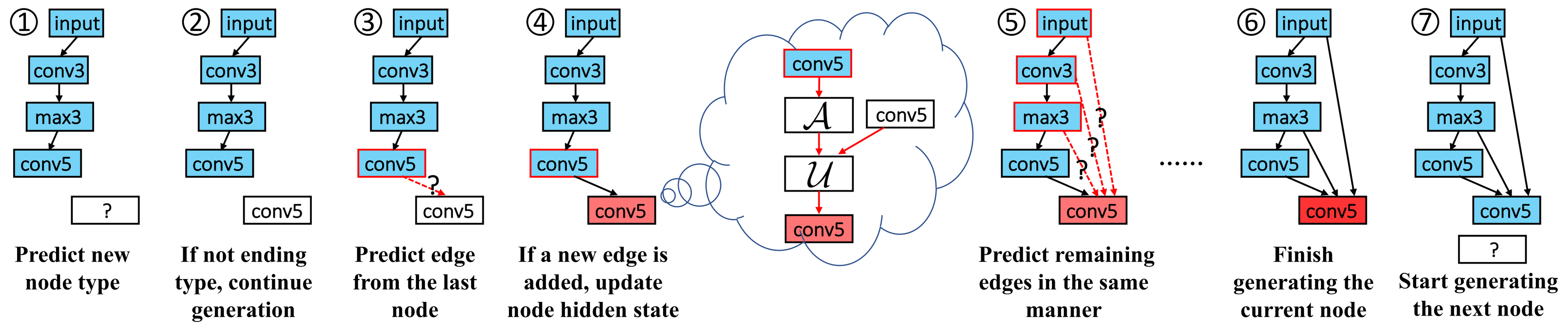
Directed Graph Grammars for Sequence-based Learning

Michael Sun[1*], Orion Foo[2], Gang Liu[3], Wojciech Matusik[1], Jie Chen[4]

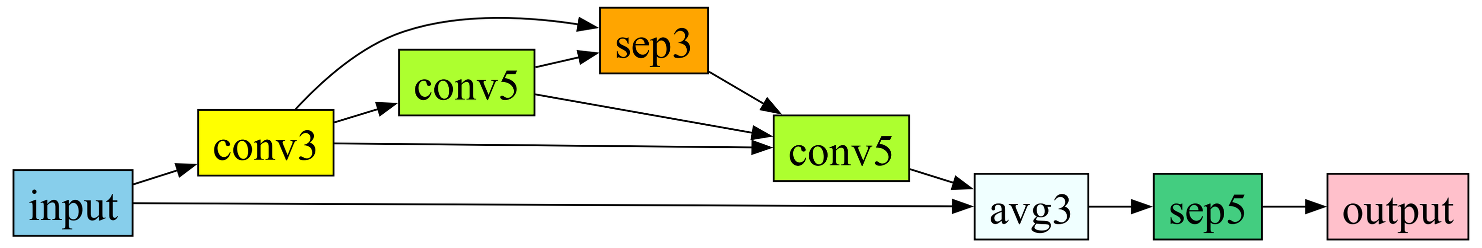
[1]MIT CSAIL, [2]MIT, [3]Notre Dame, [4]MIT-IBM Watson AI Lab, [*]Website: michaelsun.tech

Existing sequential graph generation methods lack a principled mapping between graph and sequence.

Autoregressive (AG) methods generate node-by-node. Figure from Zhang et al (2019).



Sequential Decoding (SD) methods generate an equivalent sequential description of a graph.

