Fast Relative Entropy Coding with A* Coding

Gergely Flamich* Stratis Markou*

José Miguel Hernández-Lobato

Computational and Biological Learning Lab
Department of Engineering

UNIVERSITY OF CAMBRIDGE

Motivation

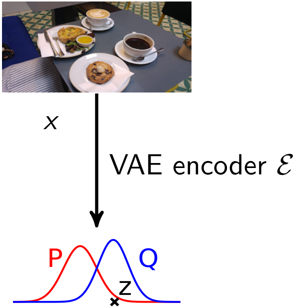
Motivation

Learned compression with VAEs

x

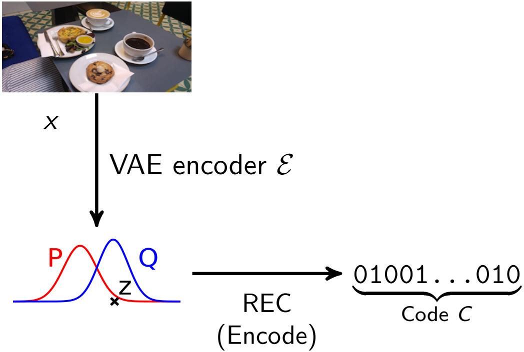
Motivation

Learned compression with VAEs



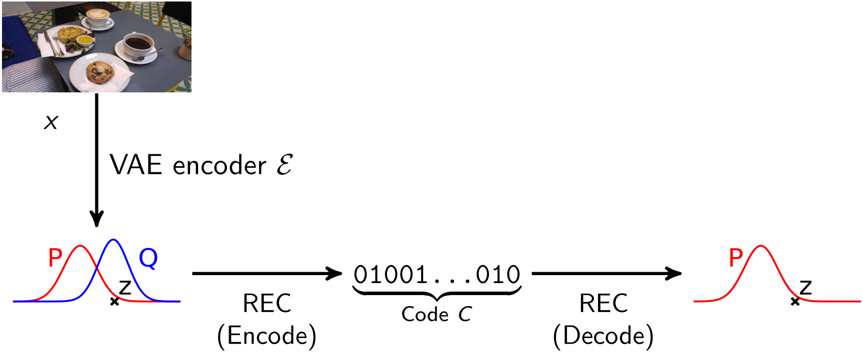
Motivation

Learned compression with VAEs



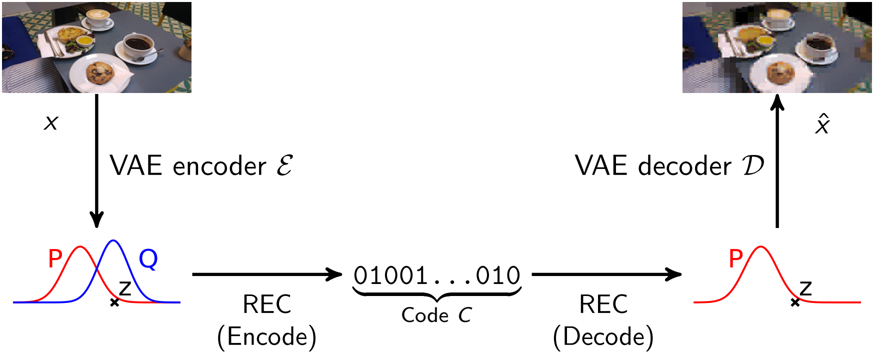
Motivation

Learned compression with VAEs



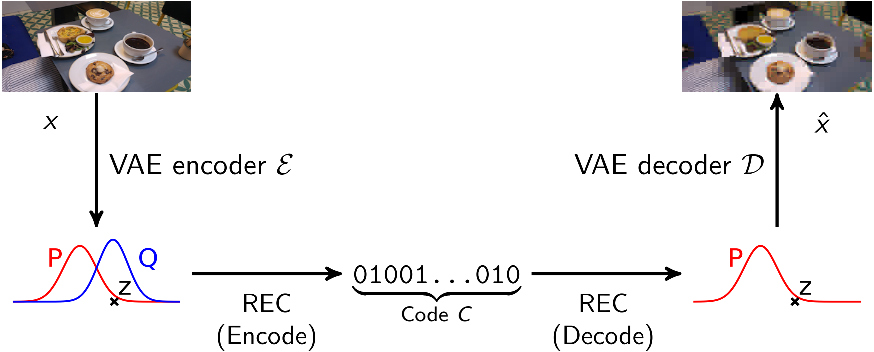
Motivation

Learned compression with VAEs



Motivation

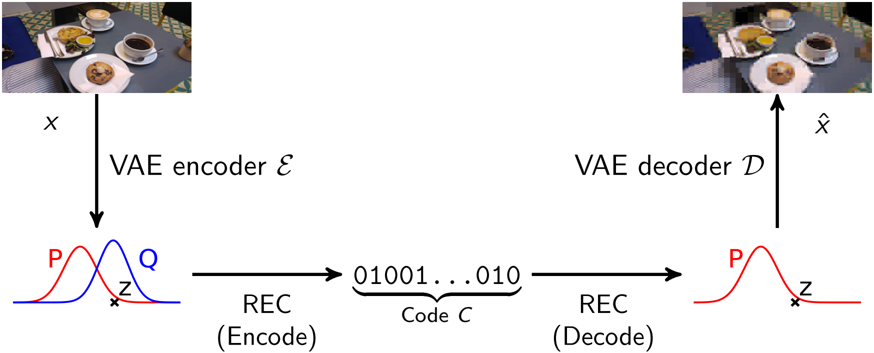
Learned compression with VAEs



✓ Does not require quantizing z.

Motivation

Learned compression with VAEs



- ✓ Does not require quantizing z.
- ✓ Lossless and lossy compression, and further applications.

Relative Entropy Coding

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q.

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

- Is decodable by Bob.

