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

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

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

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

1 Introduction

2 Generating guaranteed valid SMILES strings

2.1 Context-free grammar

2.2 Respecting valences

2.2.1 Implicit hydrogens

2.3 Cycles

2.3.1 Numeral assignment

2.3.2 Cycle size

2.3.3 Cycle chaining

2.3.4 Aromatic cycles

2.4 Extraction of Hypergraph Cliques

2.5 Rule-pair Encoding

2.6 Making sure the expansions terminate

2.7 Limits of context-free grammars

2.8 Grammar conciseness and expressiveness

3 Model choice

3.1 Reinforcement learning

3.2 Architecture

3.3 Training

3.4 Optimization and the reward function

4 Results

4.1 GuacaMol Benchmarks

4.2 Ablation Studies

Acknowledgments

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.