

# Deep Generative Models

Hongyue Li

# Sampling

# Sample a random number from uniform $[0,1]$ ?

**Pseudo-random number generator (PRNG)**, which is an algorithm that produces a sequence of numbers that *appears random*.

Linear Congruential Generator (LCG)

Parameters:

$$X_0 \text{ (seed)}, \quad a = 22,695,477, \quad c = 1, \quad m = 2^{31}$$

Step 1 — Generate next integer:

$$X_{n+1} = (aX_n + c) \bmod m$$

Step 2 — Optional normalization:

$$U_{n+1} = \frac{X_{n+1}}{m} \in [0, 1)$$

Step 3 — Repeat:

Use  $X_{n+1}$  as the new seed and iterate as many times as needed.

Sample a random number from Normal distribution  $N(0,1)$ ?

start from uniform distribution, then **transform** it to normal distribution

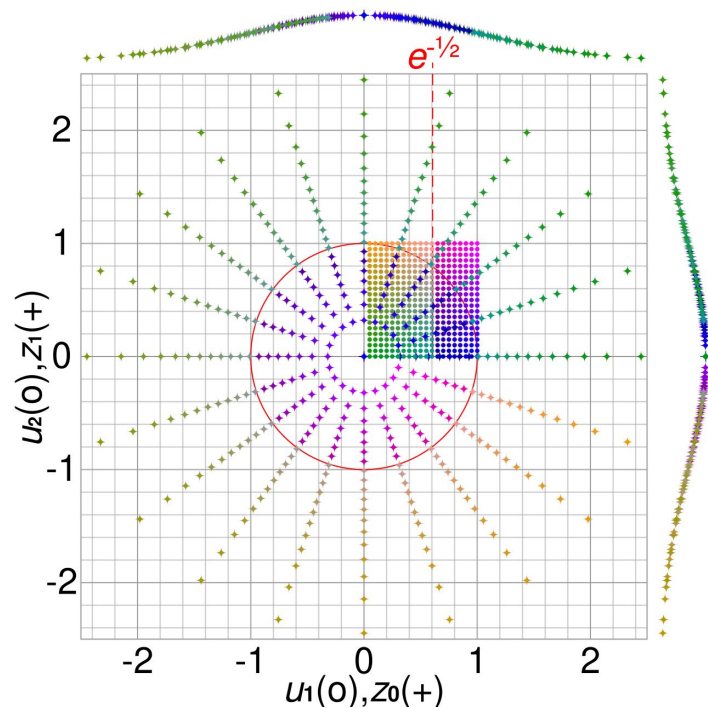
## Box-Muller Transform

$$U_1, U_2 \sim \text{Uniform}(0, 1)$$

$$Z_0 = \sqrt{-2 \ln U_1} \cos(2\pi U_2),$$

$$Z_1 = \sqrt{-2 \ln U_1} \sin(2\pi U_2)$$

$$Z_0, Z_1 \sim \mathcal{N}(0, 1), \quad \text{independent}$$



# Sample a random state from Ising model?

For spins  $s_i \in \{+1, -1\}$  on graph  $G$  with couplings  $J_{ij}$  and external fields  $h_i$ ,

$$P(s) = \frac{1}{Z} \exp \left( \beta \sum_{\langle i,j \rangle} J_{ij} s_i s_j + \beta \sum_i h_i s_i \right),$$

where  $\beta = 1/(k_B T)$ . Energy  $E(s) = - \sum_{\langle i,j \rangle} J_{ij} s_i s_j - \sum_i h_i s_i$ .

## Gibbs Sampling Steps

1. Pick a spin  $i$  to update (randomly or in sequence).
2. Compute the local field acting on that spin:

$$H_i = \sum_j J_{ij} s_j + h_i.$$

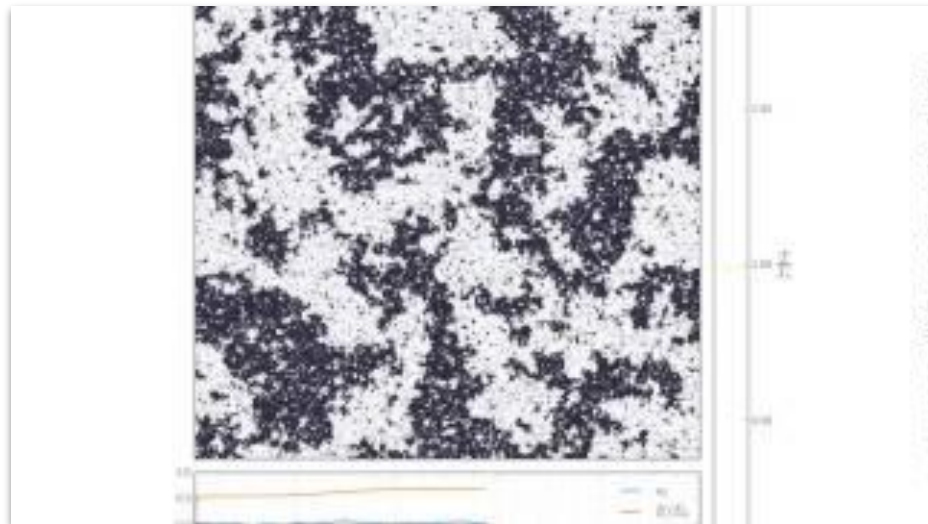
3. Compute the conditional probability that spin  $i$  is  $+1$ :

$$P(s_i = +1 \mid s_{\setminus i}) = \frac{e^{\beta H_i}}{e^{\beta H_i} + e^{-\beta H_i}} = \frac{1}{1 + e^{-2\beta H_i}}.$$

