Exploration and uncertainty in generative networks for reinforcement learning

Emilio Jorge

Chalmers University of Technology

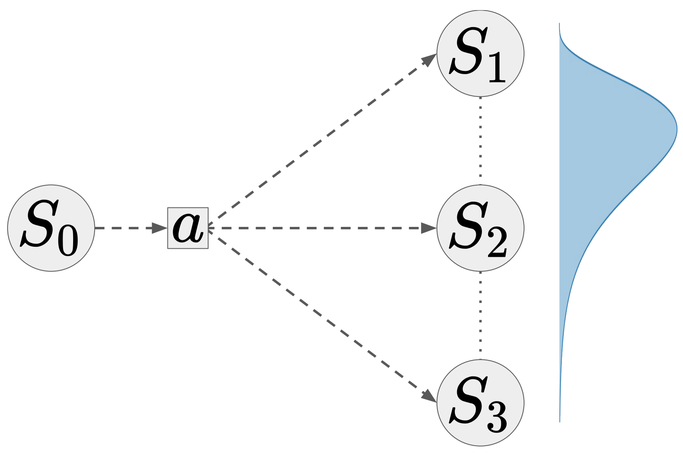
Why do we want know our uncertainty?

Safety

By avoiding what we do not know sufficiently about we can avoid behaviour that can lead to unsafe outcomes.

Exploration

In situations where safety is not an issue we can explore uncertain regions to search for potential rewarding outcomes that are still undiscovered.



What are good ways of representing uncertainty?

What do we want?

Generality. It is important for the models to be distributionally agnostic as transitions and rewards will very rarely belong to any friendly distributions with simple closed form.

Convergence. As more data is obtained it is important that the estimates converge to something that is close to the true model.

Efficiency. It needs to learn fast in terms of samples. It also needs to be able to scale reasonably with regards to computational complexity. It is especially useful if it has efficient ways of updating incrementally with data, as that is natural in reinforcement learning settings.

What might not be necessary?

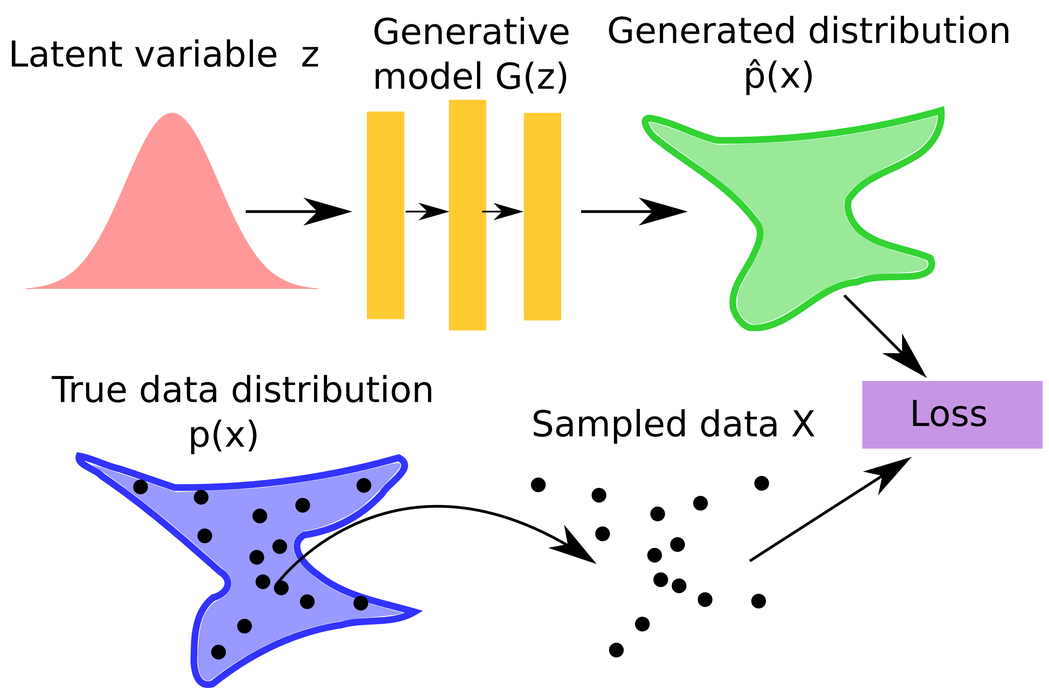
We might not always need explicit likelihoods (but it is nice to have), sampling is often enough. It might also be possible to relax some of the requirements mentioned above based on the desired application.

Generative models

Generative models allow us to approximately represent (and generate from) distribution $p(x)$ using sampled data. Popular methods are Generative Adversarial Networks (GAN) and Variational AutoEncoders (VAE).

How do they work?

They transform a sample from a latent variable z into a sample of the more complex distribution $\hat{p}(x)$ that aims to approximate the true distribution $p(x)$.



Research goals

We want to use generative models to reflect underlying uncertainty. By using this we aim to be able to create algorithms for reinforcement learning that can explore safer and/or in a more efficient way.

Any ideas or questions?
Come talk to me directly or email me at emilio.jorge@chalmers.se

CHALMERS UNIVERSITY OF TECHNOLOGY

WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM