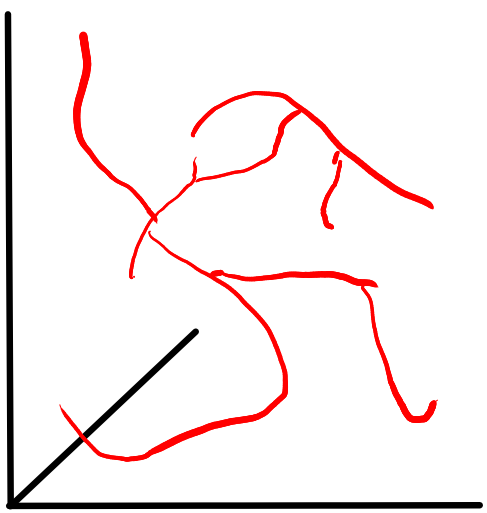
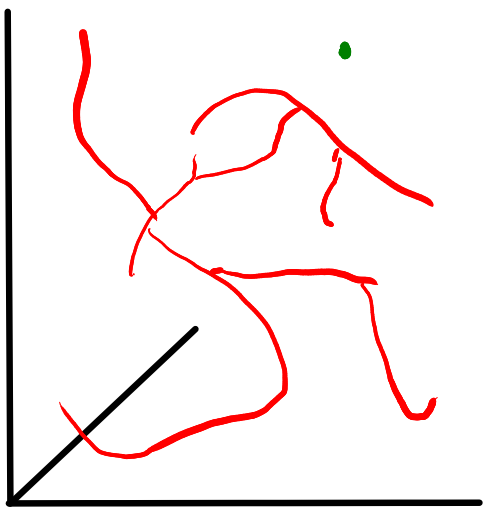
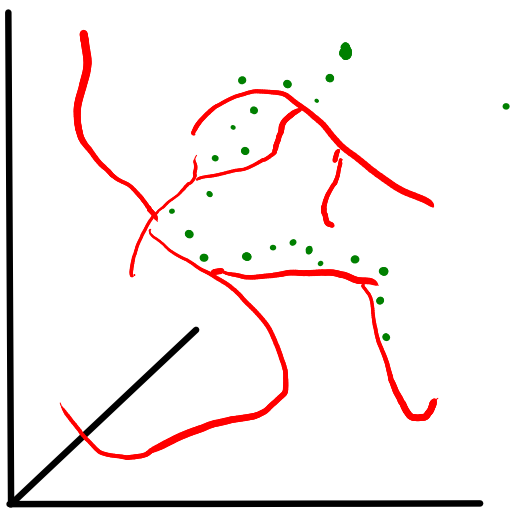
- $L\left(I|\delta\right)$: likelihood of input I given model δ
- model prior $\pi\left(\delta\right) \propto \prod_{s\in\delta} P\left(s|\mathrm{parent}(s)\right) \prod_{t\in\delta_{\Rightarrow}} \Phi_t(\varphi_t)$

Finding the best inference δ given $I \Leftrightarrow \max p(\delta|I)$.

- Initially proposed by Metropolis et al. 1953: Equations of State [3]

- Generlized by Hastings 1970: Markov chain Monte Carlo sampling methods [2]
  - Instead of sampling from distribution p
  - random walk along p's high probability areas by:
  - generating a markov chain with stationary distribution p.

- Markov chain simulation:
  - Take Markov chain with transition matrix Q
  - Sample next tentative state x^* after x_i from Q
  - accept it with probability

$$\alpha_{x_i \to x^*} = \min \left\{ 1, \frac{p\left(x^*\right) q\left(x_i | x^*\right)}{p\left(x_i\right) q\left(x^* | x_i\right)} \right\}$$

- Original formulation does not include varying dimension
- Number of parameters after grammar inference not fixed:

$$\begin{aligned} & \underline{X(3)} \\ \rightarrow\ & F(.31)[-\underline{X(2)}][+X(2)] \\ \rightarrow\ & F(.31)[-F(.55)[-\underline{X(1)}][+\underline{X(1)}]][+\underline{X(2)}] \\ \rightarrow\ & F(.31)[-F(.55)[-F(.8)][+F(.74)]][+F(.63)] \end{aligned}$$

$$\begin{aligned} & \underline{X(3)} \\ \rightarrow\ & F(.31)[-F(.55)][+F(.8)] \end{aligned}$$

- Extension by Green 1995: Reversible Jump Markov chain Monte Carlo [1]
- Adds **jump** moves:
  - changing dimension n to m
  - **dimension matching function** $f_{n \to m}$
  - has to be **reversible**
- Parametrisation:
  - Select jump with probability $j\,(m|n)$
  - Sample state after jump from $q_{m \to n}$.

Acceptance probability for diffusion:

$$\alpha_{x_i \to x^*} = \min \left\{ 1, \frac{p\left(x^*\right) q\left(x_i | x^*\right)}{p\left(x_i\right) q\left(x^* | x_i\right)} \right\}$$

Acceptance probability for jump:

$$\alpha_{n \to m} = \min \left\{ 1, \frac{p\left(m, x_m^*\right)}{p\left(n, x_n\right)} \cdot \frac{q_{m \to n}\left(u_{m,n} | m, x_m^*\right)}{q_{n \to m}\left(u_{n,m} | m, x_n\right)} \cdot \frac{j\left(n|m\right)}{j\left(m|n\right)} \cdot \det \left| \frac{\partial f_{n \to m}\left(x_m | u_{m,n}\right)}{\partial (x_m, u_{m,n})} \right| \right\}$$

with

- $j\left(\cdot|\cdot\right)$: probability of selecting the respective jump
- $q_{m \to n}$: (jump) proposal density function
- $\det \left| \frac{\partial f_{n \to m}(x_m | u_{m,n})}{\partial (x_m, u_{m,n})} \right|$: Determinant of the jacobian of the dimension matching function.

1 Introduction

2 Parametric, Stochastic, Context-free Grammars

3 Markov Chain Monte Carlo

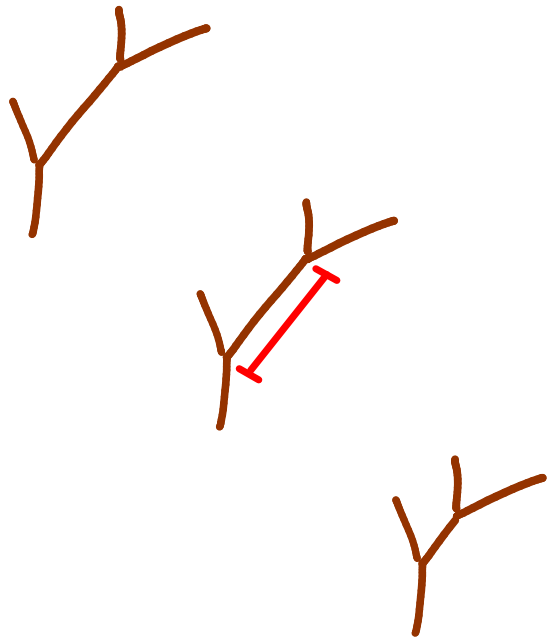
4 Reversible Jump Markov Chain Monte Carlo

5 RJMCMC applied to Grammars

6 Implementation

7 Results
- Architecture
- Trees
- Mondrian Art

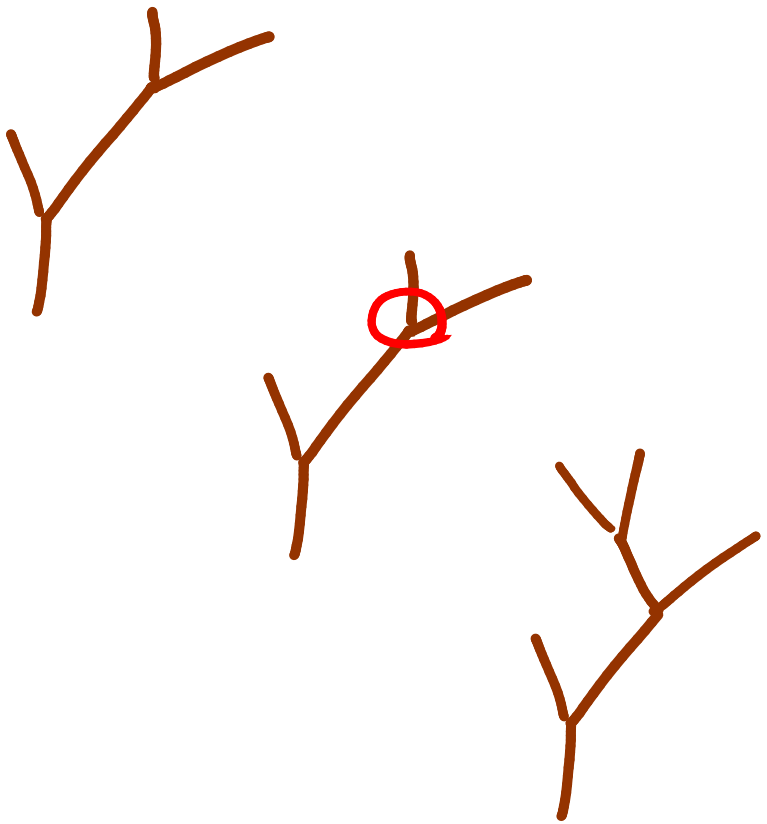
- Talton et al. 2005: Metropolis Procedural Modeling [4]
- remember: $\delta = (\tau, \varphi) \in \Delta(G)$
- Diffusion move structure:
  - pick random terminal t.
  - sample $\varphi_t' \sim \Phi_t$ to generate δ'
  - $q(\delta'|\delta) = q(\delta) = \Phi_t(\varphi_t')$
  - accept with probability