4. Sample  $s_i^{(t+1)}$  from this Bernoulli distribution:

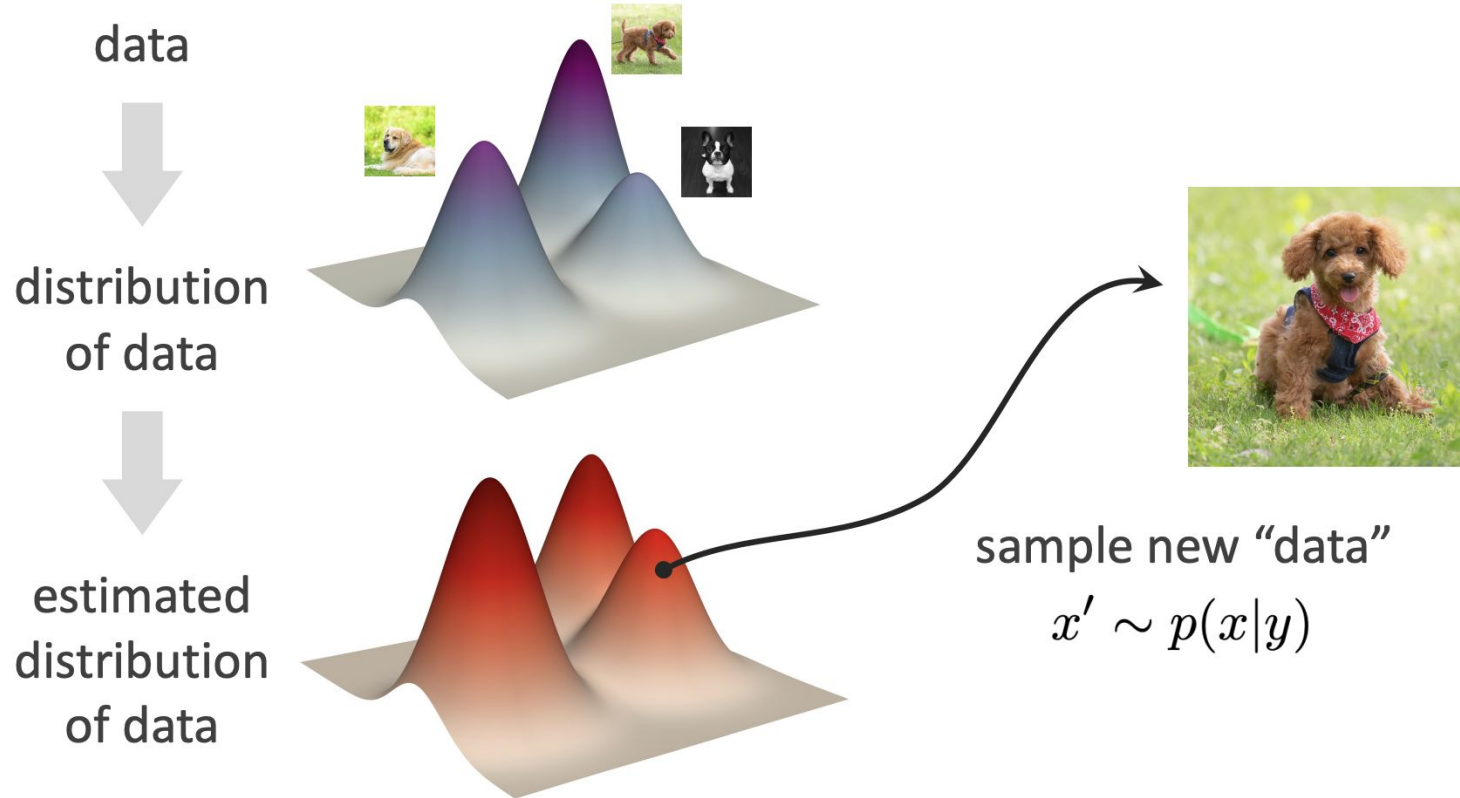
$$s_i^{(t+1)} = \begin{cases} +1 & \text{with probability } p_i = \frac{1}{1 + e^{-2\beta H_i}}, \\ -1 & \text{with probability } 1 - p_i. \end{cases}$$

5. Repeat steps 1–4 for all spins (one “sweep”) to get the next full configuration  $s^{(t+1)}$ .
6. Iterate many sweeps until samples approximate the stationary distribution  $P(s)$ .