Relative Entropy Coding

Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

- Is decodable by Bob.
- Represents exact sample from Q.

Relative Entropy Coding

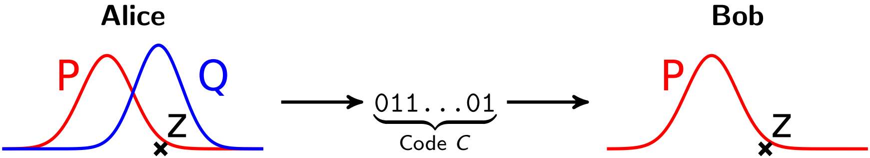
Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

- Is decodable by Bob.
- Represents exact sample from Q.



Relative Entropy Coding

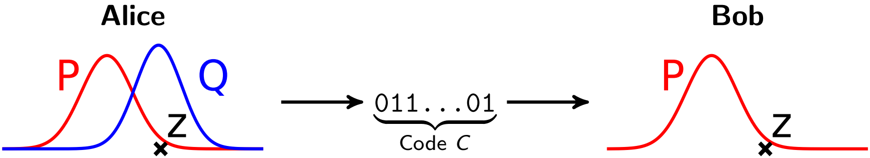
Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

- Is decodable by Bob.
- Represents exact sample from Q.



- As small codelength $|C|$ as possible.

Relative Entropy Coding

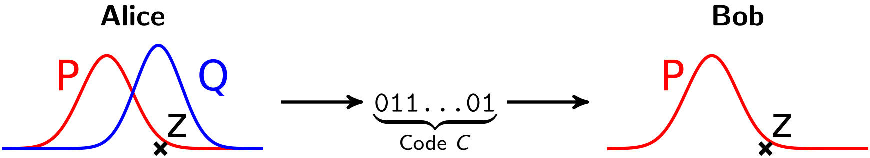
Relative Entropy Coding

Setup: Alice holds target distribution Q. Alice and Bob share

- Proposal distribution P.
- **Public** sequence of fair coin tosses $S = (s_1, s_2, \dots)$.

Goal: Alice uses P, S and Q to produce code C which

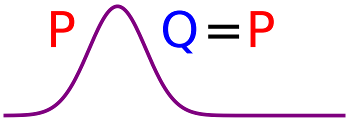
- Is decodable by Bob.
- Represents exact sample from Q.

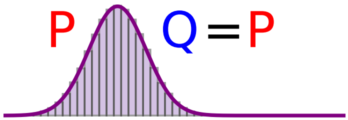


- As small codelength $|C|$ as possible.
- As short runtime as possible.

