Graph Convolutional Policy Network (GCPN) for Goal-Directed Molecular Graph Generation

PRESENTED BY: BERNARD MOUSSAD

UNIVERSITY LIBRARIES
VIRGINIA TECH

The Problem

- *Drug discovery and material science are based on the principles of designing molecular structures with specific desired properties.*
- *The space is massive though (10^23 - 10^60 drug-like molecules)*
- *Chemical space is discrete and molecular properties are highly sensitive to small changes in molecular structure.*

Progress in Molecular Design

- Molecule generation has been achieved via DL models
- Achieving objectives of chemical and biological properties has been difficult because these are highly complex and non-differentiable goals.

The Solution: Graph Convolutional Policy Network

- Combines 3 machine learning concepts:
  - Graph Representation: Used to obtain vector representaiton of the state of generated graphs
  - Reinforcement Learning: Trains the model end-to-end
  - Adversarial Training: adversarial loss used as reward to incorporate prior knowledge specified by dataset of example molecules

Graph Representation

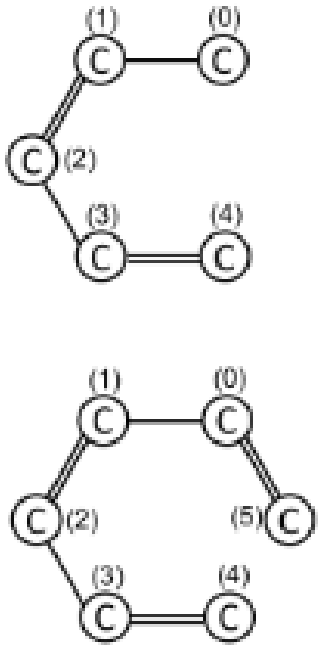
- Benefits over SMILE representation
  - Partial Graphs can be seen as substructures, but partial text-representations are not meaningful
  - If one character is changed in SMILE, it can change the entire structure or invalidate it.

Graph Representation

- The graphs are represented as (A,E,F)
  - $A \in \{0, 1\}^{n \times n}$ – Adjacency Matrix
  - $E \in \{0, 1\}^{b \times n \times n}$ - Edge-conditioned adjacency tensor where there are *b* possible edge types
    - $E_{i,j,k} = 1$ if
      there exists an edge of type i between nodes j and k
  - $F \in R^{n \times d}$ – Node Feature matrix where each node has *d* features

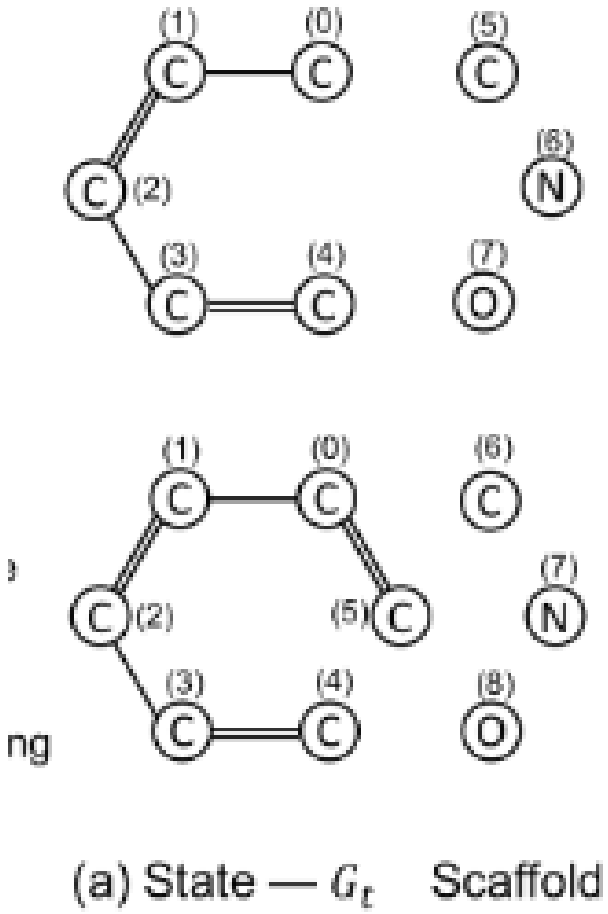
Graph Generation: State Space

- An intermediate graph G_t at time step *t*
- Initially starts as a single node that represents a carbon atom



Graph Generation: Action Space

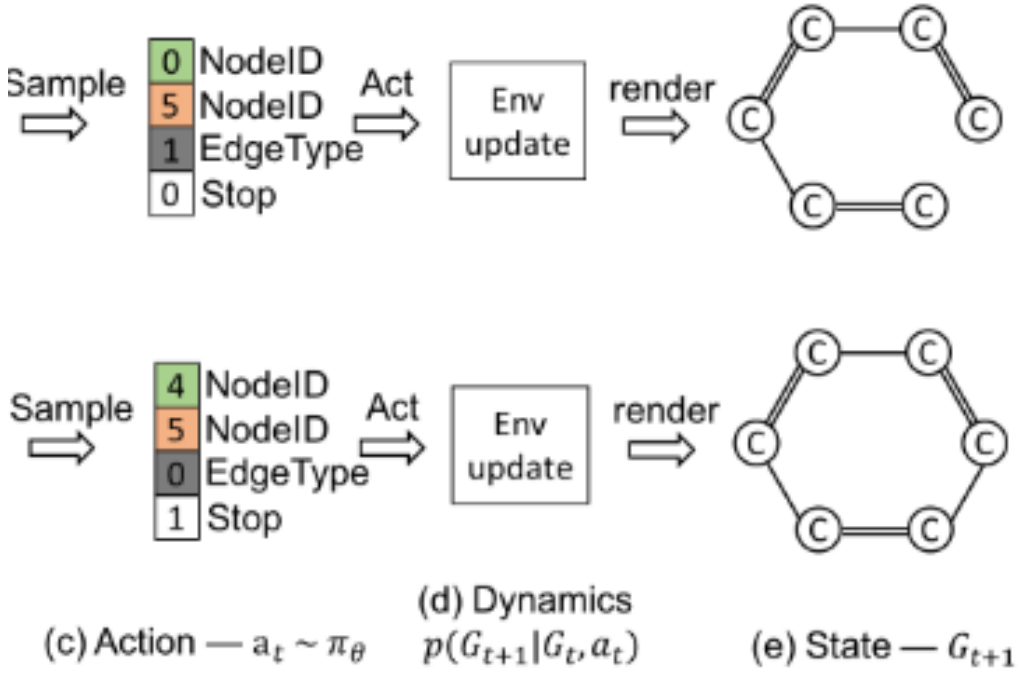
- Consists of a set of scaffold subgraphs $\{C_1, \ldots, C_s\}$
- An action can either consist of connecting a scaffold to the current graph G_t or connecting nodes already within the graph.
- A scaffold generally has single node representations of all the possible atoms desired for a molecule.



(a) State — G_t Scaffold

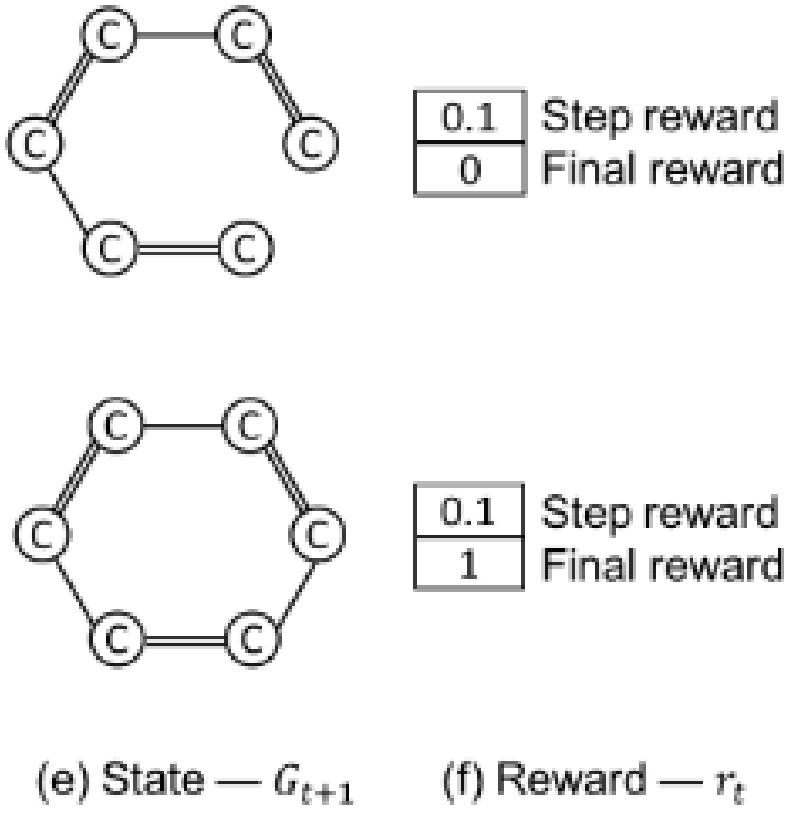
Graph Generation: State Transition Dynamics

- This controls whether an action can be considered valid given a set of rules
- If an action results in an invalid state, that action is rejected and the state remains unchanged



Graph Generation: Reward

- Domain Specific rewards
  - Final property scores
    - octanol-water partition coefficient (logP), druglikeness (QED)
  - Unrealistic Molecule Penalties
- Intermediate Rewards
  - Step-wise validity rewards
  - Adversarial rewards (via Generative Adversarial Network)



(e) State — G_{t+1} (f) Reward — r_t

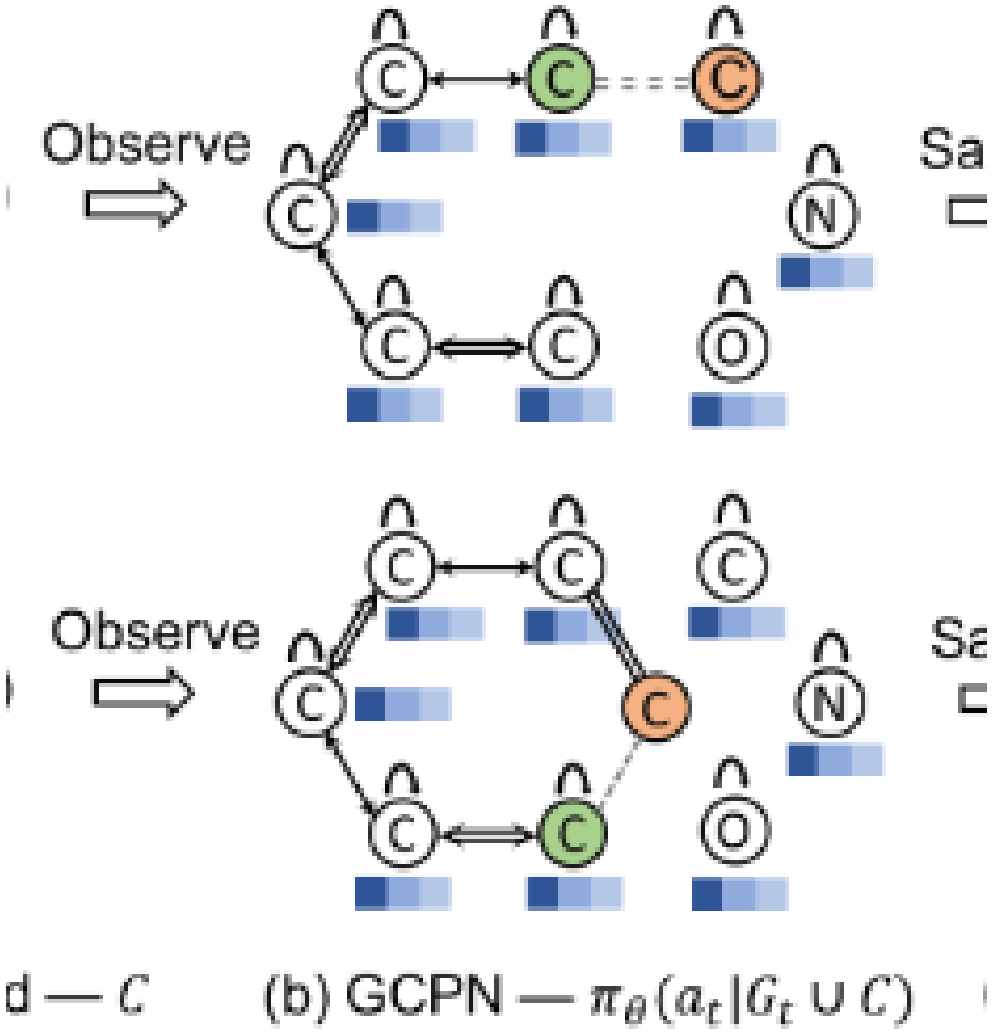
Graph Convolutional Policy Network: Link Prediction

Node embeddings are computed via Graph Convolutional Networks using information from each edge type over *L* layers

- These are aggregated at the l^{th} layer to compute the next layer node embedding.

