
Batch-Advantage Transformer with Hypergraph Optimized Grammar (BAT/HOG)

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 We present a novel approach to the issue of molecular optimization. Our approach
2 uses a hypergraph replacement grammar inferred from the ZINC database, with
3 grammar construction optimized for molecular structure creation. We treat the
4 optimization as a reinforcement learning problem, using a batch-advantage mod-
5 ification of the policy gradient algorithm - using individual rewards minus the
6 batch average reward to weight the log probability loss. The reinforcement learn-
7 ing agent is tasked with building molecules using this grammar, with the goal of
8 maximizing benchmark scores available from the literature. To do so, the agent
9 has policies both to choose the next node in the graph to expand and to select the
10 next grammar rule to apply. The policies are implemented using the Transformer
11 architecture with the partially expanded graph as the input. We achieve state of
12 the art performance on common benchmarks from the literature, such as penalized
13 logP and QED, with only hundreds of steps (without pre-training) on a budget
14 GPU instance. Competitive performance is obtained on more advanced GuacaMol
15 v2 goal-oriented benchmarks. Coupled with a Transformer based discriminator,
16 the model achieves competitive results on the GuacaMol distribution benchmarks;
17 training is stable over a range of hyperparameter values.

18	1 Introduction
19	2 Generating guaranteed valid SMILES strings
20	2.1 Context-free grammar
21	2.2 Respecting valences
22	2.2.1 Implicit hydrogens
23	2.3 Cycles
24	2.3.1 Numeral assignment
25	2.3.2 Cycle size
26	2.3.3 Cycle chaining
27	2.3.4 Aromatic cycles
28	2.4 Extraction of Hypergraph Cliques
29	2.5 Rule-pair Encoding
30	2.6 Making sure the expansions terminate
31	2.7 Limits of context-free grammars
32	2.8 Grammar conciseness and expressiveness
33	3 Model choice
34	3.1 Reinforcement learning
35	3.2 Architecture
36	3.3 Training
37	3.4 Optimization and the reward function
38	4 Results
39	4.1 GuacaMol Benchmarks
40	4.2 Ablation Studies
41	Acknowledgments
42	References

- [1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.
- [2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural Simulation System*. New York: TELOS/Springer–Verlag.
- [3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.