

**ЭВОЛЮЦИОННАЯ ГЕНЕРАЦИЯ ДАННЫХ ДЛЯ  
ОБУЧЕНИЯ МОДЕЛЕЙ РЕВЕРС-ИНЖИНИРИНГА  
ТРЕХМЕРНЫХ ОБЪЕКТОВ**

***Иванов Григорий Артемович***

*студент магистратуры*

*Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия*

*E-mail: grivind@yandex.ru*

***Научный руководитель — Конушин Антон Сергеевич***

Реверс-инжиниринг трёхмерных объектов — это восстановление параметрической САПР-модели по её геометрическому представлению (изображению, облаку точек или mesh). В отличие от простого воспроизведения формы, задача заключается в получении исполняемого кода, описывающего последовательность операций построения и параметров модели. Такой подход позволяет редактировать объект, изменять его размеры и структуру, повторно использовать конструктивные решения и интегрировать результат в инженерные пайплайны. В работе рассматривается обучение нейросети, восстанавливающей исполняемые скрипты на базе CadQuery по геометрическим данным [1,2].

Ключевая проблема — дефицит доступных обучающих данных. Существуют крупные корпуса трёхмерных моделей, такие как ShapeNet и ABC [3,4], однако они содержат только конечную геометрию без истории построения. Наборы с программным описанием моделей значительно менее распространены и охватывают ограниченный спектр сценариев моделирования (например, Fusion 360 Gallery [5]). Формирование большого числа корректных параметрических скриптов вручную требует существенных временных и экспертных ресурсов, что затрудняет масштабирование обучения.

Для решения этой задачи применяется эволюционный алгоритм синтетической генерации данных. Исходным набором служат 46 вручную написанных параметрических генераторов простых деталей. Каждый генератор представляет собой функцию с параметрами, возвращающую корректную САПР-модель. Далее выполняется итеративный цикл «модификация – проверка – отбор». На этапе модификации нейросетевая модель преобразует существующий код: добавляет новые операции построения (extrude, revolve, boolean, fillet и др.), вводит дополнительные параметры, изменяет порядок операций или комбинирует фрагменты различных генераторов.

Каждая новая версия программы проходит многоуровневую ва-

лидацию. Сначала выполняется синтаксическая проверка и компиляция кода. Затем осуществляется геометрическая проверка корректности построения (отсутствие самопересечений, топологическая целостность). Дополнительно контролируется устойчивость генератора при варьировании параметров, чтобы модель оставалась корректной в области параметрического пространства. Невалидные варианты отбрасываются или автоматически корректируются. Отобранные программы наследуют структуру родительских генераторов, что обеспечивает постепенное усложнение конструктивной истории и увеличение разнообразия форм.

После этапа эволюции применяется параметрическая выборка с оптимизацией различий (в частности, методом CMA-ES), позволяющая получать десятки геометрически отличающихся экземпляров от каждого генератора. Дополнительно используются каноникализация кода, нормализация числовых значений и семантически эквивалентные переписывания программ, что создаёт альтернативные представления одной и той же конструкции и увеличивает вариативность обучающих примеров.

Масштабирование носит мультипликативный характер. Эволюционный этап расширяет исходные 46 генераторов до нескольких тысяч валидных параметрических программ. Параметрическая выборка обеспечивает увеличение примерно в 10–20 раз, а кодовая аугментация — ещё в 2–3 раза. В результате формируется корпус из 744 780 исполняемых САПР-скриптов, что более чем в 16 000 раз превышает исходный набор. Такой объём синтетических данных создаёт основу для обучения устойчивых моделей реверс-инжиниринга трёхмерных объектов.

### Литература

1. Rukhovich D. et al. CAD-Recode: Reverse engineering CAD code from point clouds. arXiv, 2024.
2. Kolodiaznyi M. et al. cadrille: Multi-modal CAD reconstruction with reinforcement learning. arXiv, 2025.
3. Chang A. X. et al. ShapeNet: An information-rich 3D model repository. 2015.
4. Koch S. et al. ABC: A big CAD model dataset for geometric deep learning. CVPR, 2019.
5. Willis K. D. D. et al. Fusion 360 Gallery: A dataset and environment for programmatic CAD construction. ACM TOG, 2021.