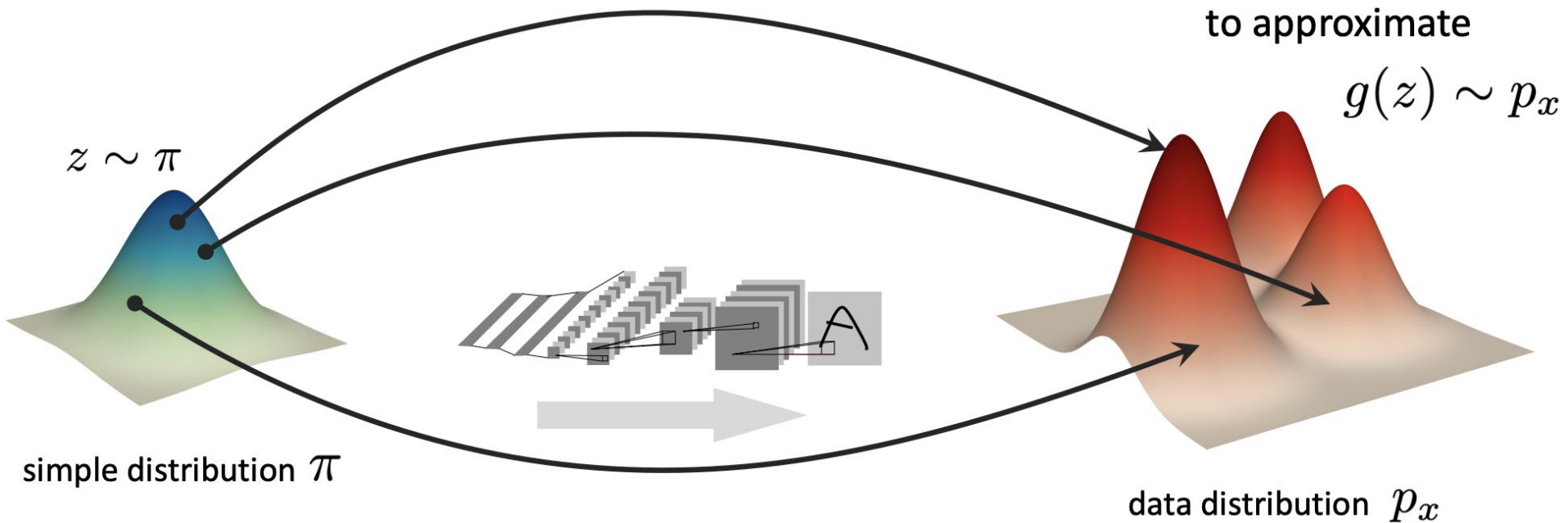


# Generative Modeling

# How to generate data sampled from some distribution?



Transform a simple distribution to  
a complex one that approximates the data distribution





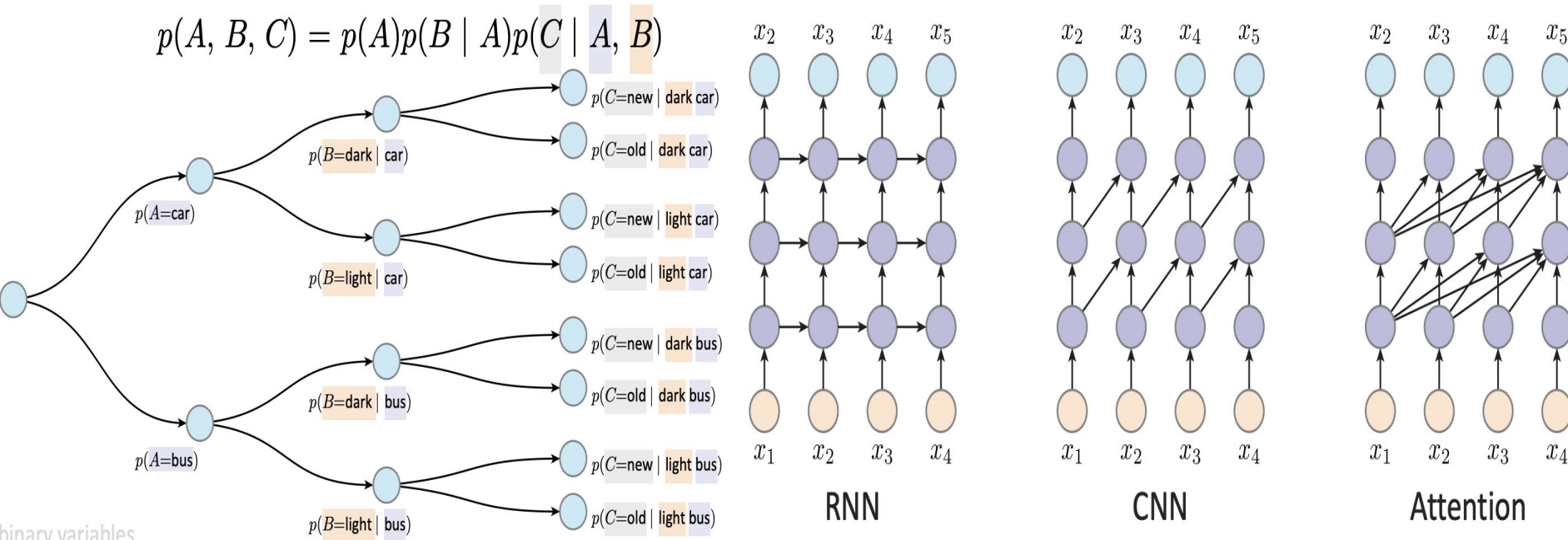
# Deep Generative Model

- Modeling & Learning
  - Formulation: frame the problem as probabilistic modeling
  - Representation: deep neural networks to represent data distribution
  - Objective: to measure how good the predicted distribution is
  - Symmetry: decompose complex distributions into simple and tractable ones
- Inference:
  - sampler: to produce new samples
  - probability density estimator (optional)

# Probabilistic Graphical Model & Autoregressive Model

Formulation: frame the problem as probabilistic modeling

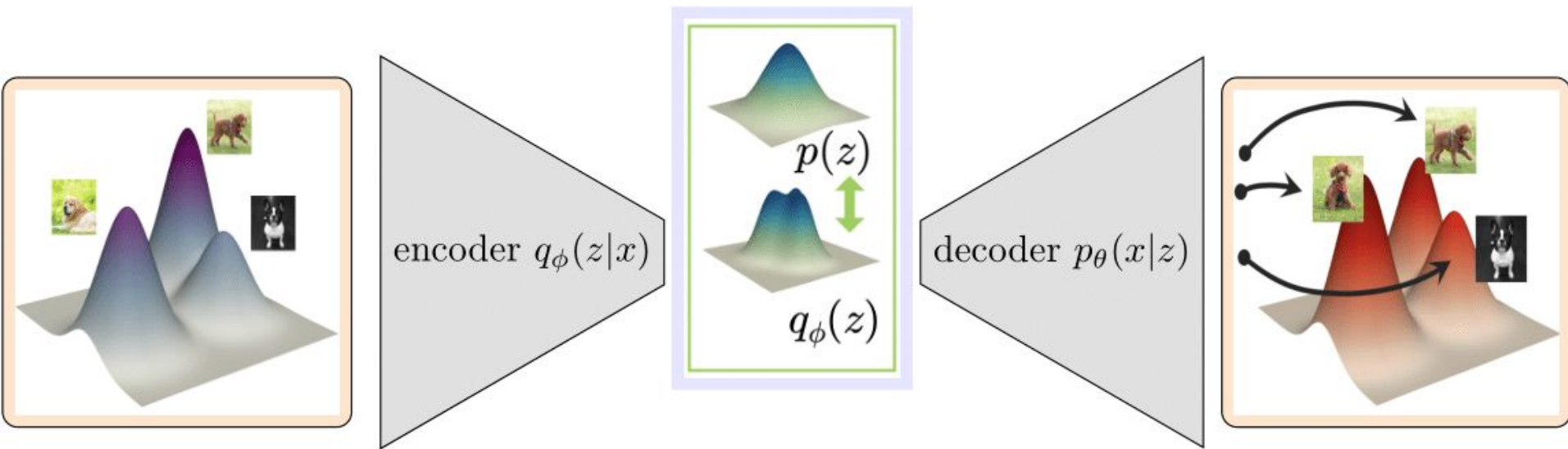
$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2 \mid x_1)...p(x_n \mid x_1, x_2, \dots, x_{n-1})$$



# Variational Autoencoder

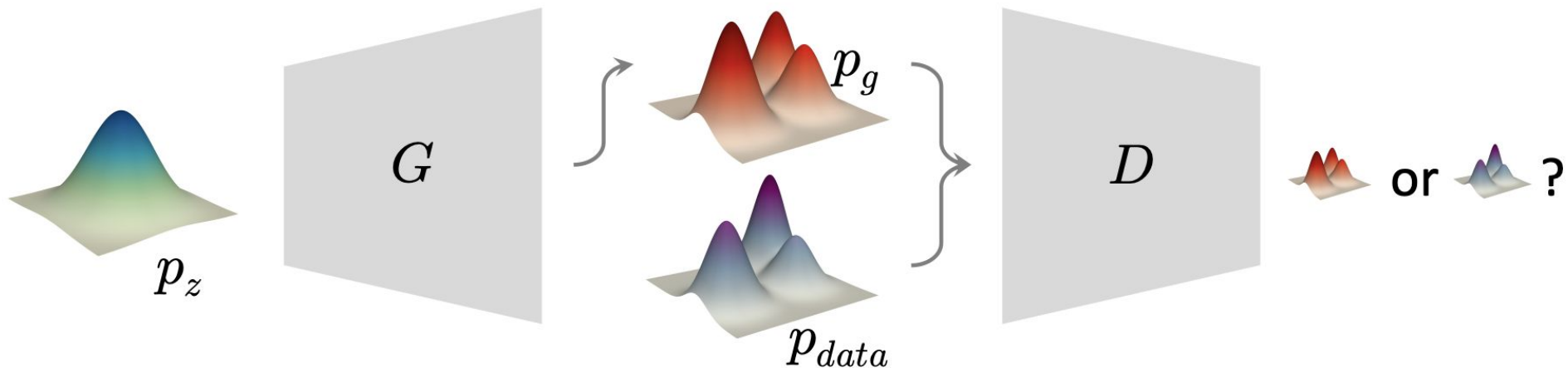
Representation: deep neural networks to represent data distribution

$$\mathcal{L}_{\theta, \phi}(x) = -\mathbb{E}_{z \sim q_{\phi}(z|x)} \left[ \log p_{\theta}(x|z) \right] + \mathcal{D}_{\text{KL}} \left( q_{\phi}(z|x) || p(z) \right)$$



# Generative Adversarial Networks

Objective: to measure how good the predicted distribution is

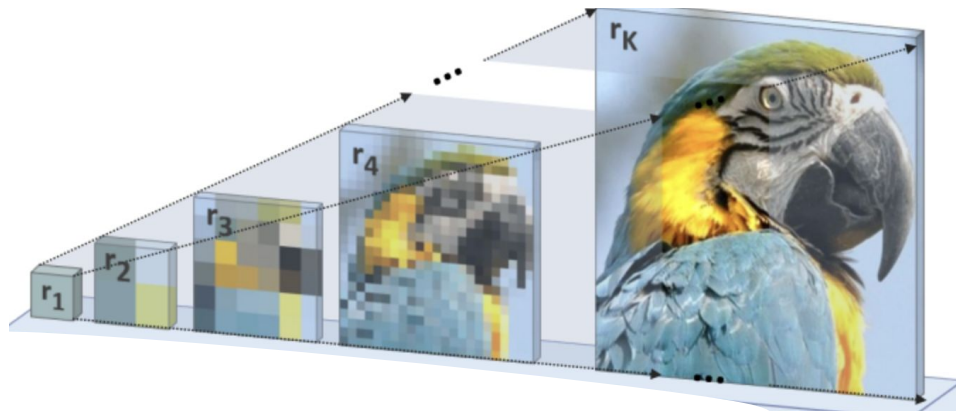
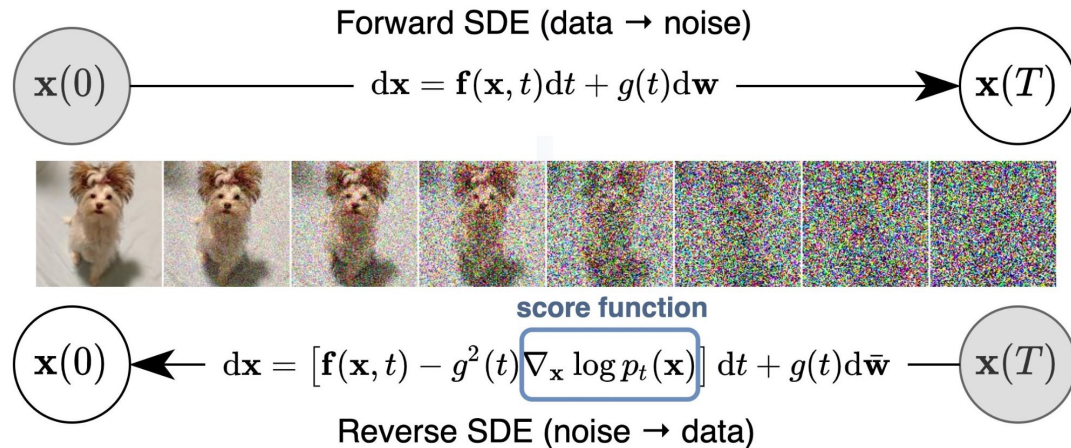


$$\text{JS}(p, q) = \frac{1}{2} \text{KL}(p \| m) + \frac{1}{2} \text{KL}(q \| m) = H(m) - \frac{1}{2} H(p) - \frac{1}{2} H(q), \quad m = \frac{p+q}{2}$$

$$W_1(p, q) = \left( \inf_{\gamma \in \Gamma(p, q)} \int d(x, y)^k d\gamma(x, y) \right)^{1/k} \Big|_{k=1} = \sup_{\|f\|_{\text{Lip}} \leq 1} \left\{ \mathbb{E}_p[f] - \mathbb{E}_q[f] \right\}$$

