Deep Generative Models
LMS Code 013672598

Dr. Zulqarnain Khan

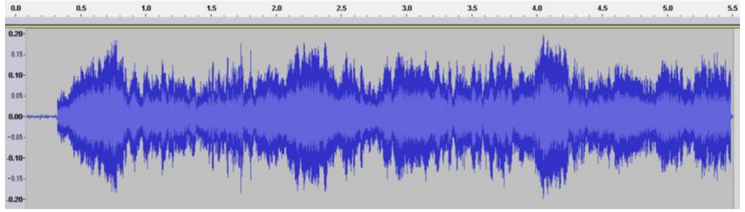
Slides from:

Stefano Ermon

Introduction

Challenge: understand complex, unstructured inputs

Computer Vision

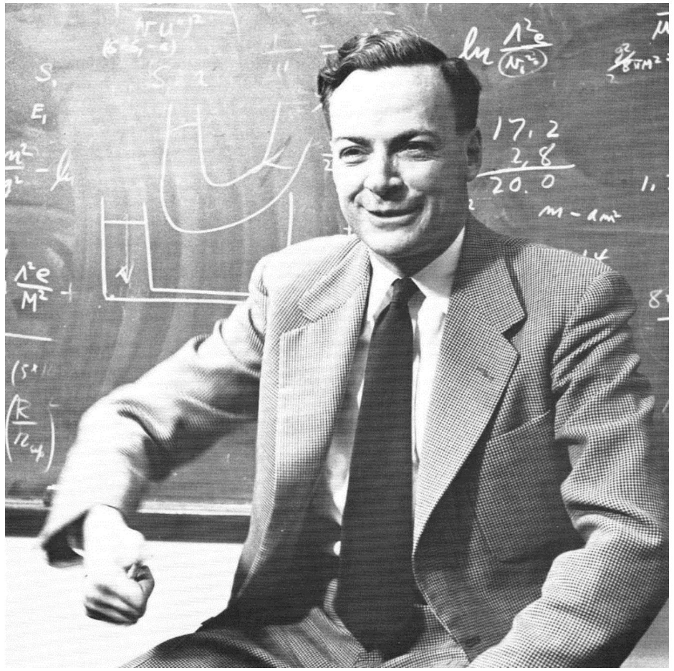


Speech

Natural Language Processing

Robotics

Introduction



Richard Feynman: "*What I cannot create, I do not understand*"

Generative modeling: "*What I understand, I can* ***create***"

Generative Modeling: Computer Graphics

How to generate natural images with a computer?

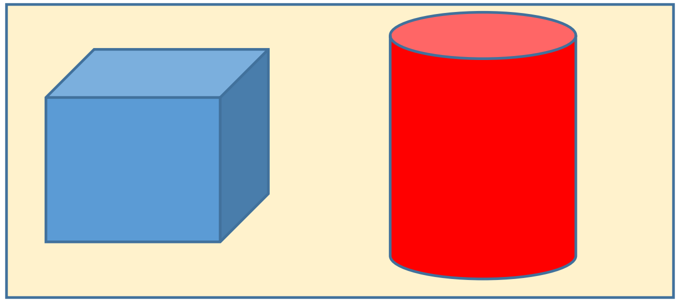
High level description

Cube(color=blue, position=(x,y,z), size=…)
Cylinder(color=red, position=(x’,y’,z’), size=..)

Generation (graphics)

Inference (vision as inverse graphics)

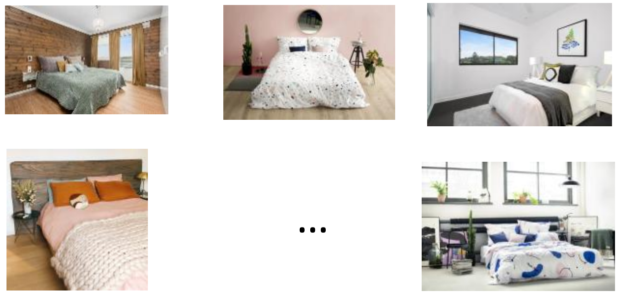
Raw sensory outputs



Many of our models will have **similar structure (generation + inference)**

Statistical Generative Models

Statistical generative models are **learned from data**



Data
(e.g., images of bedrooms)

+

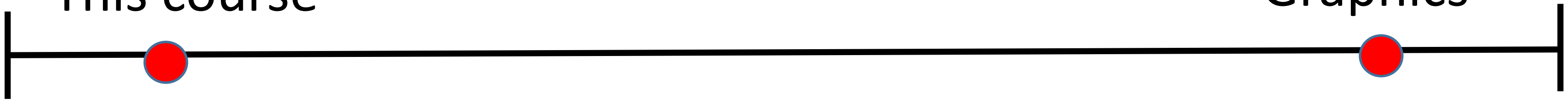
Prior Knowledge
(e.g., physics, materials, ..)

Priors are always necessary, but there is a spectrum

This course

Graphics

Data



Prior
Knowledge

Statistical Generative Models

A statistical generative model is a **probability distribution** p(x)

- **Data:** samples (e.g., images of bedrooms)
- **Prior knowledge:** parametric form (e.g., Gaussian?), loss function (e.g., maximum likelihood?), optimization algorithm, etc.

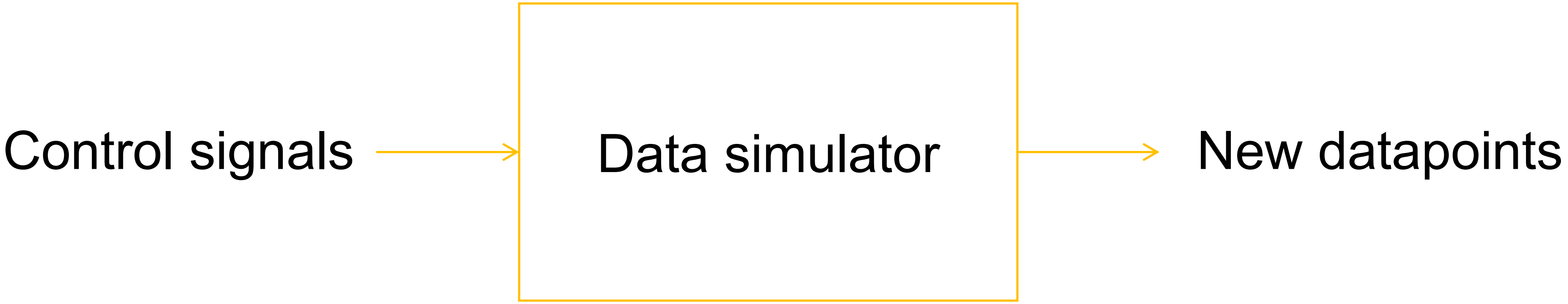
Image x  → A probability distribution p(x) → scalar probability p(x)