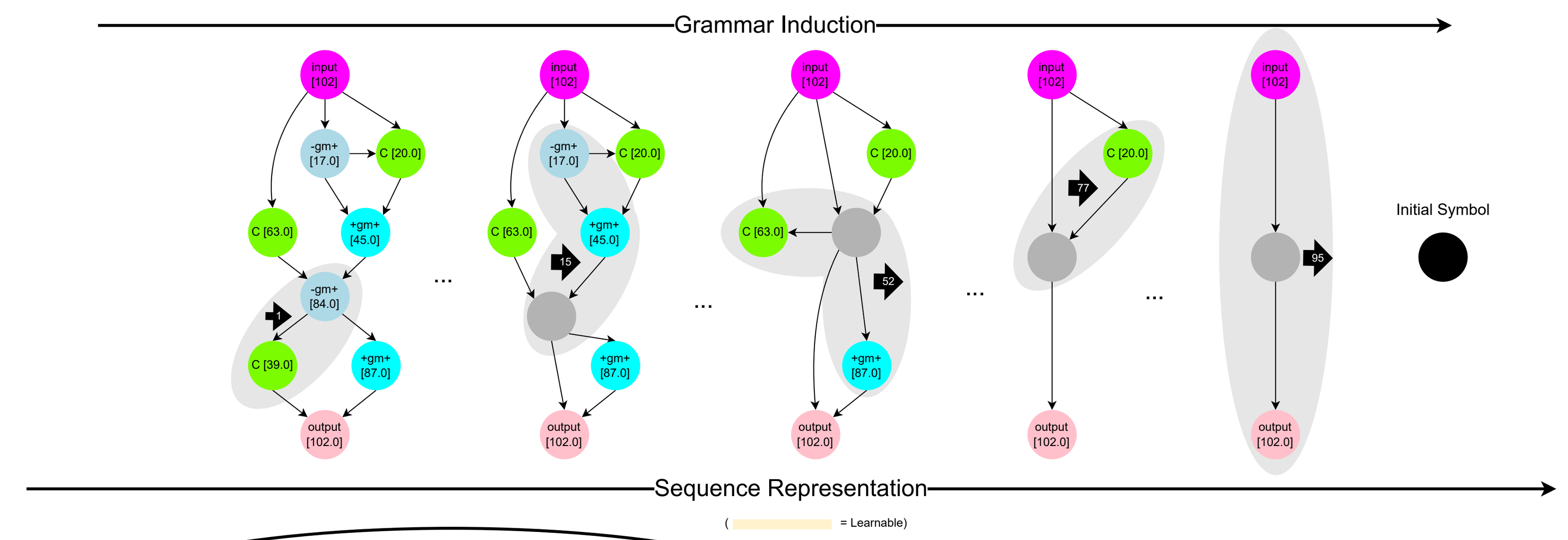


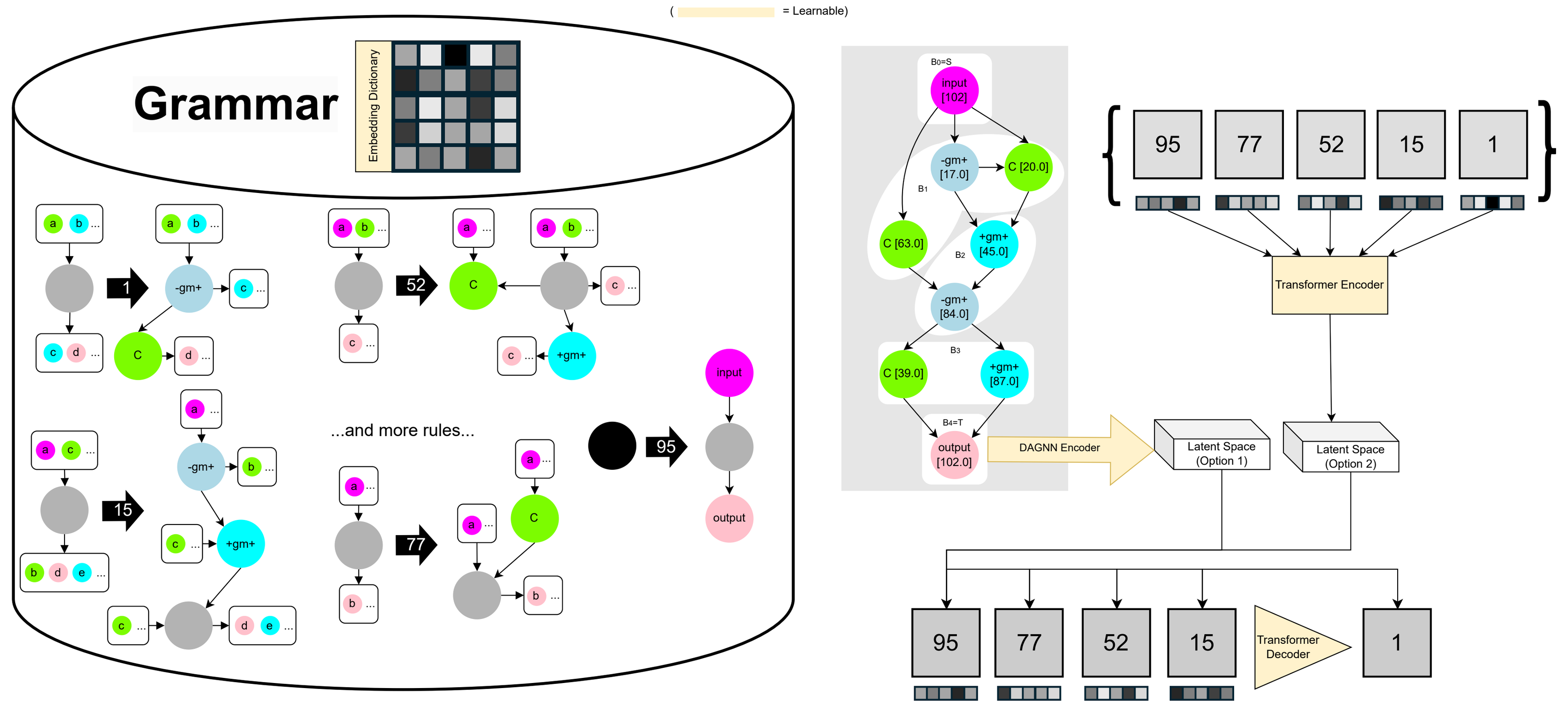
Sequential description: [0] + [2,0] + [1,0,1] + [2,0,1,1] + [4,1,0,0,0] + [3,0,0,0,0,0].
Format: Concat [node type, *skip connections] over all nodes.
Node types: [conv3, sep3, conv5, sep5, avg3].