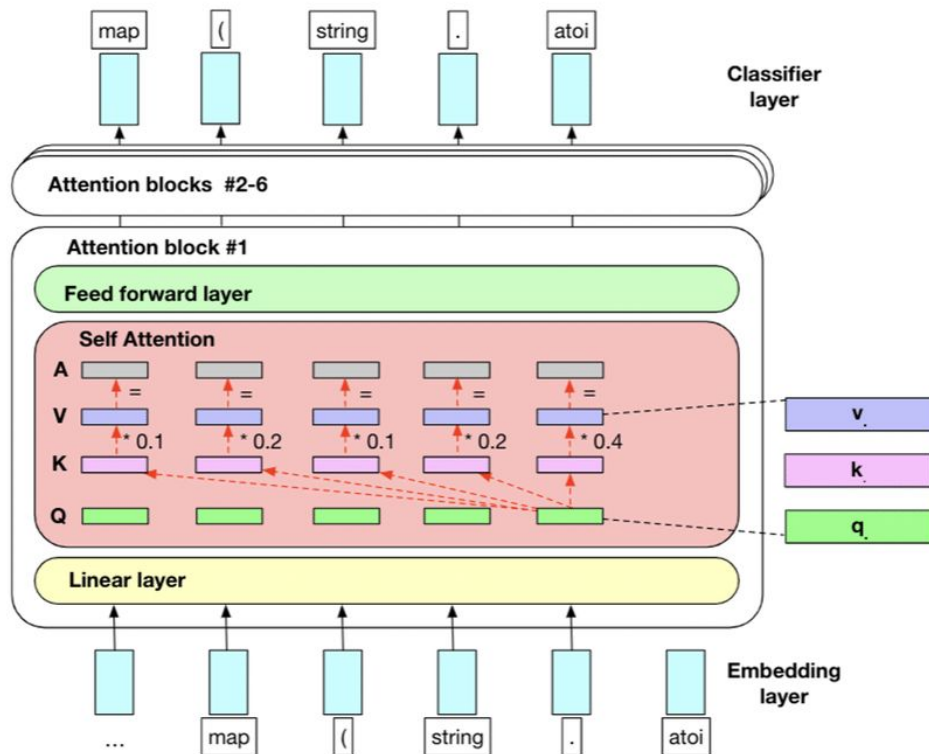
# Diffusion Model and beyond

Symmetry: decompose complex distributions into simple and tractable ones



# Applications

# Language Generation



+ New chat

BE

Brief me about Elon Musk within 50 words

History is temporarily unavailable.  
We're working to restore this feature  
as soon as possible.



Elon Musk is a billionaire entrepreneur and the CEO of Tesla, SpaceX, Neuralink, and The Boring Company. He is known for his ambitious goals of colonizing Mars, revolutionizing the electric car industry, and connecting human brains to computers.