$$\begin{aligned}
\alpha_{\delta \to \delta'} &= \min\left\{1, \frac{p(\delta'|I)}{p(\delta|I)} \frac{q(\delta|\delta')}{q(\delta'|\delta)}\right\} \\
&= \min\left\{1, \frac{p(\delta'|I)}{p(\delta|I)} \frac{\Phi_t(\varphi_t)}{\Phi_t(\varphi_t')}\right\} \\
&= \min\left\{1, \frac{L(I|\delta')}{L(I|\delta)}\right\}
\end{aligned}$$

$$\because p(\delta|I) \propto L(I|\delta)\,\pi(\delta)$$

Jump Moves



Dimension Matching:

- Each $\varphi_t \sim \Phi_t$ has corresponding $\widetilde{\varphi}_t \in [0,1]^k$, $k \in \mathbb{N}$
- all components of $\widetilde{\varphi}_t$ are $\sim U_{[0,1]}$!
- Therefore: Model $\varphi \in \mathbb{R}^q$ as $\widetilde{\varphi} \in [0,1]^w$, $q \leq w$.

Dimension Matching:

- Each $\varphi_t \sim \Phi_t$ has corresponding $\widetilde{\varphi}_t \in [0,1]^k$, $k \in \mathbb{N}$
- all components of $\widetilde{\varphi}_t$ are $\sim U_{[0,1]}$!
- Therefore: Model $\varphi \in \mathbb{R}^q$ as $\widetilde{\varphi} \in [0,1]^w$, $q \leq w$.
- Now, $f_{n \to m}$ for increasing dimension:

$$\varphi' = f_{n \to m}(\varphi) = (\varphi, u)$$

  where $u \in [0,1]^k$ with appropriately chosen k.
- Determinant of Jacobian:

$$\det \left| \frac{\partial f_{n \to m}(x_m | u_{m,n})}{\partial (x_m, u_{m,n})} \right| = \det \left| \frac{\partial (\varphi, u)}{\partial \varphi'} \right| = 1$$

Dimension Matching:

- Each $\varphi_t \sim \Phi_t$ has corresponding $\widetilde{\varphi}_t \in [0,1]^k$, $k \in \mathbb{N}$
- all components of $\widetilde{\varphi}_t$ are $\sim U_{[0,1]}$!
- Therefore: Model $\varphi \in \mathbb{R}^q$ as $\widetilde{\varphi} \in [0,1]^w$, $q \leq w$.
- Now, $f_{n \to m}$ for increasing dimension:

$$\varphi' = f_{n \to m}(\varphi) = (\varphi, u)$$

  where $u \in [0,1]^k$ with appropriately chosen k.
- Determinant of Jacobian:

$$\det \left| \frac{\partial f_{n \to m}(x_m | u_{m,n})}{\partial (x_m, u_{m,n})} \right| = \det \left| \frac{\partial (\varphi, u)}{\partial \varphi'} \right| = 1$$