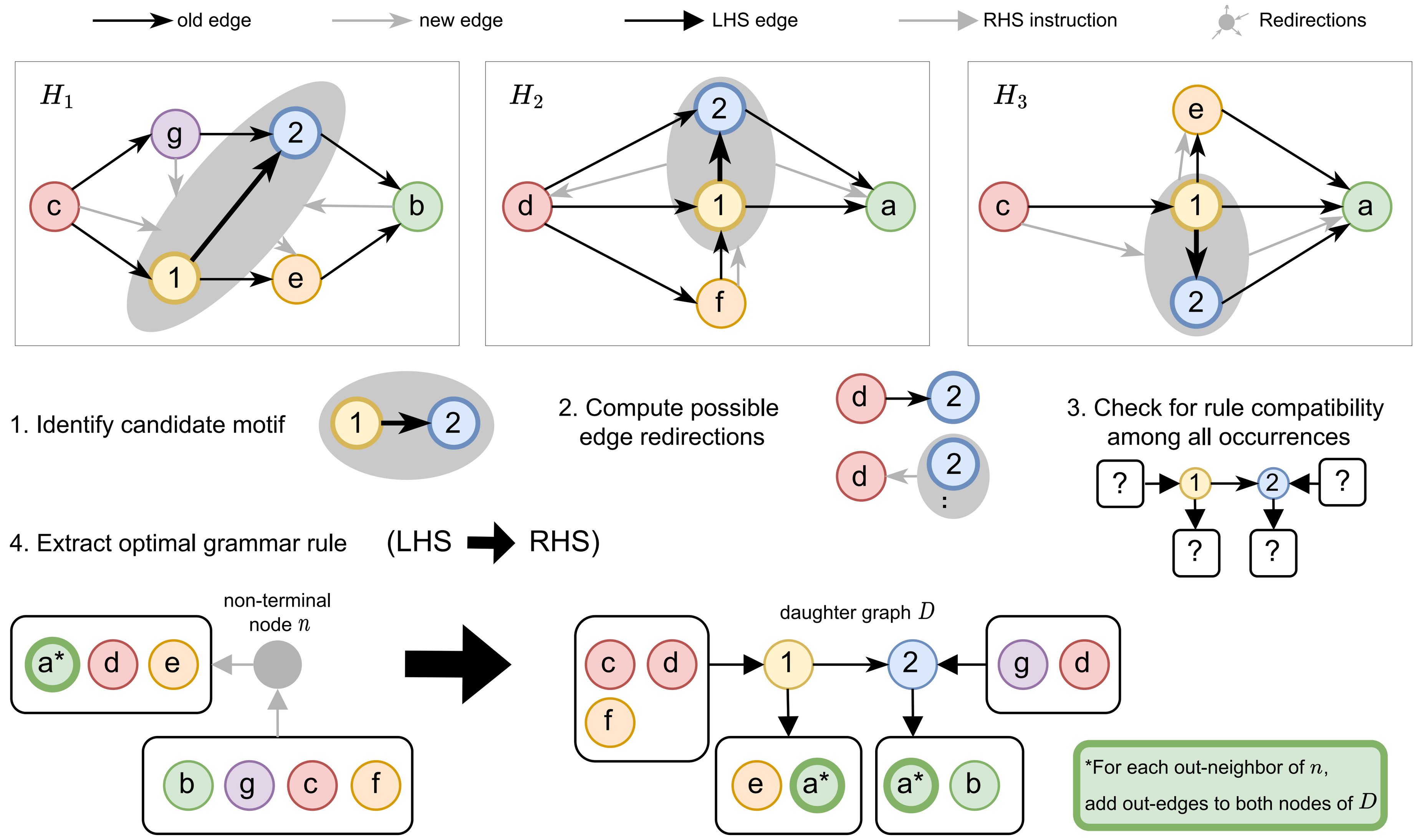
Methods	One-to-one?	Onto?	Deterministic?	Valid?	Stateless?
AG	✗	✓	✗	✓	✗
SD	✗	✓	✓	✗	✓
DIGGED	✓	✓	✓	✓	✓