Upgrade to Plus

NEW

Dark mode

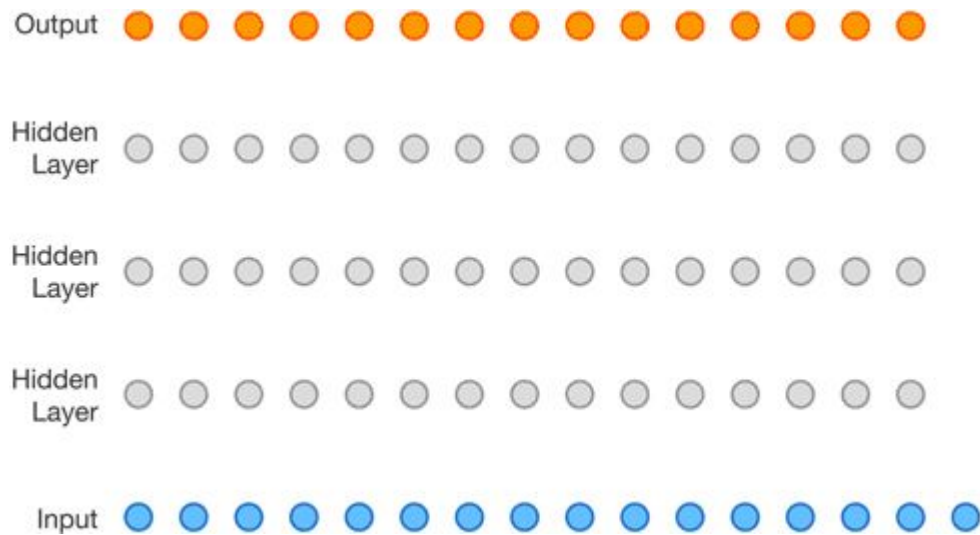
Updates & FAQ

Log out

Regenerate response

ChatGPT Mar 14 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

# Audio Generation





# Image Generation



2014



2015



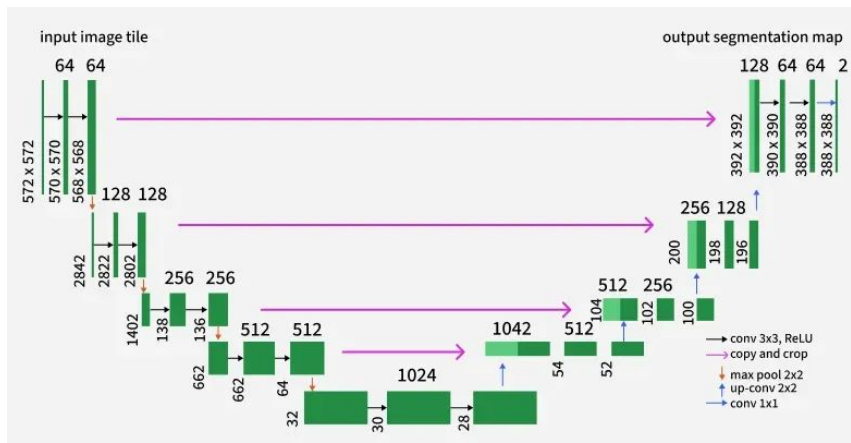
2016



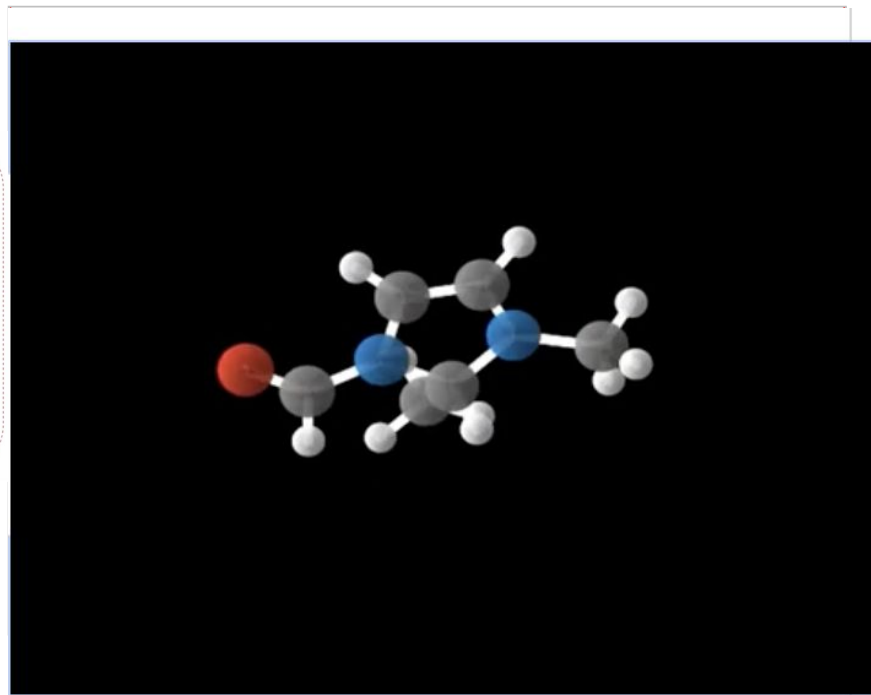
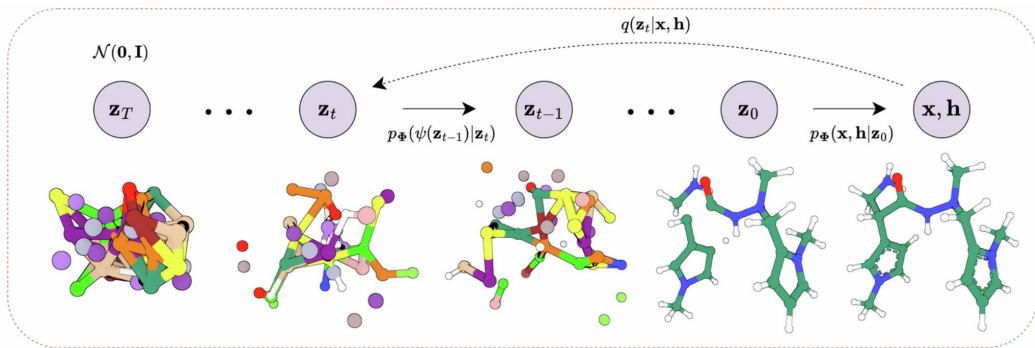
2017



2018

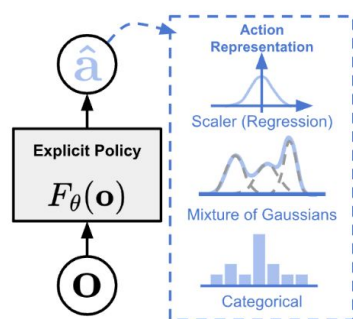


# Molecule Generation

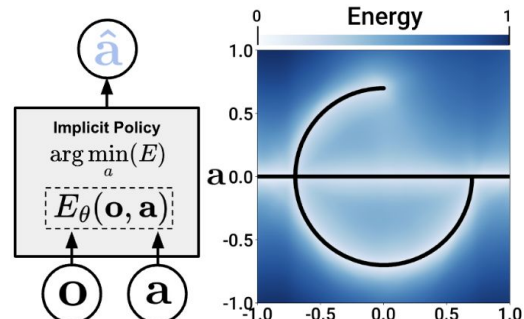


# Robot Learning

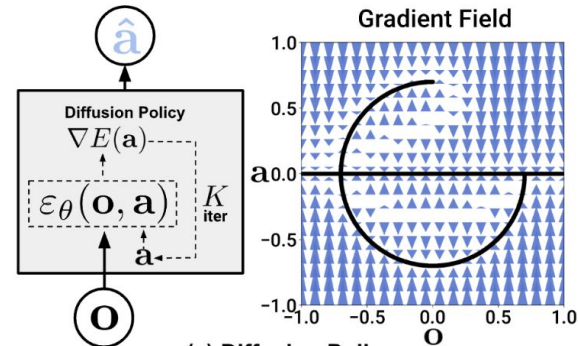
$P(\text{actions} \mid \text{past observations})$



