a)

text-to-image autoencoder

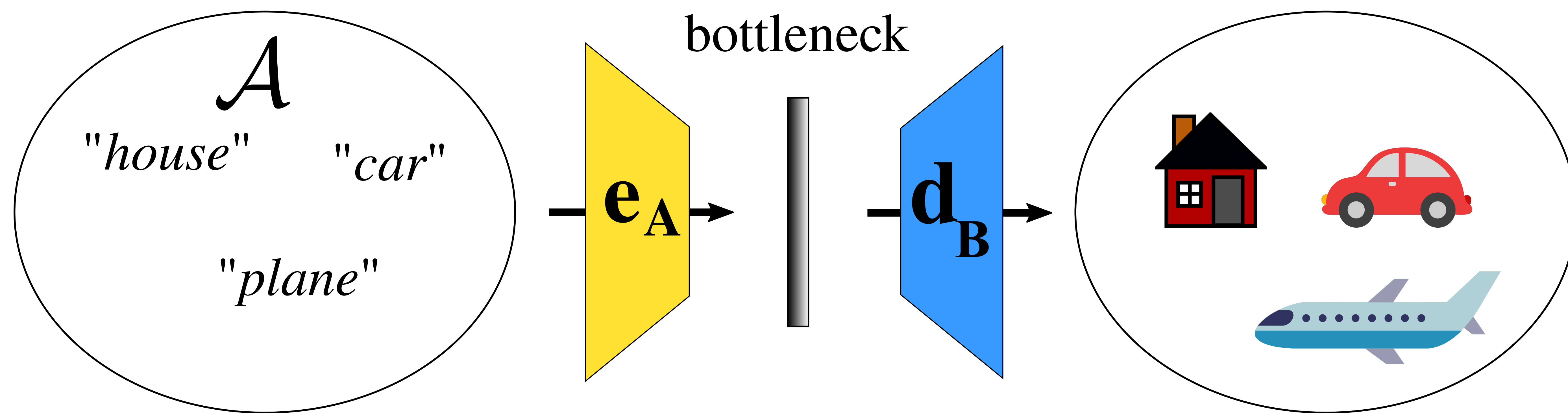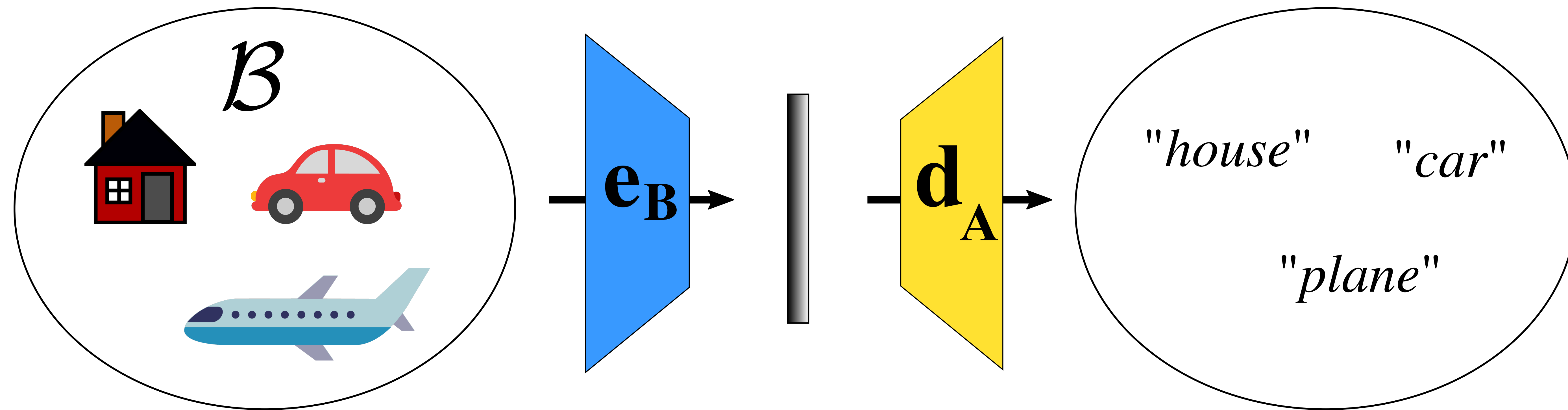$\mathcal{A}$
"house" "car"
"plane"

information bottleneck

e_A

d_B

image-to-text autoencoder

$\mathcal{B}$

e_B

d_A

"house" "car"
"plane"

b)

training
Fritz

$\mathcal{Y}$
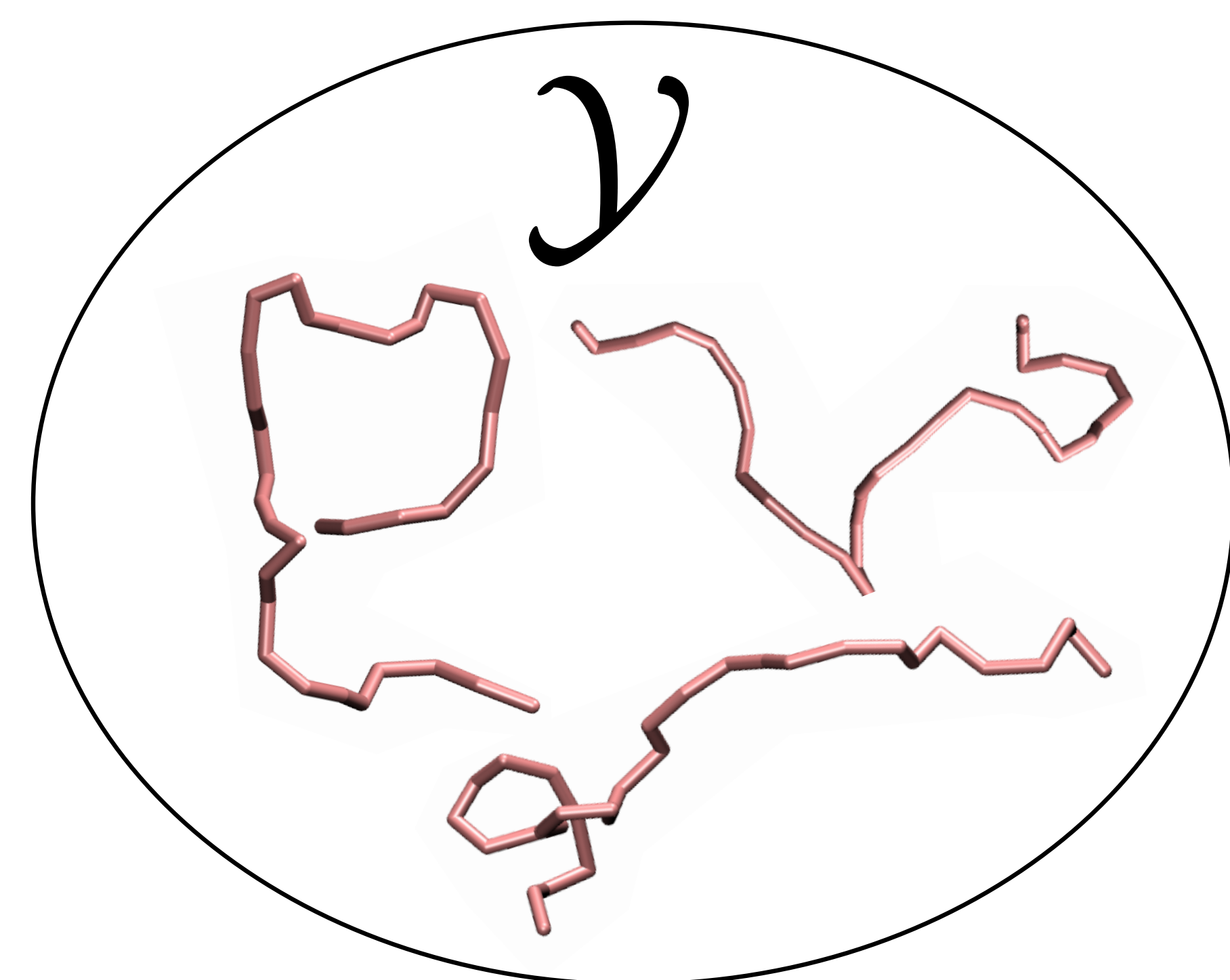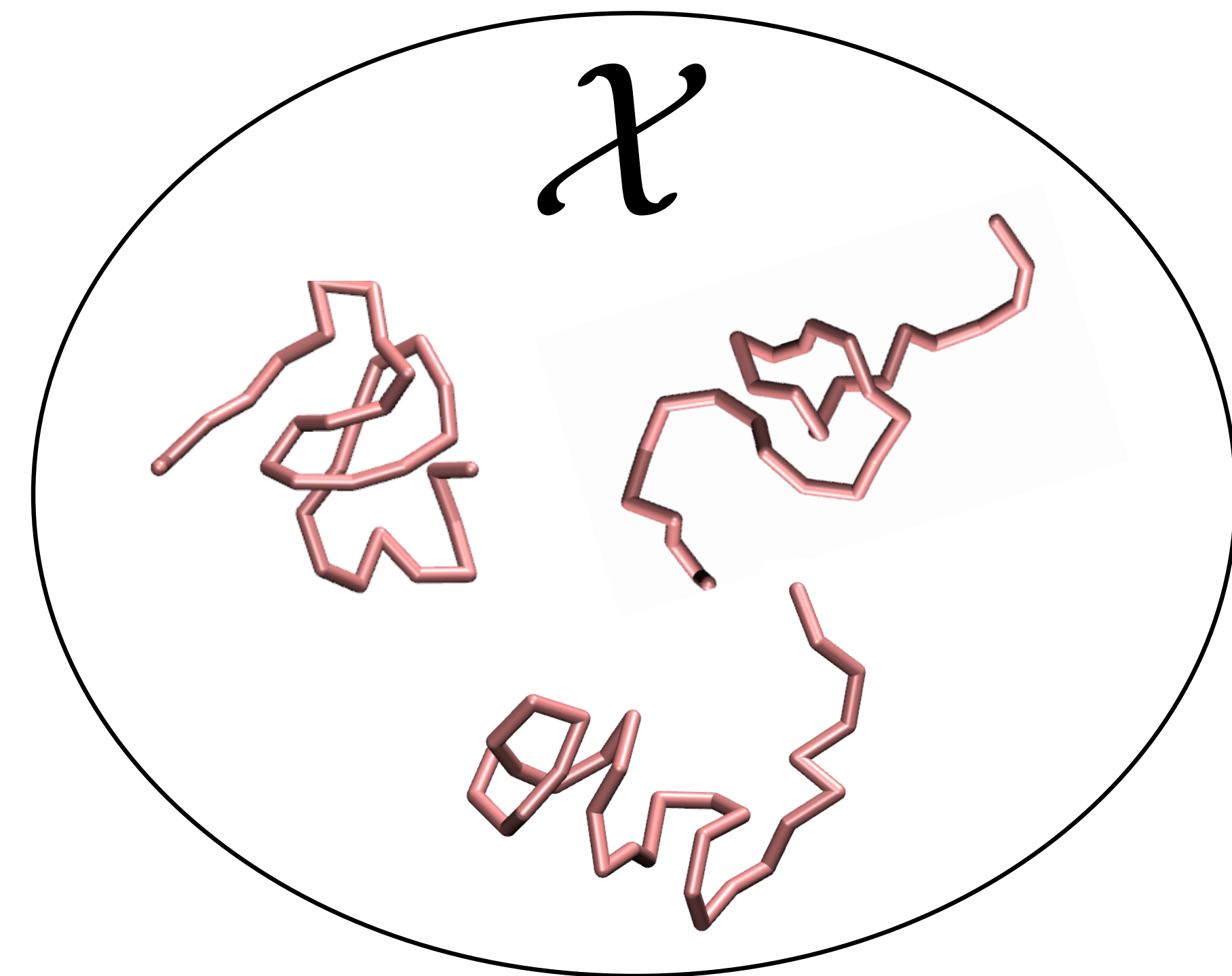
resolution bottleneck
$s = 6$

e_6

g_6

morphing
Kremer-Grest

$\mathcal{X}$

e_6