One-to-one (dataset D):
DIGGED's disambiguation procedure assures there is exactly one parse per graph in the dataset.
AG methods can generate in (up to exponential) ways.
SD methods rely on an arbitrary ordering of the nodes.
Onto (dataset D):
DIGGED's grammar induction is also a parsing algorithm, so every graph in the dataset has a parse.
AG and SD can always generate the dataset.
Deterministic:
DIGGED's grammar is deterministic.
SD enforces teacher-forcing.
AG methods can take many action trajectories to arrive at the same final state.
Valid:
DIGGED uses context-free grammar (CFG) , so an arbitrary derivation still produces a valid graph.
AG can ensure only actions that retain validity are taken.
SD methods do not guarantee an arbitrary description encodes a valid graph.
Stateless:
CFGs are stateless. Generation terminates when a terminal rule is selected.
SD is an equivalent encoding.
AG methods require tracking the intermediate graph as a state to filter out invalid actions.

Graph Grammar is a formal framework for describing graphs as a sequence of rewrite rules.







Results show DIGGED's superiority on downstream generation, prediction and optimization tasks.
Unconditional Generation – How well we can generate samples by decoding from a Gaussian prior.
Predictive Performance – How well the latent embeddings of DAGs can predict their performances.
Bayesian Optimization – How well the learned latent space can be used for searching for high-performance DAGs through BO.

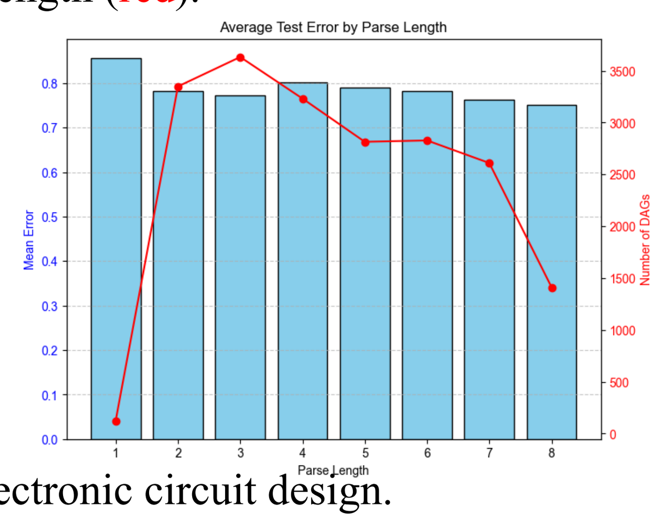
(Left) Prior validity, uniqueness and novelty (%). (Right) Predictive performance of latent representation on ENAS & BN test set.

Methods	Neural architectures				Bayesian networks			
	Accuracy	Validity	Uniqueness	Novelty	Accuracy	Validity	Uniqueness	Novelty
D-VAE	99.96	**100.00**	37.26	**100.00**	99.94	98.84	38.98	98.01
S-VAE	99.98	**100.00**	37.03	99.99	99.99	**100.00**	35.51	99.70
GraphRNN	99.85	99.84	29.77	**100.00**	96.71	**100.00**	27.30	98.57
GCN	98.70	99.53	34.00	**100.00**	99.81	99.02	32.84	99.40
DeepGMG	94.98	98.66	46.37	99.93	47.74	98.86	57.27	98.49
DIGGED (GNN)	**100**	**100**	**98.7**	99.9	**100**	**100**	97.6	**100**
DIGGED (TOKEN)	**100**	**100**	25.4	37.8	**100**	**100**	**98.67**	26.67

