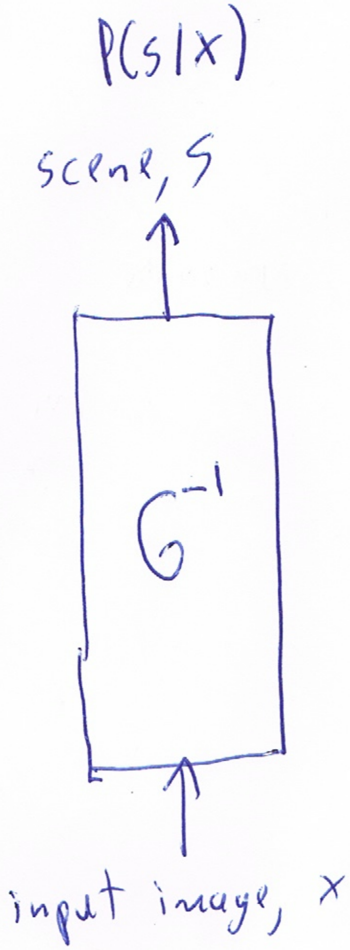
discriminative approach

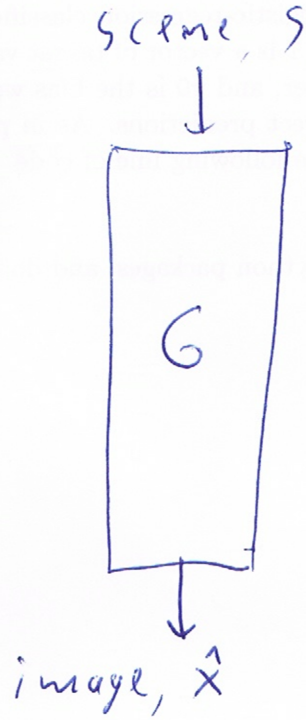
(supervised)



$$\dim(s) \ll \dim(x)$$

s = scene representation.

Generative Approach (unsupervised)



input image, x

find s that makes $\hat{x} \approx x$.

more generally, find $p(s|x)$.

~~the closer $\hat{x}$ is to x, the more~~ like

$p(s|x)$ is high iff $\hat{x} \approx x$,

Intractable in general.

use variational approximations (later).

Note:

Use Bayes Rule to compute $P(S|X)$:

prior on scenes

$$P(S|X) = \frac{P(X|S)\cdot P(S)}{P(X)}$$

$$P(X) = \sum_S P(X,S) = \sum_S P(X|S)\cdot P(S)$$

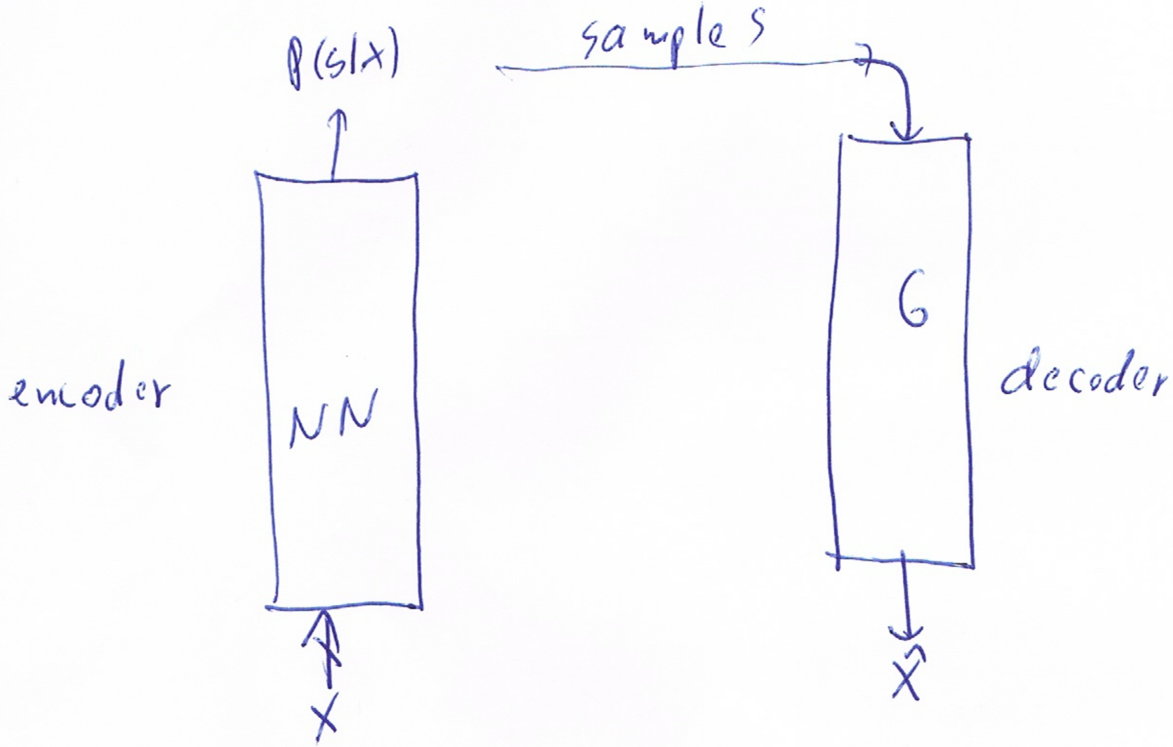
$$P(X|S) = P(X|\hat{X}(S))$$ usually simple (eg. Gaussian)

~~P(X~~ eg. $P(X|\hat{X}) = \dfrac{e^{-\|X-\hat{X}\|^2/2\sigma^2}}{\sqrt{2\pi}\,\sigma}$

Variational Autoencoder

Simplifying assumption:

~~$p(s|x)$~~ Train a NN to compute (an approximation to) $p(s|x)$



train NN to make $\hat{x} \approx x$ (for all training points)

~~Can train N~~

G can be given or can be trained ~~with NN~~ simultaneously with NN

~~Note: must~~

Note: must perform gradient descent through both G & NN (& the sampling operation).

No problem if G is a NN.

If G is a graphics program, we must be able to differentiate it. (Hence "differentiable rendering").

Note: unlike discriminative approach, vae is unsupervised.

No ~~scene~~ labels needed.