It is generative because **sampling from p(x) generates new images**

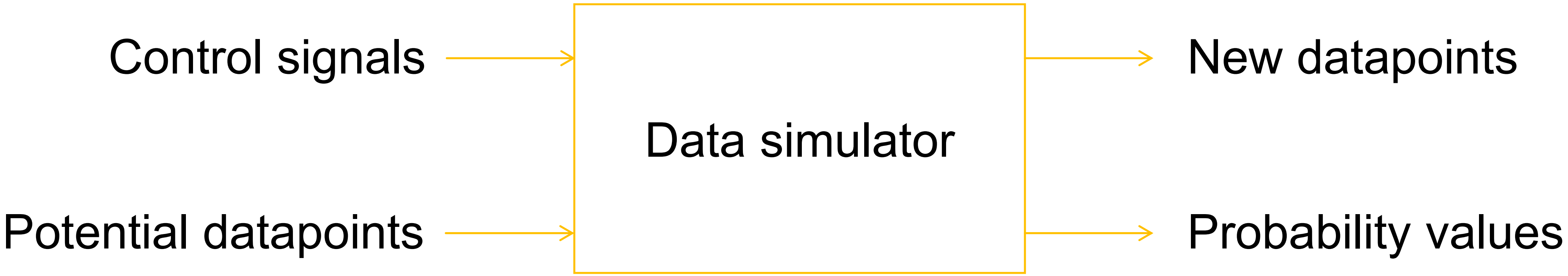
 … 

Building a simulator for the data generating process

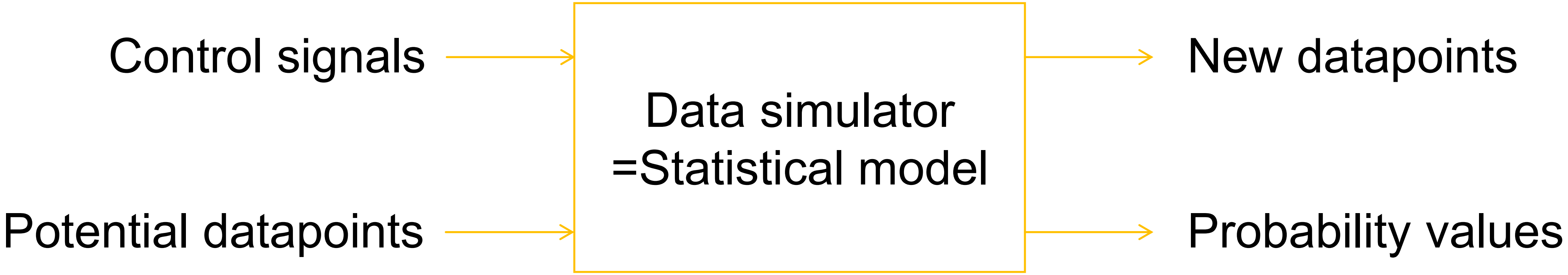
Building a simulator for the data generating process



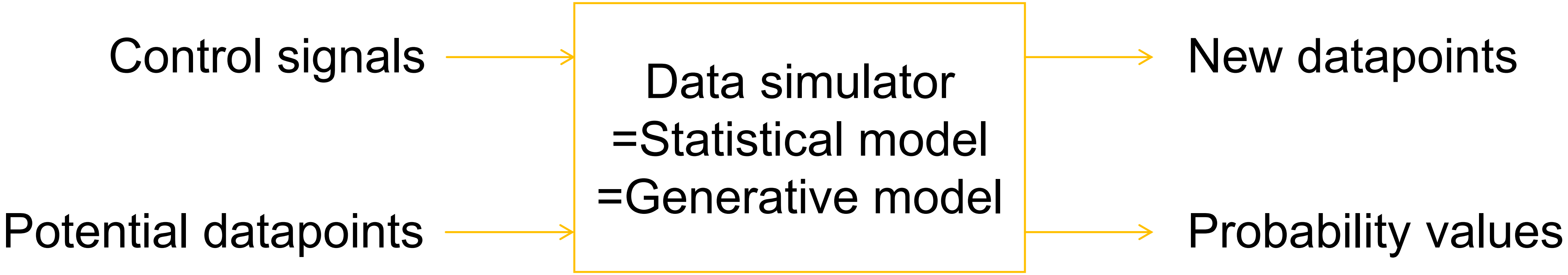
Building a simulator for the data generating process



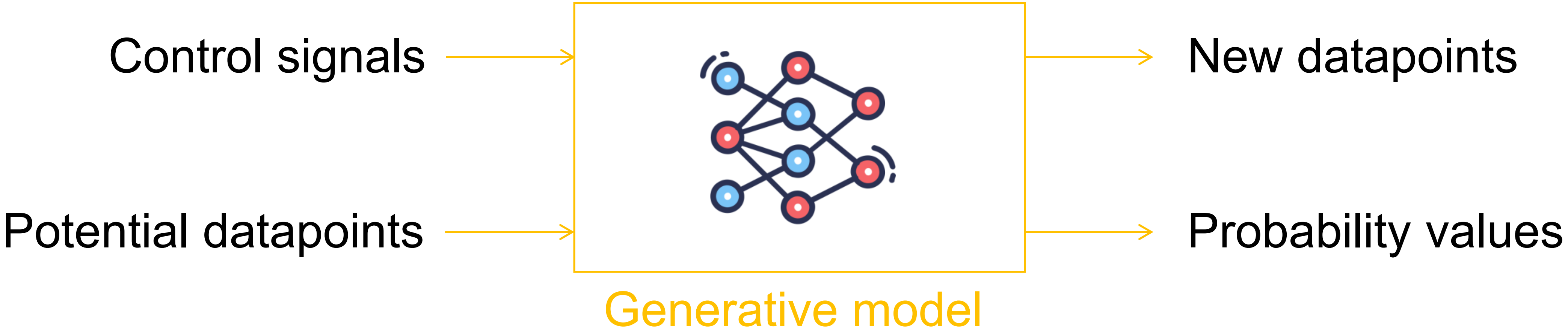
Building a simulator for the data generating process



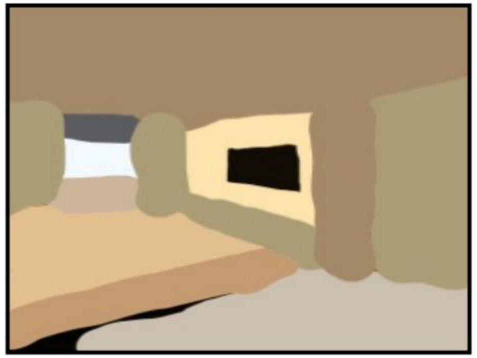
Building a simulator for the data generating process

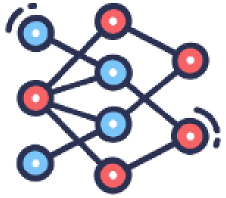


Building a simulator for the data generating process



Data generation in the real world



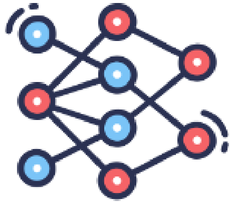


Generative model
of realistic images

Generate

Stroke paintings to realistic images
[Meng, He, Song, et al., ICLR 2022]

"Ace of Pentacles"

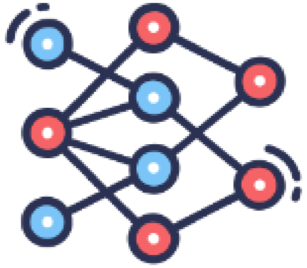


Generative model
of paintings

Generate

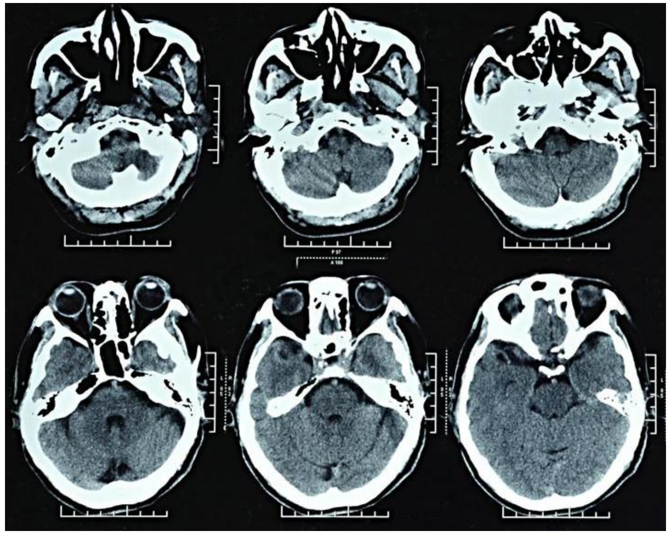
Language-guided artwork creation
https://chainbreakers.kath.io @RiversHaveWings

Solving inverse problems with generative models



Generative model
of medical images

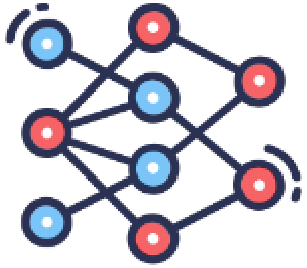
Generate



Medical image reconstruction
[Song et al., ICLR 2022]

Outlier detection with generative models

High probability



Generative model of traffic signs

Low probability

Outlier detection

[Song et al., ICLR 2018]

Progress in Generative Models of Images -- GANs

Ian Goodfellow, 2019

Progress in Generative Models of Images – Diffusion Models

Song et al., Score-Based Generative Modeling through Stochastic Differential Equations, 2021

Text2Image Diffusion Models

User input:

An astronaut riding a horse

Text2Image Diffusion Models

User input:

A perfect Pakistani meal

Text2Image Diffusion Models

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

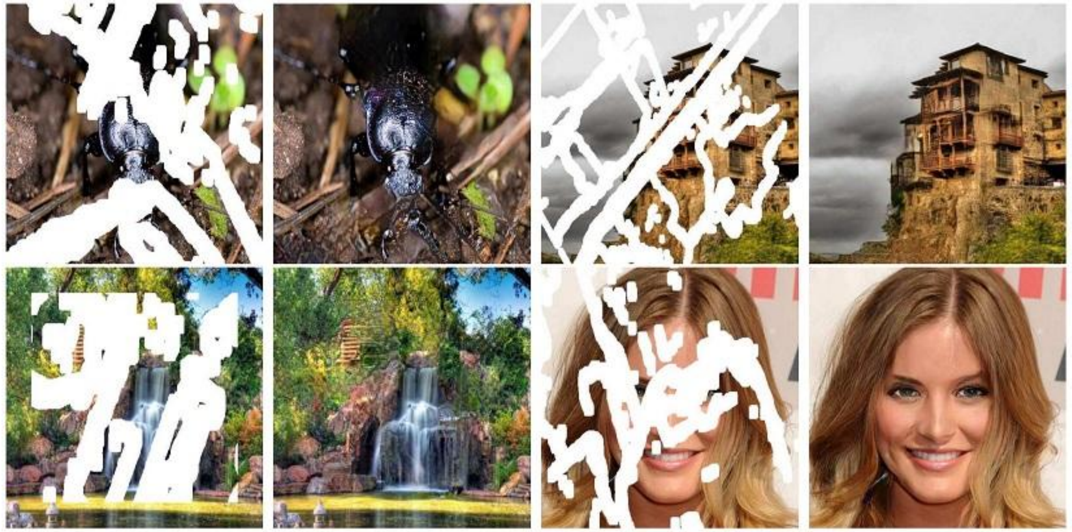
A model of a cafe adorned with indoor plants. Wooden beams crisscross above, and a cold brew station stands out with tiny bottles and glasses

Progress in Inverse Problems

P(high resolution | low resolution)

Menon et al, 2020

P(full image| mask)



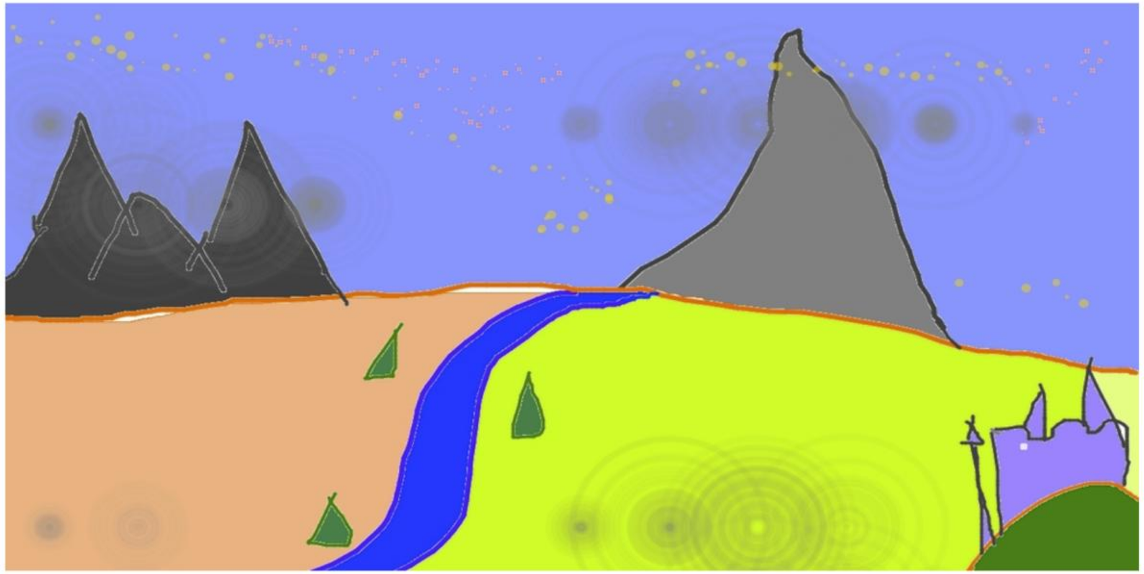
Liu al, 2018

P(color image| greyscale)