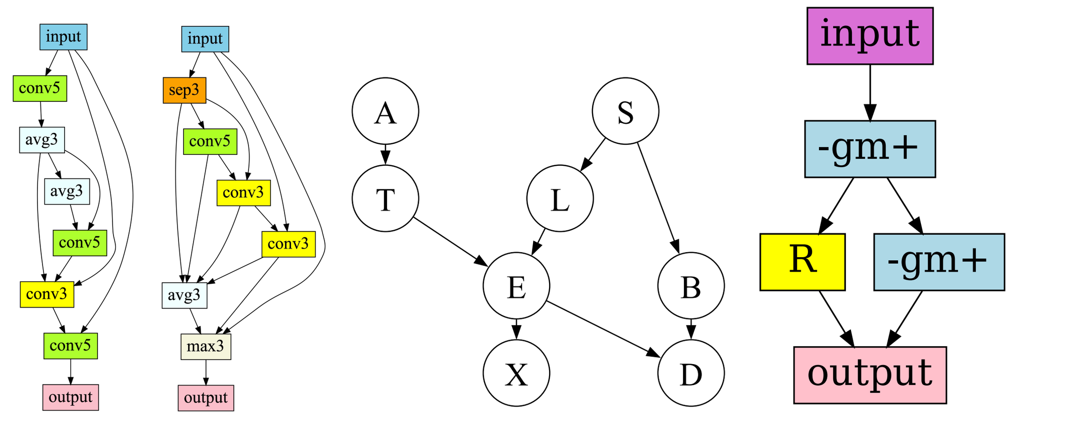
Model	ENAS		BN	
	RMSE ↓	Pearson's r ↑	RMSE ↓	Pearson's r ↑
S-VAE	**0.644 ± 0.003**	**0.762 ± 0.002**	0.896 ± 0.003	0.442 ± 0.001
GraphRNN	0.695 ± 0.002	0.707 ± 0.001	**0.881 ± 0.002**	0.453 ± 0.001
GCN	0.681 ± 0.003	0.739 ± 0.001	0.914 ± 0.002	0.394 ± 0.001
DeepGMG	0.976 ± 0.003	0.140 ± 0.002	0.970 ± 0.003	0.236 ± 0.001
D-VAE	0.890 ± 0.003	0.352 ± 0.001	0.926 ± 0.003	0.251 ± 0.001
DAGNN	0.882 ± 0.004	0.433 ± 0.001	0.933 ± 0.003	0.247 ± 0.001
DIGGED (GNN)	0.912 ± 0.001	0.386 ± 0.001	0.953 ± 0.052	**0.712 ± 0.013**
DIGGED (TOKEN)	0.987 ± 0.001	0.049 ± 0.006	0.989 ± 0.0001	0.129 ± 0.002

(Left) Predictive Performance of Latent Representations on CktBench101. (Right) Test error distribution across the parse length (blue). For reference, we also include a count of the number of test set examples of each parse length (red).

Evaluation Metric	Gain		BW		PM		FoM	
	RMSE ↓	Pearson's r ↑	RMSE ↓	Pearson's r ↑	RMSE ↓	Pearson's r ↑	RMSE ↓	Pearson's r ↑
PACE	0.644 ± 0.003	0.762 ± 0.002	0.896 ± 0.003	0.442 ± 0.001	0.970 ± 0.003	0.226 ± 0.001	0.889 ± 0.003	0.423 ± 0.001
DAGNN	0.695 ± 0.002	0.707 ± 0.001	0.881 ± 0.002	0.453 ± 0.001	0.969 ± 0.003	0.231 ± 0.002	0.877 ± 0.003	0.442 ± 0.001
D-VAE	0.681 ± 0.003	0.739 ± 0.001	0.914 ± 0.002	0.394 ± 0.001	**0.956 ± 0.003**	0.301 ± 0.002	0.897 ± 0.003	0.374 ± 0.001
GCN	0.976 ± 0.003	0.140 ± 0.002	0.970 ± 0.003	0.236 ± 0.001	0.993 ± 0.002	0.171 ± 0.001	0.974 ± 0.003	0.217 ± 0.001
GIN	0.890 ± 0.003	0.352 ± 0.001	0.926 ± 0.003	0.251 ± 0.001	0.985 ± 0.004	0.187 ± 0.002	0.910 ± 0.003	0.284 ± 0.001
NGNN	0.882 ± 0.004	0.433 ± 0.001	0.933 ± 0.003	0.247 ± 0.001	0.984 ± 0.004	0.196 ± 0.002	0.926 ± 0.002	0.267 ± 0.001
Pathformer	0.816 ± 0.003	0.529 ± 0.001	0.895 ± 0.003	0.410 ± 0.001	0.967 ± 0.002	0.297 ± 0.001	0.887 ± 0.002	0.391 ± 0.001
CktGNN	**0.607 ± 0.003**	**0.791 ± 0.001**	0.873 ± 0.003	0.479 ± 0.001	0.973 ± 0.002	0.217 ± 0.001	0.854 ± 0.003	0.491 ± 0.002
DIGGED (GNN)	0.630 ± 0.005	0.771 ± 0.004	**0.635 ± 0.006**	**0.784 ± 0.001**	0.990 ± 0.001	**0.314 ± 0.001**	**0.627 ± 0.002**	**0.787 ± 0.001**
DIGGED (TOKEN)	—	—	—	—	—	—	1.005 ± 0.0002	0.199 ± 0.001