- dual for dimension reduction
- still no semantic comparison

- select a non-terminal $v \in \tau$ of δ with probability $q_\tau(v)$.
- rederive grammar starting at v to get τ'.
- apply dimension matching funtcion: $\varphi' = f_{n \to m}(\varphi)$.
- short calculation gives:

$$\alpha_{\delta \to \delta'} = \min \left\{ 1, \frac{q_{\tau'}(v)}{q_\tau(v)} \frac{L(I|\delta')}{L(I|\delta)} \frac{\prod_{t \in \tau'_v} \Phi_t(\varphi'_t)}{\prod_{t \in \tau_v} \Phi_t(\varphi_t)} \right\}$$

Now the algorithm is surprisingly simple:

- **Given:** Grammar G, Input specification I, Likelihood function L, computational budget.
- Derive one $\delta \in \Delta(G)$
- While still within budget:
  - Select randomly between jump and diffusion move
  - Perform move
  - Compute acceptance probability
  - Accept or reject move

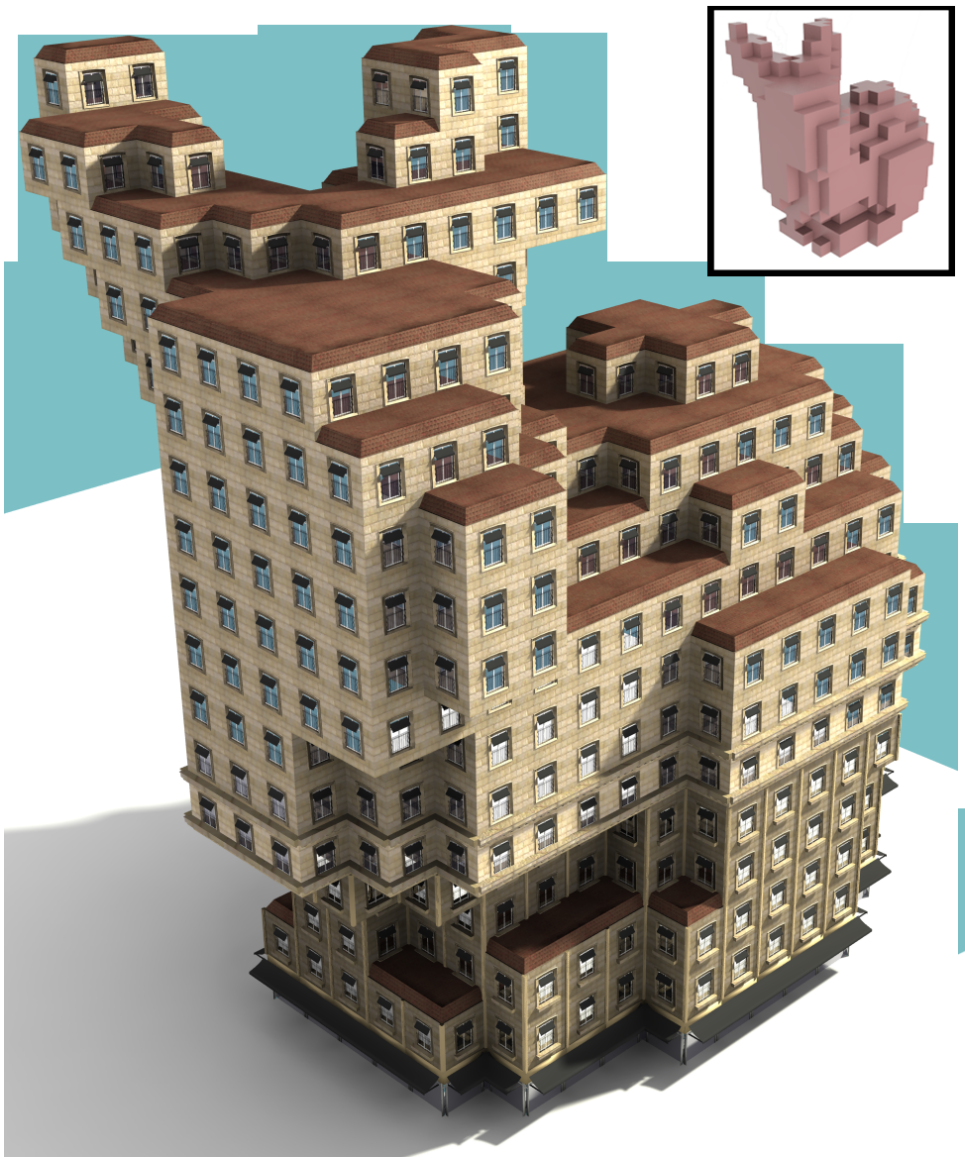
- implemented likelihood functions:
  - Image / Volume based
  - Analytical
- supplied grammars:
  - Architecture: Houses and Cities
  - Trees
  - Mondrian Paintings