Actions are then predicted by selection of two nodes and prediction if there will be an edge or termination

- Multilayer Perceptrons (MLP) are used to achieve this



Policy Gradient Training

- Allows for optimization of policy networks
- Specifically adopted Proximal Policy Optimization (PPO)
  - PPO computes an update at each step that minimizes the cost function while ensuring the deviation from the previous policy is relatively small
- Network can be pretrained on known molecules
  - A sampling of a graph *G* and randomly selecting one of its subgraphs *G`* results in a state s_t
  - An action a_t can be defined as the addition of any atom or bond to the G` thus resulting in the pair (s_t, a_t)

Experiments: Evaluation Tasks

- Three tasks were used for evaluation of GCPN against state-of-the-art molecule generation algorithms
  - Property Optimization - Generate novel molecules whose specified molecular properties are optimized
  - Property Targeting – Generate novel molecules whose specified molecular properties are as close to the target scores as possible
  - Constrained Property Optimization - Generate novel molecules whose specified molecular properties are optimized, while also containing a specified molecular substructure

Experiments: Evaluation Tasks

- Three tasks were used for evaluation of GCPN against state-of-the-art molecule generation algorithms
  - Property Optimization - Generate novel molecules whose specified molecular properties are optimized
  - Property Targeting – Generate novel molecules whose specified molecular properties are as close to the target scores as possible
  - Constrained Property Optimization - Generate novel molecules whose specified molecular properties are optimized, while also containing a specified molecular substructure

Experiments: Setup

- Dataset: ZINC250K which contains 250,000 drug like commercially available molecules
- Molecule Environment: OpenAI Gym using RDKit and adapted to ZINC250k (Max atom # = 38, 9 atom types, 3 edge types)
- Rewards
  - [−4, 4] for final chemical property reward
  - [−2, 2] for final chemical filter reward
  - [−1, 1] for final adversarial reward
  - [−1, 1] for intermediate adversarial reward
  - [−1, 1] for intermediate validity reward

Experiments: Baselines

- Junction Tree VAE (JT-VAE) – Combines graph representation and a VAE framework
- ORGAN – RL-based molecule generation algorithm using a text-based representation of molecules

Results

Table 1: Comparison of the top 3 property scores of generated molecules found by each model.

Method	Penalized logP				QED			
	1st	2nd	3rd	Validity	1st	2nd	3rd	Validity
ZINC	4.52	4.30	4.23	100.0%	0.948	0.948	0.948	100.0%
Hill Climbing	–	–	–	–	0.838	0.814	0.814	100.0%
ORGAN	3.63	3.49	3.44	0.4%	0.896	0.824	0.820	2.2%
JT-VAE	5.30	4.93	4.49	100.0%	0.925	0.911	0.910	100.0%
GCPN	**7.98**	**7.85**	**7.80**	**100.0%**	**0.948**	**0.947**	**0.946**	**100.0%**

Results

Table 2: Comparison of the effectiveness of property targeting task.

Method	$-2.5 \leq \text{logP} \leq -2$		$5 \leq \text{logP} \leq 5.5$		$150 \leq \text{MW} \leq 200$		$500 \leq \text{MW} \leq 550$	
	Success	Diversity	Success	Diversity	Success	Diversity	Success	Diversity
ZINC	0.3%	0.919	1.3%	0.909	1.7%	0.938	0	–
JT-VAE	11.3%	**0.846**	7.6%	0.907	0.7%	0.824	16.0%	0.898
ORGAN	0	–	0.2%	**0.909**	15.1%	0.759	0.1%	0.907
GCPN	**85.5%**	0.392	**54.7%**	0.855	**76.1%**	**0.921**	**74.1%**	**0.920**

The diversity of a set of molecules is defined as the average pairwise Tanimoto distance between the Morgan fingerprints of the molecules

Table 3: Comparison of the performance in the constrained optimization task.

δ	JT-VAE			GCPN		
	Improvement	Similarity	Success	Improvement	Similarity	Success
0.0	1.91 ± 2.04	0.28 ± 0.15	97.5%	$\mathbf{4.20 \pm 1.28}$	$\mathbf{0.32 \pm 0.12}$	**100.0%**
0.2	1.68 ± 1.85	0.33 ± 0.13	97.1%	$\mathbf{4.12 \pm 1.19}$	$\mathbf{0.34 \pm 0.11}$	**100.0%**
0.4	0.84 ± 1.45	$\mathbf{0.51 \pm 0.10}$	83.6%	$\mathbf{2.49 \pm 1.30}$	0.47 ± 0.08	**100.0%**
0.6	0.21 ± 0.71	0.69 ± 0.06	46.4%	$\mathbf{0.79 \pm 0.63}$	$\mathbf{0.68 \pm 0.08}$	**100.0%**