

Генеративные модели для структуры и взаимодействия протеинов

Выполнил: Клыпа Р.С.

Научный руководитель: Грудинин С.В.

МФТИ

16 декабря 2023 г.

Цели исследования

Цель работы

Предсказание структуры молекулярных соединений (например, РНК) при взаимодействии с протеином - $P(\text{geom}_{RNA} | \text{geom}_{pr}, \text{seq}_{pr}, \text{seq}_{RNA})$.

Проблема

Существующие решения обладают рядом технических и теоретических недостатков.

Метод решения

Дизайн архитектуры модели и диффузионного процесса с учетом биологических и физико-химических особенностей задачи.

Постановка задачи

Дано

Полная структура и последовательность протеина, последовательность РНК, потенциально известен интерфейс взаимодействия.

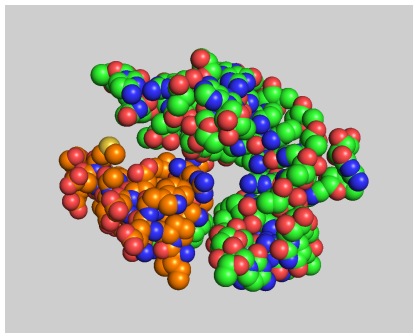
Варианты постановки задачи

- Inpaint task - протеин и РНК считаются 'равноправными', диффузируется только РНК. Модель идентично взаимодействует с обеими компонентами.
- Условная диффузия - информация о протеине используется как дополнительное условие при генерировании структуры РНК.

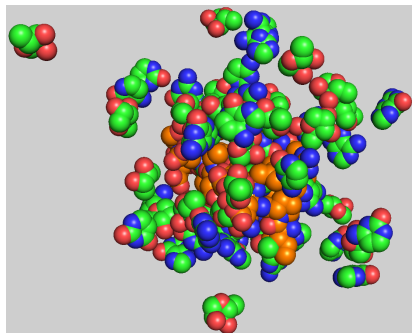
Общая цель

Предсказать структуру РНК, максимально близкую к экспериментальной. Измерение близости является одной из подзадач.

Диффузия в $SE(3)$



РНК-протеиновый комплекс.

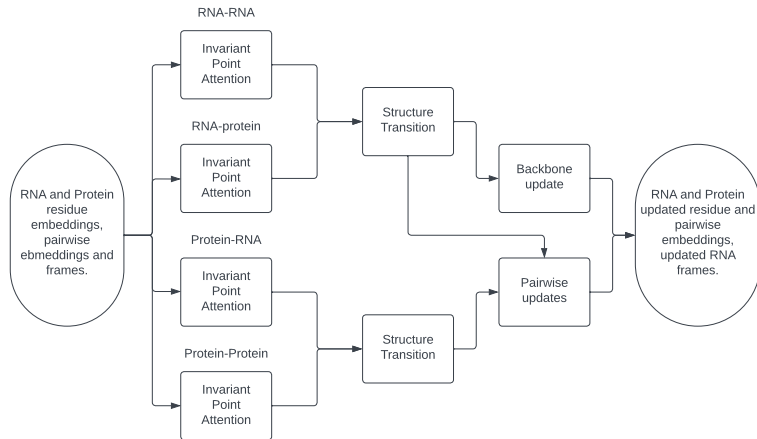


Комплекс с зашумленной РНК.

Была исследована сходимость в зависимости от точности модели для DDPM в $R(3)$, планируется повторить для RSGM в $R(3)$ и $SO(3)$.

<https://arxiv.org/pdf/2202.02763.pdf>, <https://arxiv.org/pdf/2302.02277.pdf>

Архитектура модели



<https://www.nature.com/articles/s41586-021-03819-2>,

<https://www.biorxiv.org/content/10.1101/2022.12.21.521526v1>

- Были проведены эксперименты с inpaint и условной диффузией, наилучший результат был получен для условной.
- Несмотря на ошибку, превышающую стандартные требования отрасли, модель проявляет выдающуюся адекватность и интерпретируемость в получаемых результатах. Это подчеркивает успешное обучение модели, ее способность осваивать информацию и эффективно усваивать ключевые концепции.
- Планируется усовершенствование модели и дополнительные эксперименты.