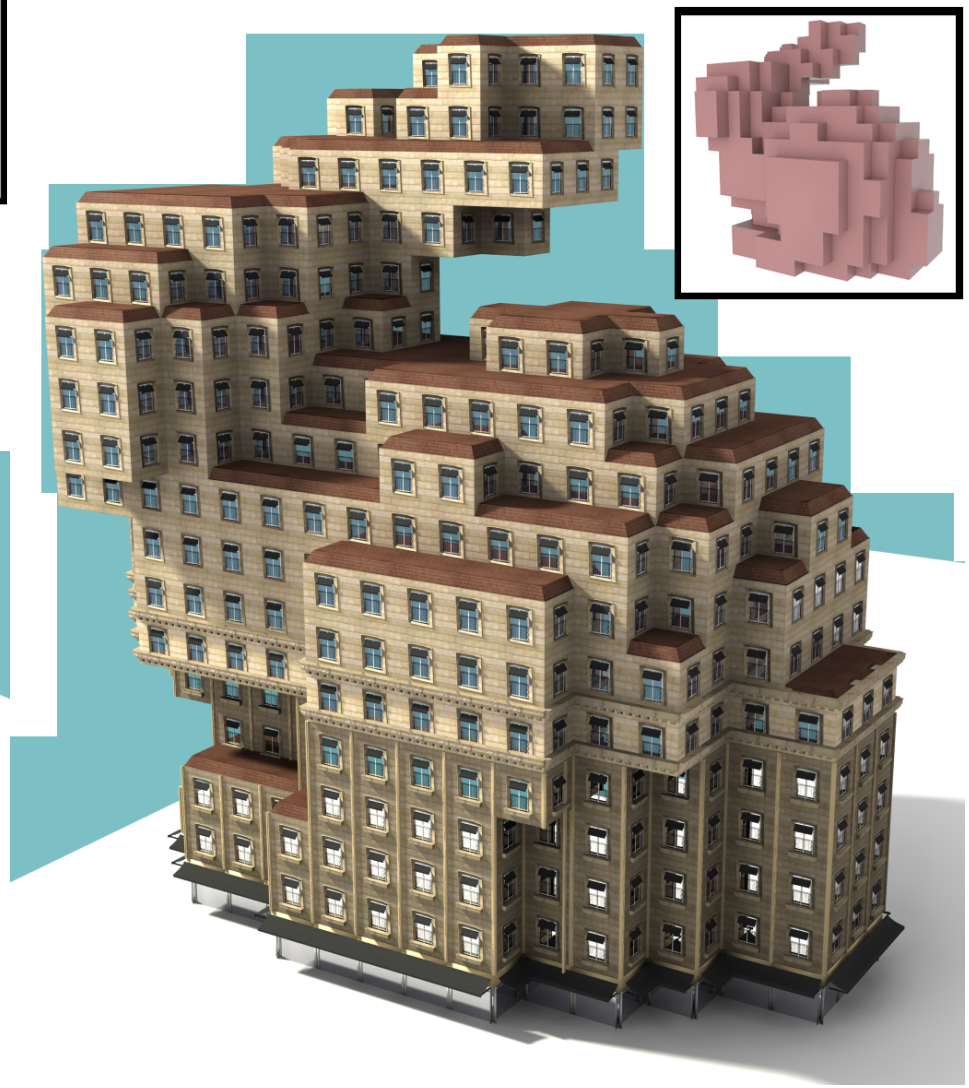


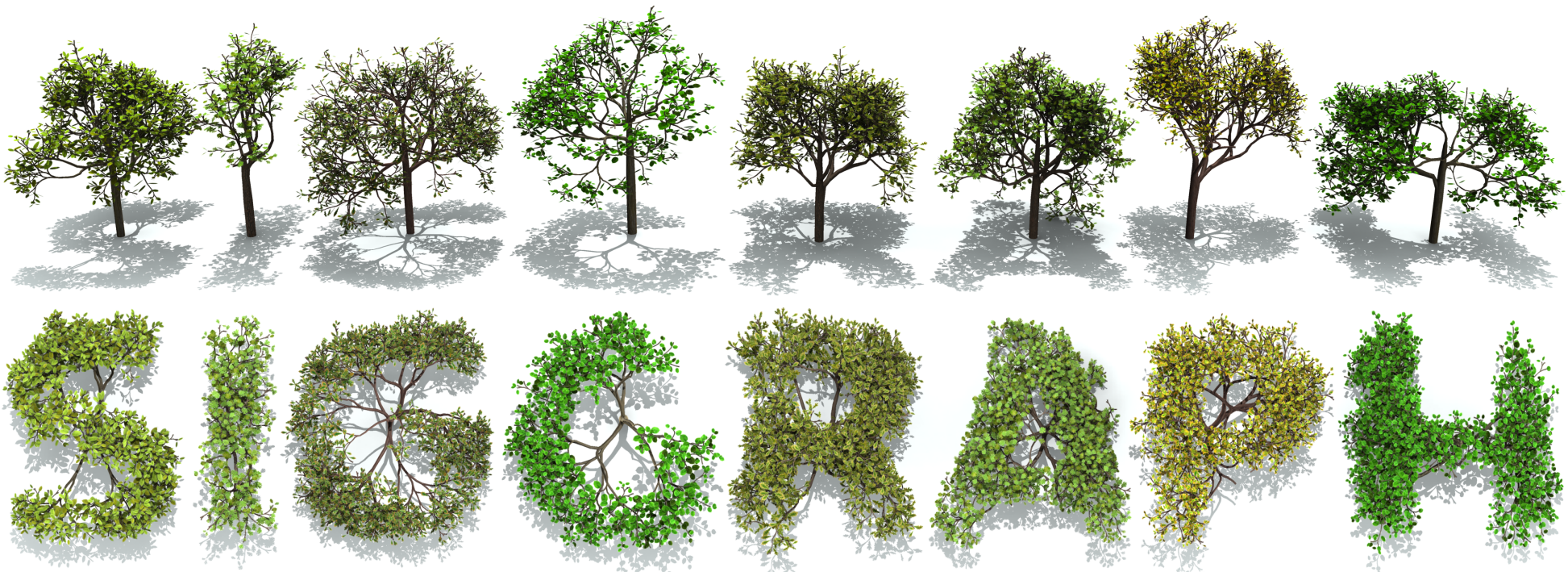


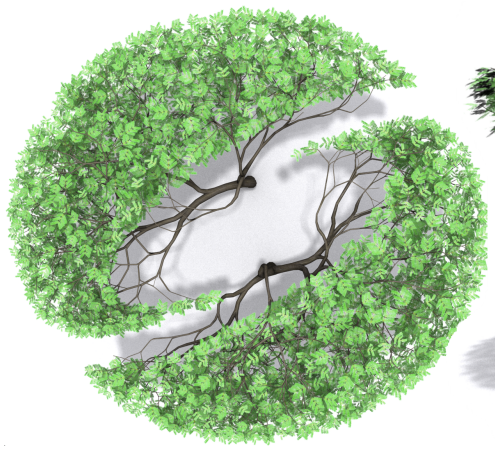










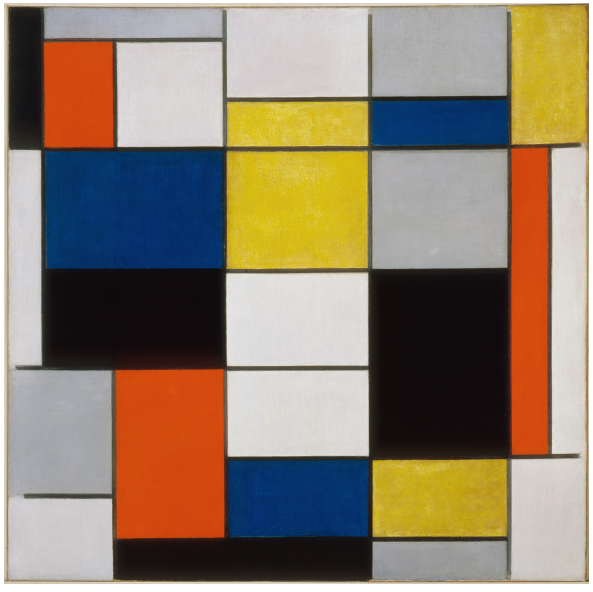


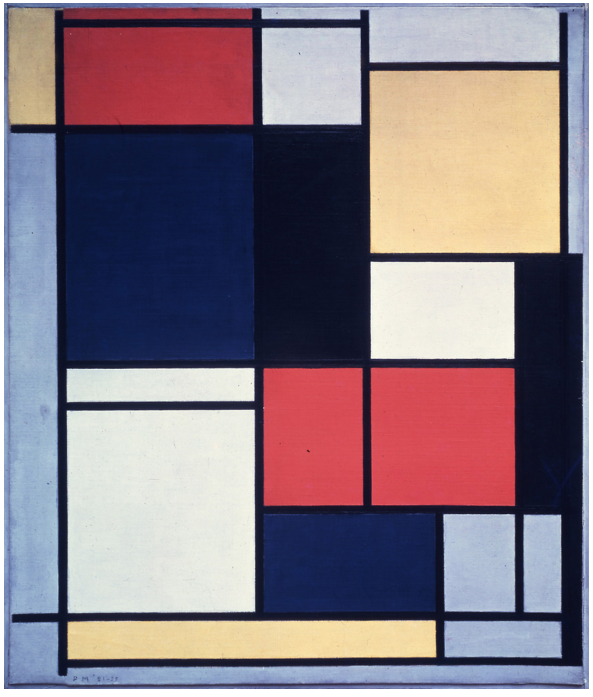


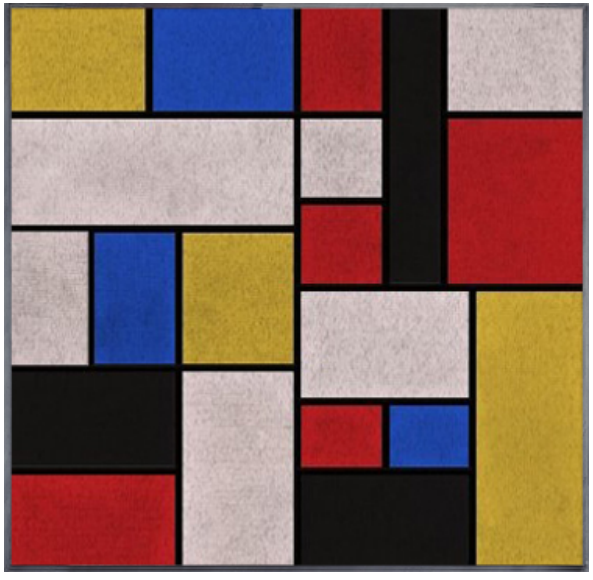

Insufficient Grammar Support