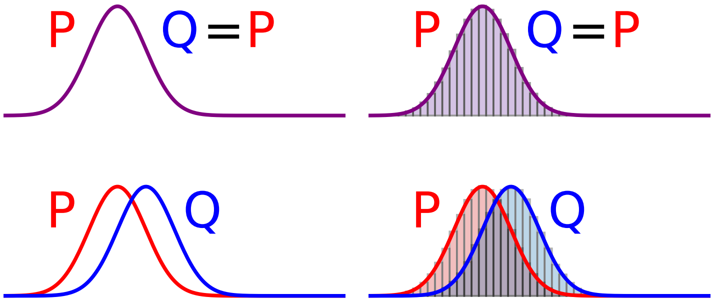
Intuition

Intuition

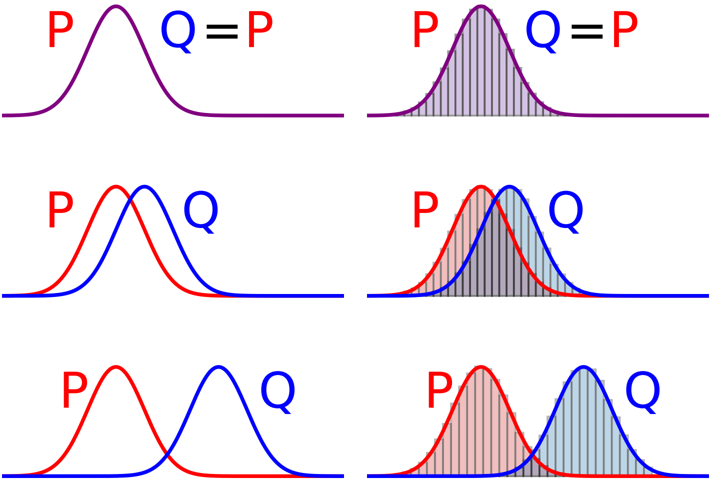




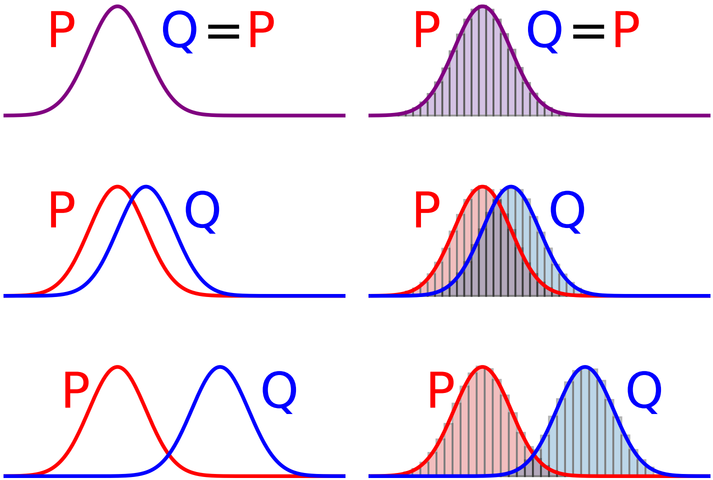
Intuition



Intuition



Intuition



Challenge and our solution

Challenge and our solution

Runtime of general REC (Agustsson and Theis, 2020)

Without additional assumptions, any REC scheme will have

$$\Omega(\exp(D_{\mathrm{KL}}[Q \| P]))$$

expected runtime.

Challenge and our solution

Runtime of general REC (Agustsson and Theis, 2020)

Without additional assumptions, any REC scheme will have

$$\Omega(\exp(D_{\mathrm{KL}}[Q \| P]))$$

expected runtime.

Our Solution

A* coding, a REC algorithm based on A* sampling.

Theoretical results

Theoretical results

Codelength of A* coding (informal)

Let C be the code returned by A* coding. Then

$$\mathbb{E}\left[|C|\right] = \mathcal{O}(D_{\mathrm{KL}}[Q \| P]).$$

Theoretical results

Codelength of A* coding (informal)

Let C be the code returned by A* coding. Then

$$\mathbb{E}\left[|C|\right] = \mathcal{O}(D_{\mathrm{KL}}[Q\|P]).$$

Runtime of AS* coding (informal)

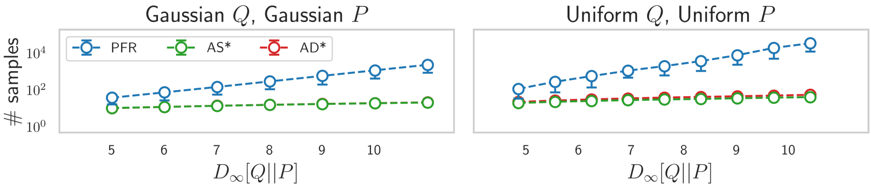
For unimodal q/p, the expected runtime of AS* coding is

$$\mathbb{E}\left[T\right] = \mathcal{O}(D_{\infty}[Q\|P]) = \mathcal{O}\left(\log \sup_{z\in\mathbb{R}} \frac{q(z)}{p(z)}\right).$$

Empirical results

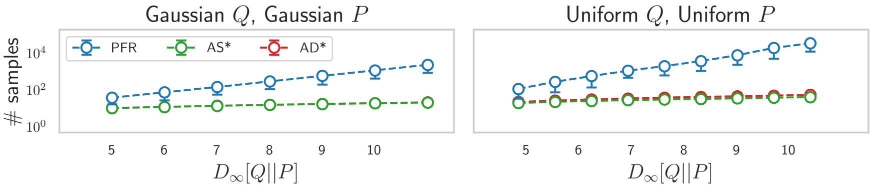
Empirical results

Synthetic Experiments



Empirical results

Synthetic Experiments



Compression with VAEs (MNIST)

# Latent	Neg. ELBO	A* Coding
20	1.43 ± 0.01	1.53 ± 0.01
50	1.40 ± 0.01	1.66 ± 0.01

Summary

We introduce A* coding, a practically fast REC algorithm.

Summary

We introduce A* coding, a practically fast REC algorithm.

For more info, find us in **Hall E** at poster **#703** or read our paper.