Protein Diffusion with Reinforcement Learning

1 Forward Diffusion

$$p_{position}(x_t|x_0) = \mathcal{N}(\alpha_t x_0, \sigma_t^2 I)$$
$$p_{identity}(x_t|x_0) = \mathcal{N}(\alpha_t x_0, \sigma_t^2 I)$$

2 Loss Function

$$L_{position} = \frac{1}{2}\mathbb{E}_{\epsilon_{pos},t}[\gamma'(t)||\epsilon_{pos} - \epsilon_{pos}(x_t, t; \theta)||_2^2]$$
$$L_{identity} = \frac{1}{2}\mathbb{E}_{\epsilon_{id},t}[\gamma'(t)||\epsilon_{id} - \epsilon_{id}(x_t, t; \theta)||_2^2]$$

3 Diffusion RL

Energy and Reward function:

$$E(d) = 3((\frac{2.828}{d})^4 - (\frac{2.828}{d})^2)$$
$$R(d) = \exp(-\frac{1}{N}\sum_i^N E(d_i)) - 1$$

Loss function:

$$L_{reinforcement} = \mathbb{E}_{p_\theta}[A(\hat{x}_0)\sum_t^T \frac{1}{2c\sigma_s^2}(x_s - \frac{\alpha_s}{\alpha_t}(x_t - \sigma_t c\epsilon(x_t, t; \theta)))^2]$$

Loss function and policy gradient:

$$\begin{aligned}
\nabla_\theta J_{total\,reward} &= \mathbb{E}_{p_\theta}[A(\hat{x}_0)\sum_t^T \nabla_\theta \log p_\theta(x_s|x_t)] \\
&= \mathbb{E}_{p_\theta}[A(\hat{x}_0)\sum_t^T -\frac{1}{2c\sigma_s^2}\nabla_\theta(x_s - \frac{\alpha_s}{\alpha_t}(x_t - \sigma_t c\epsilon(x_t, t; \theta)))^2] \\
&= -\nabla_\theta L_{reinforcement}
\end{aligned}$$

Gradient of log likelihood:

$$
\begin{aligned}
\nabla_\theta \log p_\theta(x_s|x_t) &= \nabla_\theta \log(\mathcal{N}(x_s; \mu_t(x_t, t; \theta), \sigma_Q^2)) \\
&= \nabla_\theta \log(\frac{1}{\sigma_Q\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{x_s - \mu_t(x_t, t; \theta)}{\sigma_Q})^2)) \\
&= \nabla_\theta(-\frac{1}{2}(\frac{x_s - \mu_t(x_t, t; \theta)}{\sigma_Q})^2) \\
&= -\frac{1}{2\sigma_Q^2}\nabla_\theta(x_s - \mu_t(x_t, t; \theta))^2 \\
&= -\frac{1}{2c\sigma_s^2}\nabla_\theta(x_s - \frac{\alpha_s}{\alpha_t}(x_t - \sigma_t c\epsilon(x_t, t; \theta)))^2
\end{aligned}
$$

4 Extra

$$
\begin{aligned}
&\alpha_{t|s}^2 = \frac{\sigma_t^2 - \sigma_{t|s}^2}{\sigma_s^2} = \frac{\alpha_t^2}{\alpha_s^2} \\
&\sigma_Q^2 = \sigma_s^2(-(e^{\gamma(s)-\gamma(t)} - 1)) \\
&c = -(e^{\gamma(s)-\gamma(t)} - 1) = \frac{\sigma_{t|s}^2}{\sigma_t^2} \\
&\mu_t = \frac{(1-c)\alpha_s}{\alpha_t}x_t + c\alpha_s\hat{x}_\theta \\
&L_{RL} = \mathbb{E}_{p_\theta(\tau)}[R(\hat{x}_0)\sum_t^T \frac{1}{2\sigma_Q^2(t)}(x_{t-1} - \mu_t(x_t; \theta))^2] \\
&L_{Diff} = \sum_t^T \mathbb{E}_{q(x_t|x_0)}[\frac{1}{2\sigma_Q^2(t)}(\mu_q(x_t, x_0) - \mu_t(x_t; \theta))^2]
\end{aligned}
$$