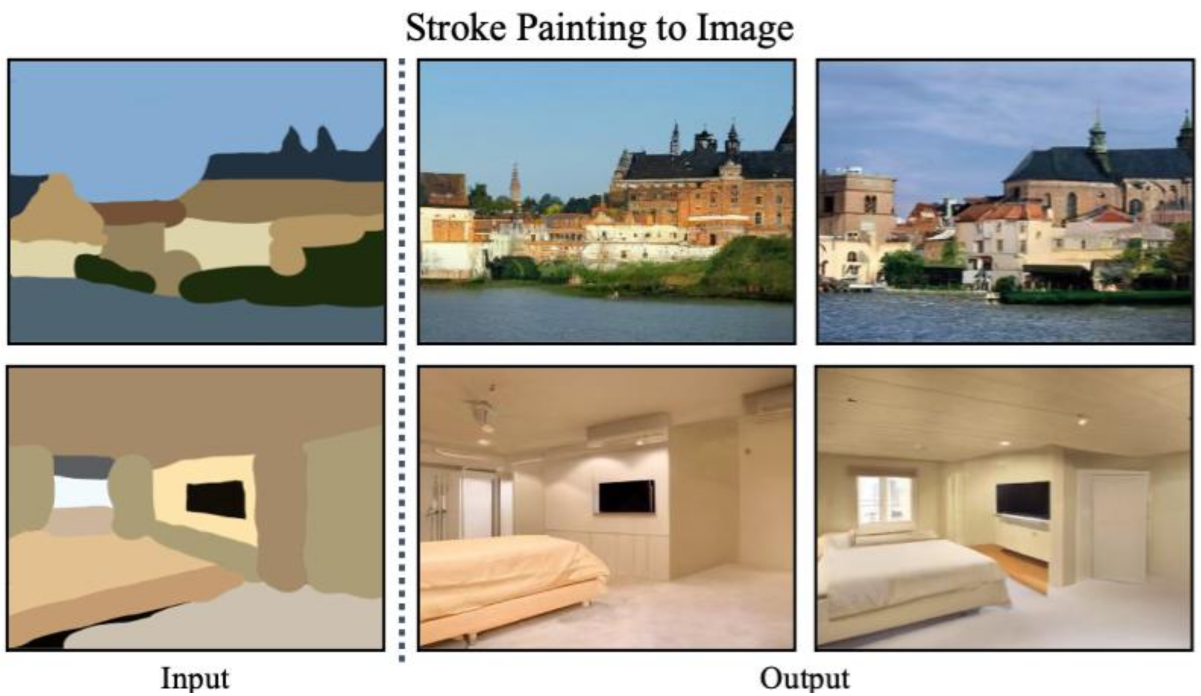
Antic, 2020

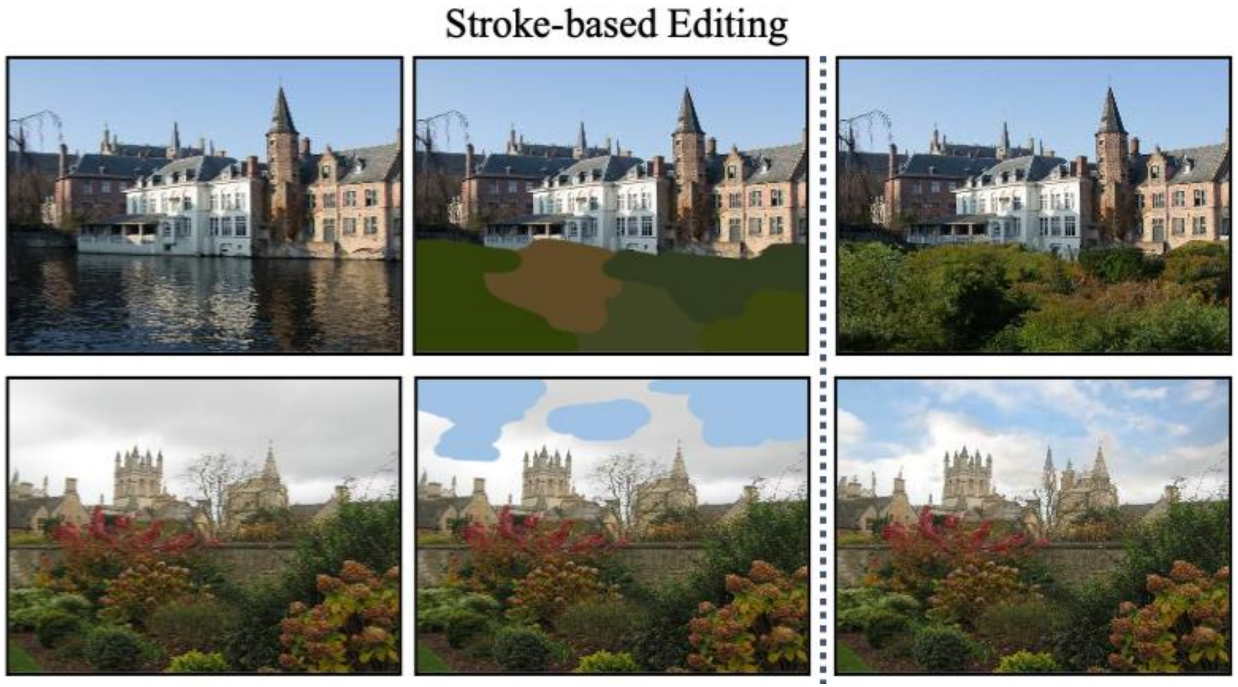
Progress in Inverse Problems

User input:

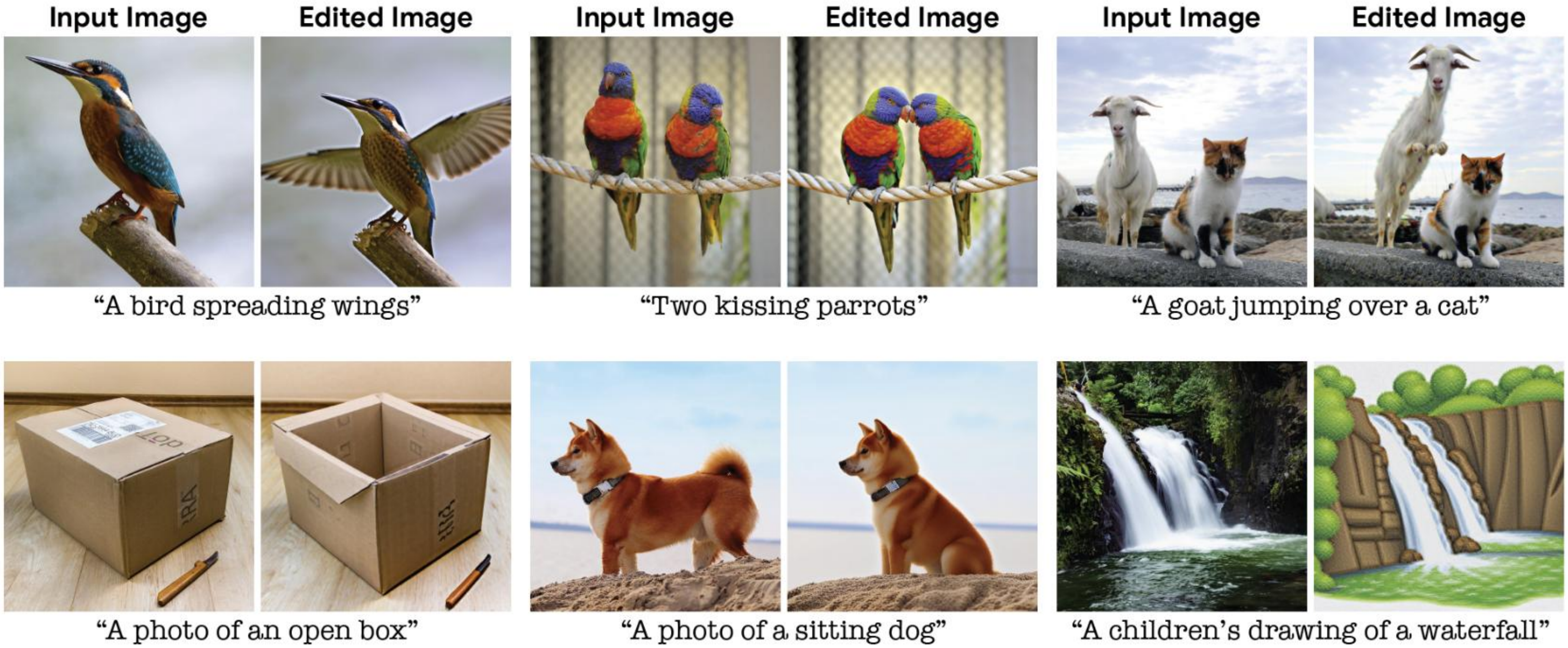


Progress in Inverse Problems



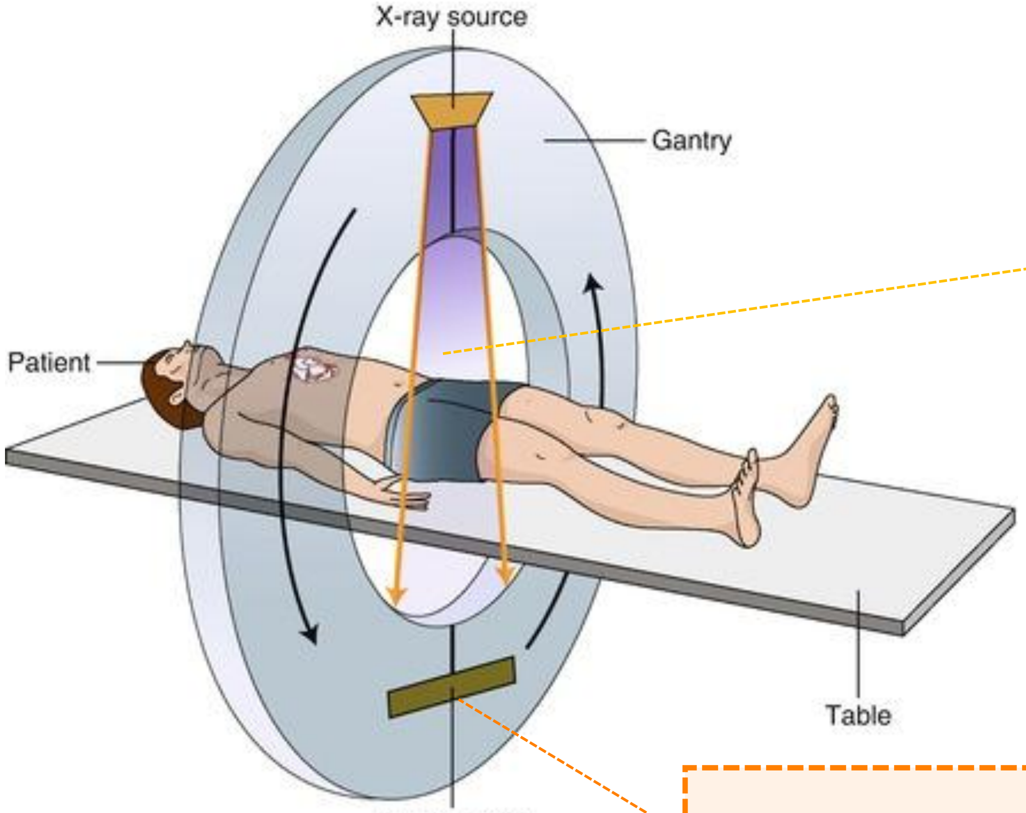


Progress in Inverse Problems



Kawar et al., 2023

Medical image reconstruction



Cross-sectional image

$\mathbf{x} \mid \mathbf{y}$

$\mathbf{y}$ 

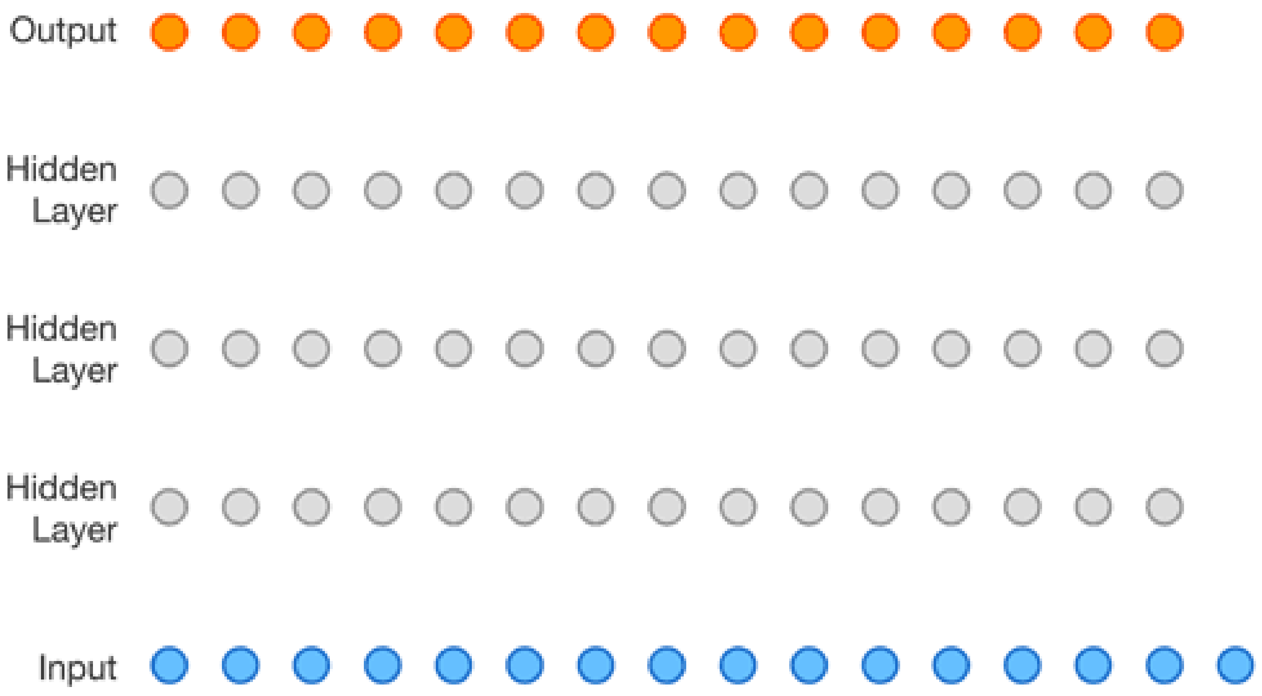
Sparse-view sinogram

Sparse-view computed tomography (CT)

Forward model $p(\mathbf{y} \mid \mathbf{x})$ is given by physical simulation

WaveNet

Generative model of speech signals



Text to Speech

- Parametric
- Concatenative
- WaveNet

Unconditional

Music

van den Oord et al, 2016c

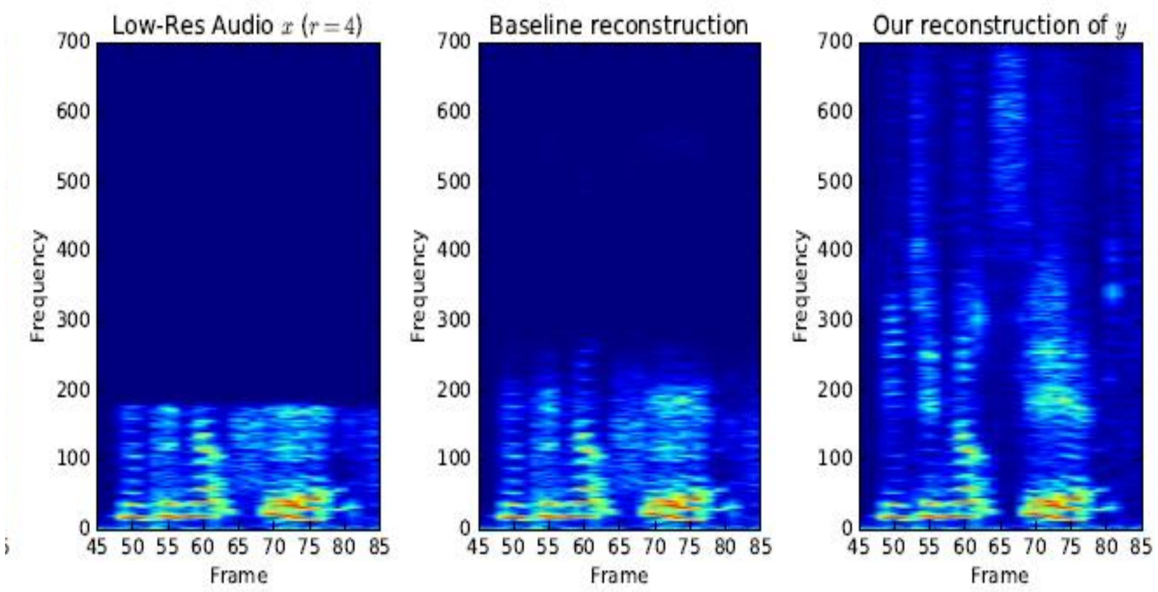
Diffusion Text2Speech

Generative model of speech signals

Betker, Better speech synthesis through scaling 2023

Audio Super Resolution

Conditional generative model P(high-res signal | low-res audio signal)



Low res signal

High res audio signal

Kuleshov et al., 2017

Language Generation

Custom prompt

To get an A+ in deep generative models, students have to

Completion

To get an A+ in deep generative models, students have to be willing to work with problems that are a whole lot more interesting than, say, the ones that most students work on in class. If you're a great student, the question above can be avoided and you'll be able to do great work, but if you're not, you will need to go beyond the basics before getting good.

Now to be clear, this advice is not just for the deep-learning crowd; it is good advice for any student who is taking his or her first course in machine learning.

The key point is that if you have a deep, deep brain of a computer scientist, that's just as important to you.

P(next word | previous words)

Radford et al., 2019
Demo from talktotransformer.com

Language Generation – ChatGPT

You can try the same prompt in ChatGPT, notice the improvement!

Machine Translation

Conditional generative model P(English text| Chinese text)

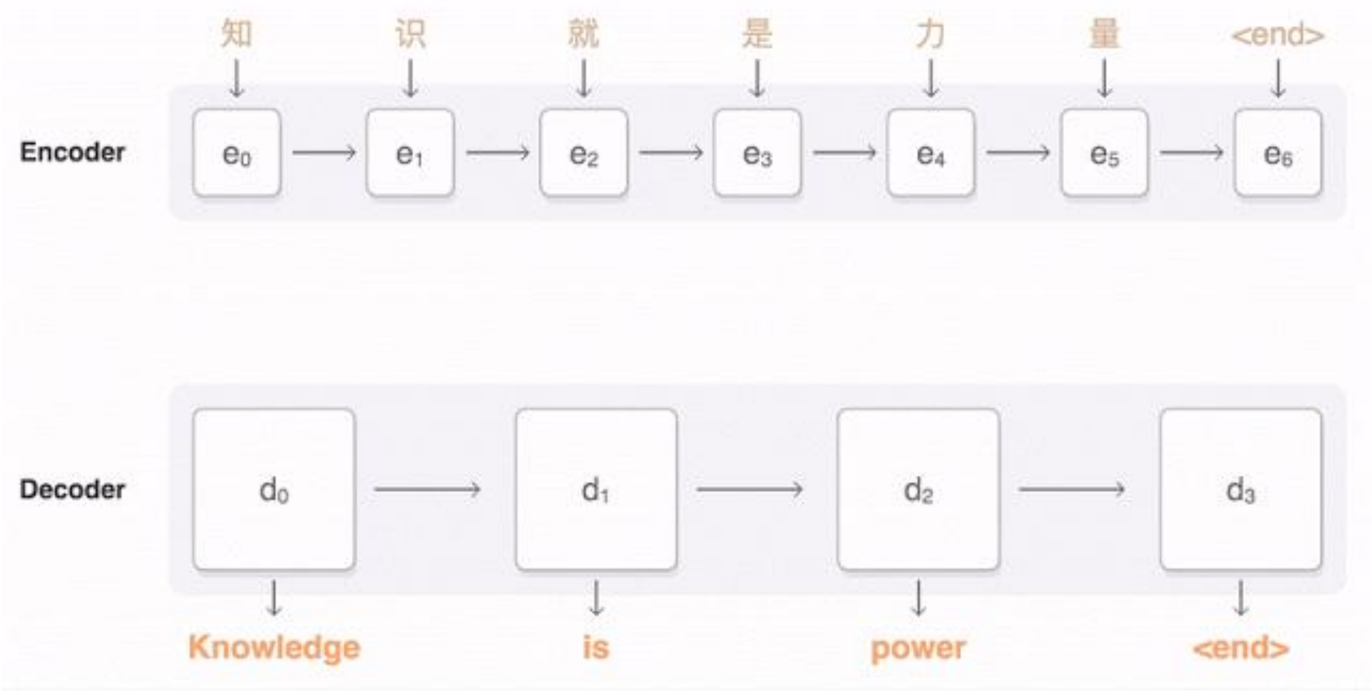


Figure from Google AI research blog.

Code Generation

parse_expenses.py | write_sql.go | sentiment.ts | addresses.rb

```python
import datetime

def parse_expenses(expenses_string):
    """Parse the list of expenses and return the list of triples (date, va
```

OpenAI Codex

Video Generation

Suddenly, the walls of the embankment broke and there was a huge flood

Video Generation

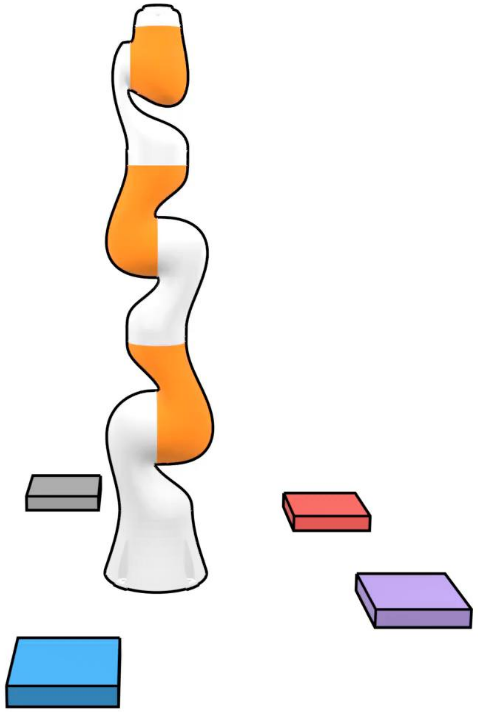
a couple sledding down a snowy hill on a tire roman chariot style

Video Generation

Imitation Learning

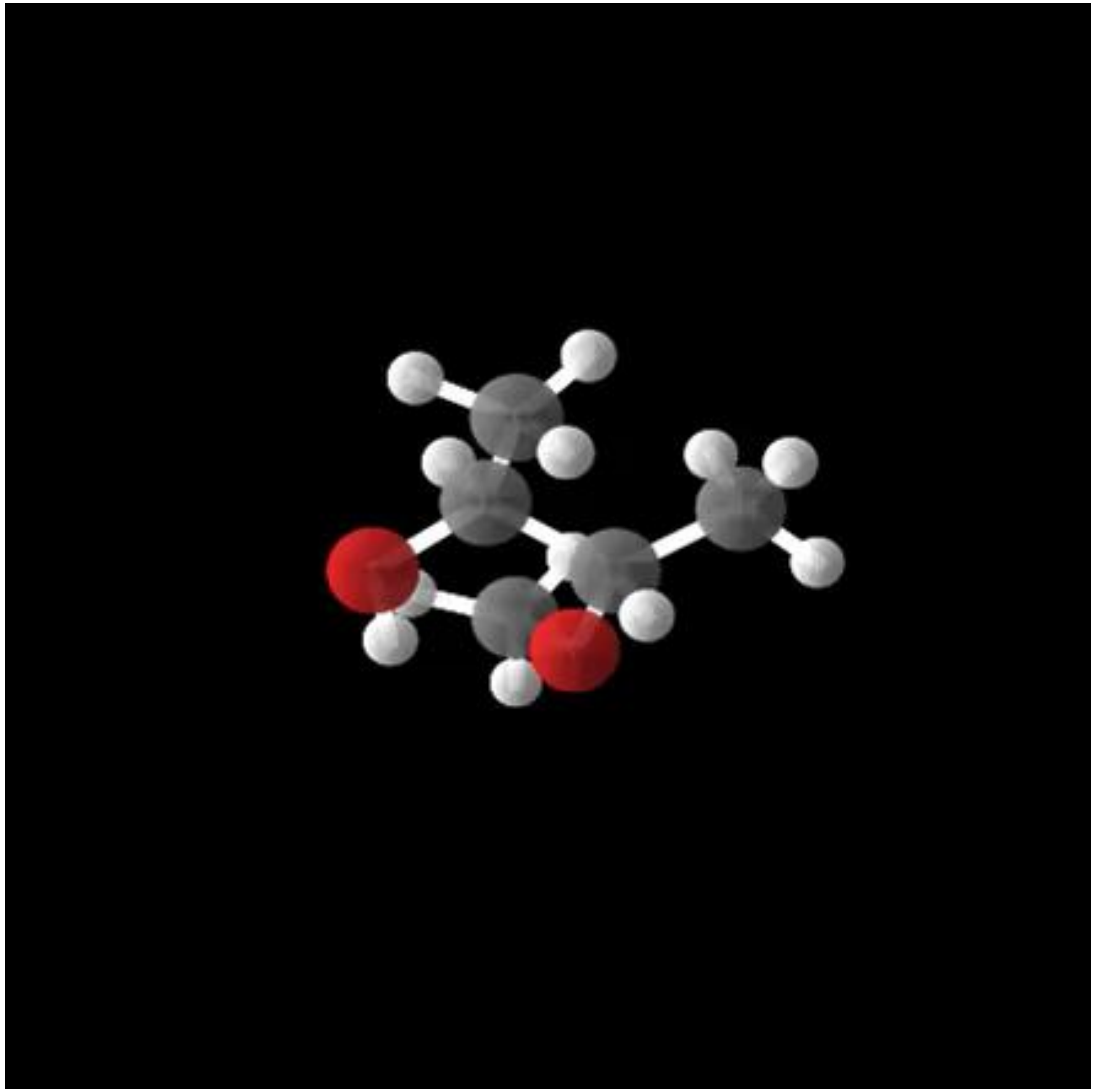
Conditional generative model P(actions | past observations)