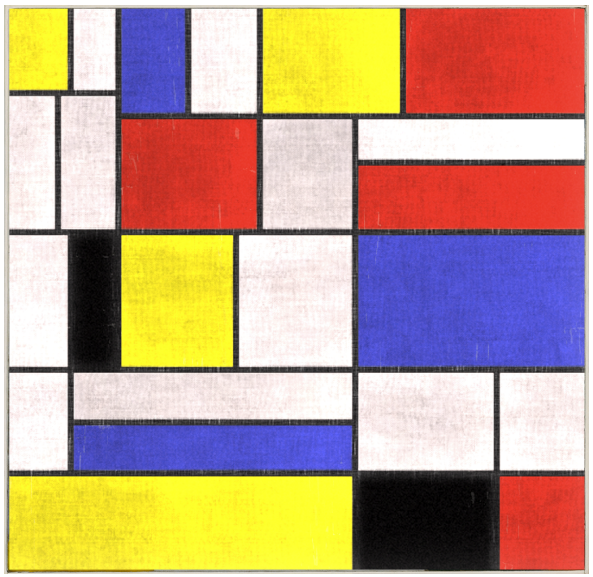
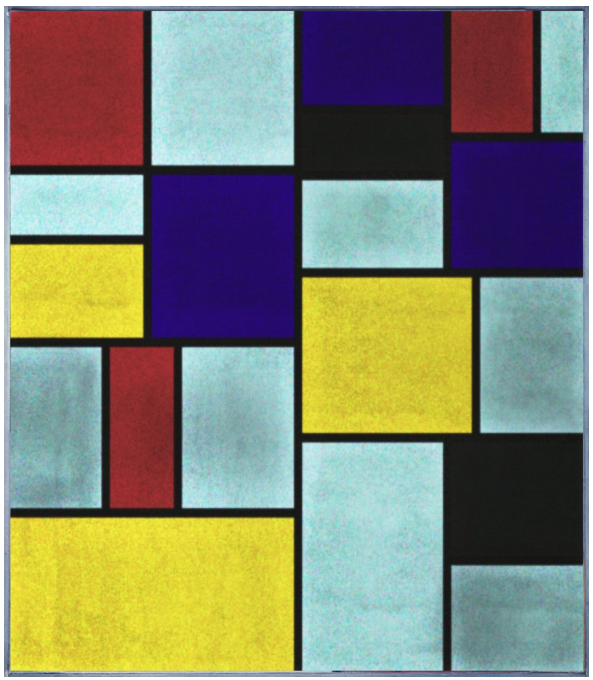
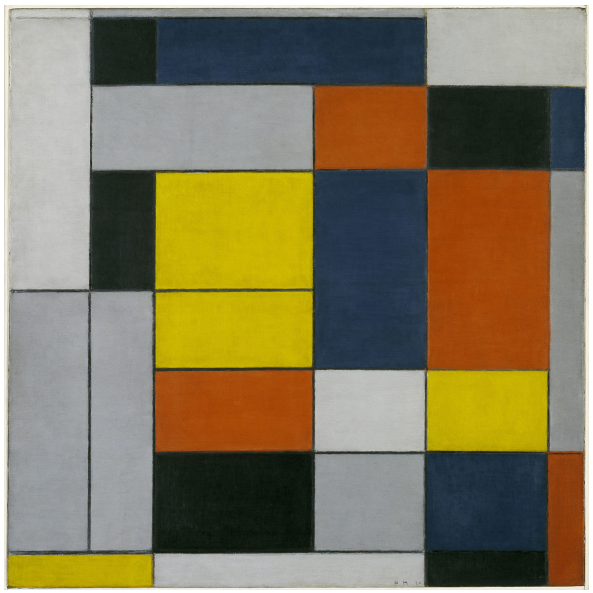










Fin.


Peter J. Green.
Reversible jump markov chain monte carlo computation and bayesian model determination.
Biometrika, 82:711–732, 1995.


W. K. Hastings.
Monte carlo sampling methods using markov chains and their applications.
Biometrika, 57(1):97–109, 1970.


N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller.
Equation of State Calculations by Fast Computing Machines.
J. Chem. Phys., 21:1087–1092, June 1953.



Jerry O. Talton, Yu Lou, Steve Lesser, Jared Duke, Radomír Měch, and Vladlen Koltun.
Metropolis procedural modeling.
ACM Trans. Graph., 30(2):11:1–11:14, April 2011.

$$\alpha_{n\to m} = \min\left\{1, \frac{p\left(m, x_m^*\right)}{p\left(n, x_n\right)} \cdot \frac{q_{m\to n}\left(u_{m,n}|m, x_m^*\right)}{q_{n\to m}\left(u_{n,m}|m, x_n\right)} \cdot \frac{j\left(n|m\right)}{j\left(m|n\right)} \cdot \mathcal{J}_{n\to m}\right\}$$

$$\alpha_{\delta\to\delta'} = \min\left\{1, \frac{p(\delta'|I)}{p(\delta|I)} \frac{j(\tau_v|\tau_v')}{j(\tau_v'|\tau_v)} \frac{U_{[0,1]}(u')}{1} \mathcal{J}_{\delta\to\delta'}\right\}$$

$$p\left(\delta|I\right) \propto L\left(I|\delta\right)\pi\left(\delta\right),\ \pi\left(\delta\right) \propto \prod_{s\in\delta} P\left(s|\text{parent}(s)\right) \prod_{t\in\delta_\Rightarrow} \Phi_t(\varphi_t)$$

$$= \min\left\{1, \frac{L\left(I|\delta'\right)}{L\left(I|\delta\right)} \frac{\pi(\delta')}{\pi(\delta)} \frac{q_{\tau'}(v)}{q_\tau(v)} \frac{\prod_{s\in\tau_v} P(s|\text{parent}(s))}{\prod_{s\in\tau_{v'}} P(s|\text{parent}(s))}\right\}$$

$$= \min\left\{1, \frac{q_{\tau'}(v)}{q_\tau(v)} \frac{L\left(I|\delta'\right)}{L\left(I|\delta\right)} \frac{\prod_{t\in\tau_v'} \Phi_t\left(\varphi_t'\right)}{\prod_{t\in\tau_v} \Phi_t\left(\varphi_t\right)}\right\}$$