NEAT: Neighborhood-Guided, Efficient, Autoregressive Set Transformer for 3D Molecular Generation

Daniel Rose[1,2,3 *], Roxane Axel Jacob[1,2,3 *], Johannes Kirchmair[1,2], Thierry Langer[1,2]

[1]*Department of Pharmaceutical Sciences, Faculty of Life Sciences, University of Vienna, Josef-Holaubek-Platz 2, 1090 Vienna, Austria*

[2]*Christian Doppler Laboratory for Molecular Informatics in the Biosciences, Department of Pharmaceutical Sciences, University of Vienna, Josef-Holaubek-Platz 2, 1090 Vienna, Austria*

[3]*Vienna Doctoral School of Pharmaceutical, Nutritional and Sport Sciences, University of Vienna, Josef-Holaubek-Platz 2, 1090 Vienna, Austria*

[*] *Equal contribution*

The *de novo* generation of molecules with desired biological and physicochemical properties is emerging as a promising alternative to virtual screening approaches. Current state-of-the-art 3D molecular generation models, which are predominantly based on diffusion or flow-matching architectures [1], have demonstrated strong performance. However, their inference runtime remains a significant bottleneck, and they impose limitations on the number of atoms in the generated designs. Transformer-based autoregressive models present a compelling alternative: they are computationally efficient and they allow for flexible atom counts in the generated molecules [2]. However, a key challenge in applying transformers to 3D molecular generation lies in their reliance on sequential token ordering. While such ordering is naturally defined for words in a sentence or pixels in an image, it is not uniquely defined for atoms in a molecule, as there is no universally correct way to number atoms in a molecular graph. Previous approaches have addressed this issue by employing canonical atom orderings [2, 3], but this strategy fails to ensure atom permutation invariance and introduces data-representation artifacts. These limitations become particularly evident in tasks such as prefix completion, where models are tasked with building full molecules around a given scaffold. In such cases, the canonical ordering of atoms in the scaffold often conflicts with their ordering in the final molecule, leading to invalid designs generated by existing transformer-based models. In this talk, we will discuss contemporary 3D molecular generation paradigms and their trade-offs. We will highlight novel developments from the computer vision domain for training performant autoregressive image generators. We will then show how to apply these techniques to the 3D molecular domain and introduce our new model NEAT: a Neighborhood-guided, Efficient, Autoregressive Set Transformer [4]. NEAT treats molecular graphs as sets of atoms and learns an order-agnostic distribution of admissible tokens at the graph boundary, ensuring atom-level permutation invariance by design. It achieves efficient, state-of-the-art performance in 3D molecular generation, demonstrating its ability to generate chemically valid and diverse molecules while overcoming the limitations of canonical ordering.

Bibliography:

[1] I. Dunn, D. R. Koes. FlowMol3: Flow Matching for 3D de novo Small-Molecule Generation. *Digit. Discov.* **2026** doi: 10.1039/d5dd00363f.
[2] A. H. Cheng, C. Sun, A. Aspuru-Guzik. Scalable Autoregressive 3D Molecule Generation. *arXiv* **2025** 2505.13791.
[3] H. Li, W. Du, Y. Li, H. Guo, S. Liu. InertialAR: Autoregressieve 3D Molecule Generation with Inertial Frames. *arXiv* **2025** 2510.27497.
[4] D. Rose, R. A. Jacob, J. Kirchmair, T. Langer. NEAT: Neighborhood-Guided, Efficient, Autoregressive Set Transformer for 3D Molecular Generation. *arXiv* **2026** 2512.05844.