(a) Explicit Policy



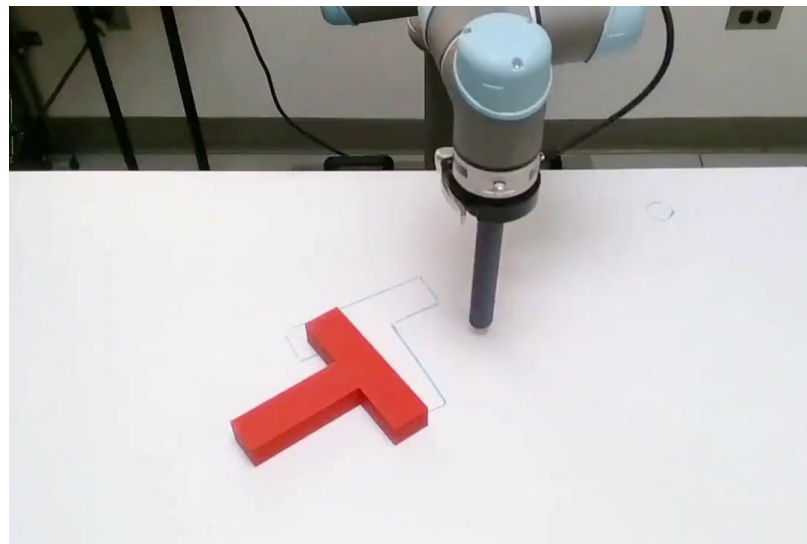
(b) Implicit Policy



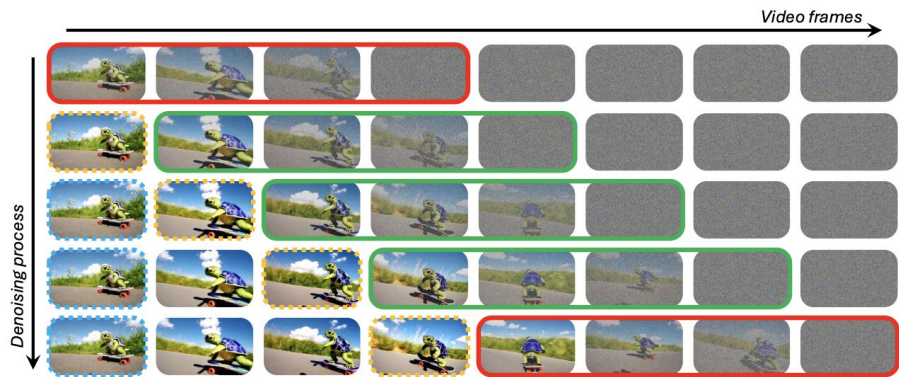
(c) Diffusion Policy



Diffusion Policy



# Video Generation





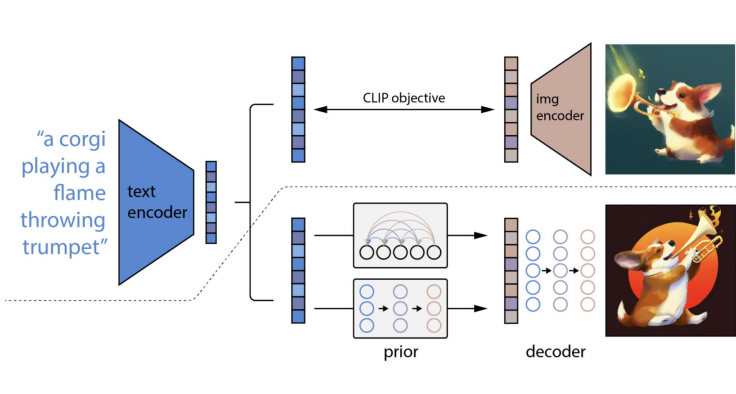
# Multimodality

## Example: Text to Image

User Input:

泰迪熊穿着戏服，站在太和殿前唱京剧

A teddy bear, wearing a costume, is standing in front of the Hall of Supreme Harmony and singing Beijing opera



# Future Studies in Generative Modeling

Courses:

[Stanford CS236 Deep Generative Models](#) by Prof. Stefano Ermon

[MIT 6.S978 Deep Generative Models](#) by Prof. Kaiming He

To play around with generative models:

<https://github.com/li-hong-yue/GenerativeModelsZoo>

Thank you!

# References

[Stanford CS236 Deep Generative Models](#)

[MIT 6.S978 Deep Generative Models](#)