Li et al., 2017

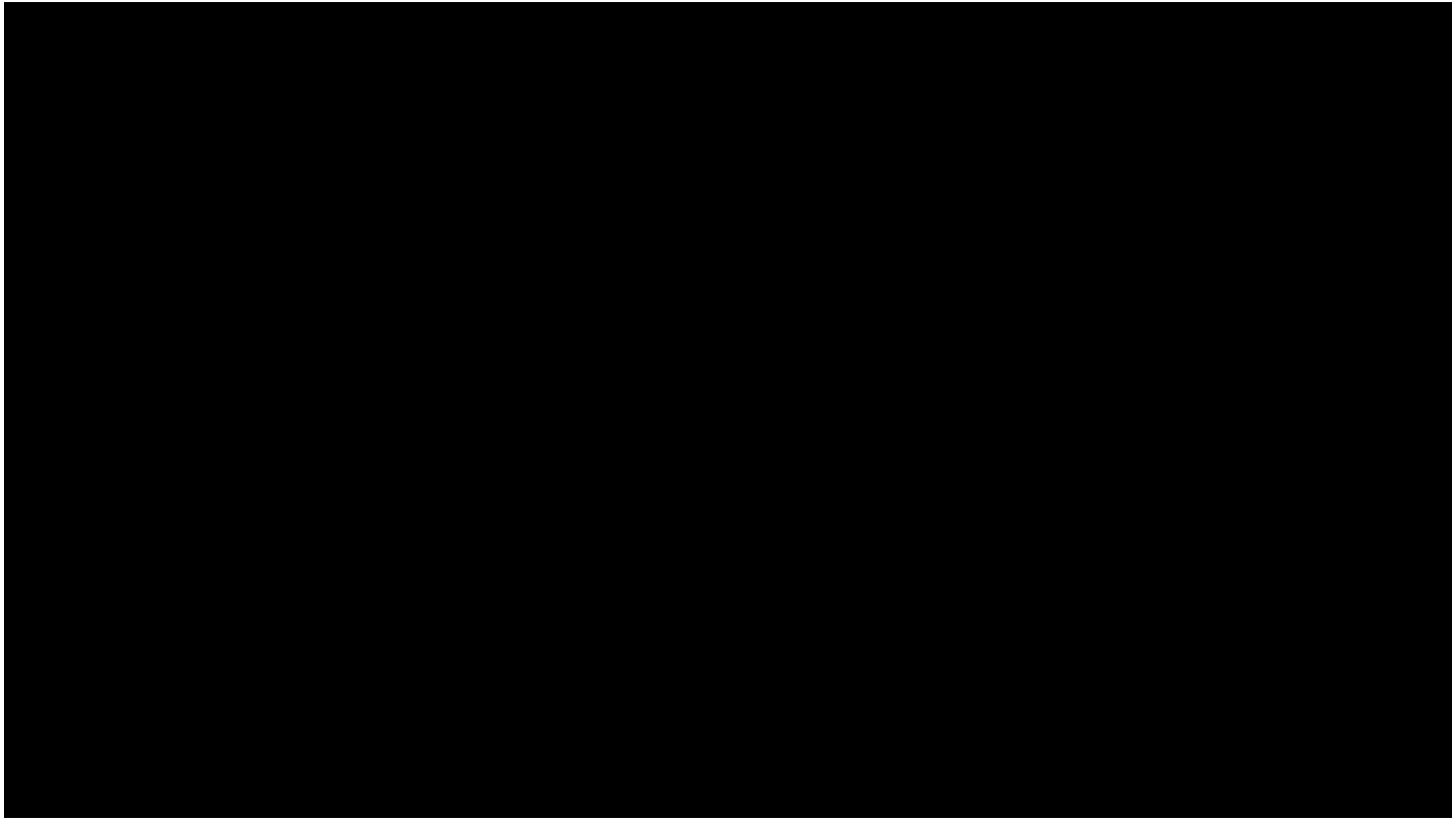


Janner et al., 2022

Molecule generation



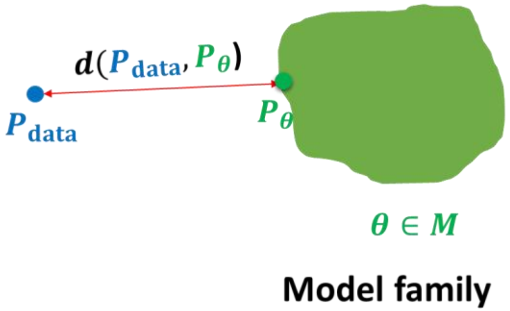
DeepFakes



Roadmap and Key Challenges

- **Representation**: how do we model the joint distribution of many random variables?
  - Need compact representation
- **Learning**: what is the right way to compare probability distributions?

$\mathbf{x}_i \sim P_{\text{data}}$
$i = 1, 2, \ldots, n$



- **Inference**: how do we invert the generation process (e.g., vision as inverse graphics)?
  - Unsupervised learning: recover high-level descriptions (features) from raw data

Syllabus

- Fully observed likelihood-based models
  - Autoregressive
  - Flow-based models
- Latent variable models
  - Variational learning
  - Inference amortization
  - Variational autoencoder
- Implicit generative models
  - Two sample tests, embeddings, F-divergences
  - Generative Adversarial Networks
- Energy Based Models
- Score-based Diffusion Generative Models
- Learn about algorithms, theory & applications

Prerequisites

- Basic knowledge about machine learning.
- Basic knowledge of probabilities and calculus:
  - Gradients, gradient-descent optimization, backpropagation
  - Random variables, independence, conditional independence
  - Bayes rule, chain rule, change of variables formulas
- Proficiency in some programming language, preferably Python, required.

Logistics

- There is no required textbook. Reading materials and course notes will be provided.
- Suggested Reading: *Deep Learning* by Ian Goodfellow, Yoshua Bengio, Aaron Courville. Online version available free here.