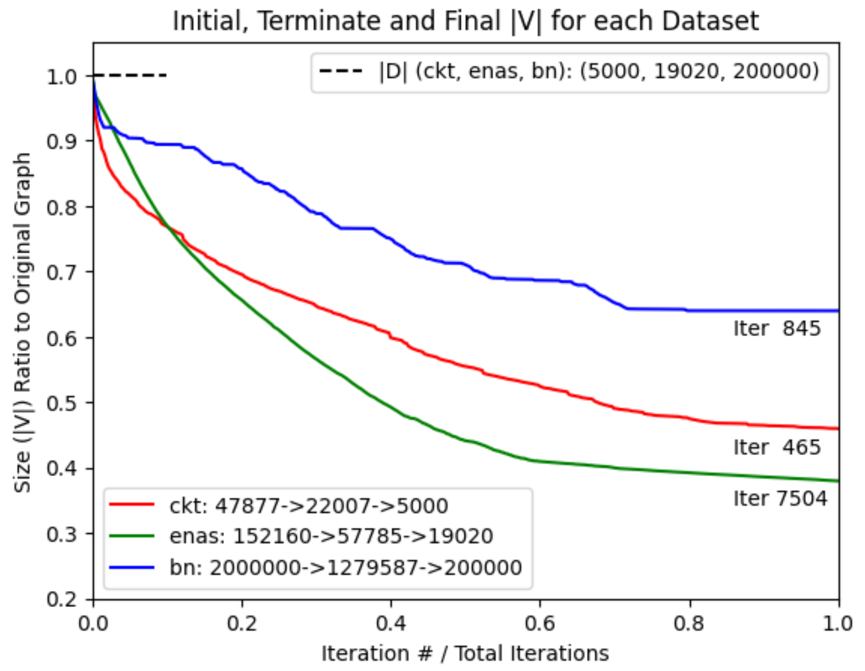
(Left) We visualize the best discovered designs from BO. (Right) Effectiveness in real-world electronic circuit design.



(a) ENAS (WS-Acc: 74.8, 74.9) (b) BN (Ground-truth) (c) CKT (FoM: 306.32)

Methods	Valid DAGs (%) ↑	Valid circuits (%) ↑	Novel circuits (%) ↑	BO (FoM) ↑
PACE	83.12	75.52	97.14	33.2742
DAGNN	83.10	74.21	**97.19**	33.2742
D-VAE	82.12	73.93	97.15	32.3778
GCN	81.02	72.03	97.01	31.6244
GIN	80.92	73.17	96.88	31.6244
NGNN	82.17	73.22	95.29	32.2827
Graphormer	82.81	72.70	94.80	32.2827
CktGNN	98.92	98.92	92.29	33.4364
CktGNN (CktBench301)	—	—	—	190.2354
DIGGED (GNN)	**100**	**100**	78.80	**310.2635**
DIGGED (TOKEN)	92.2	92.2	60.7	—

(Left) We show graph size |H| aså function of iteration (same as the #rules induced). Lower legend follows the format initial → pre-termination → post-termination. (Right) Controlled study comparing with simpler node-order encodings.



		Valid	Unique	Novel	RMSE	Pearson's r	1st	2nd	3rd
Graph2NS-Default	ENAS	96.1	99.17	100	0.746	0.656	0.746	0.744	0.743
	BN	95.8	96.4	94.8	0.498	0.869	-11590	-11685	-11991
	CKT	80.2	71.0	96.8	0.695	0.738	220.96	177.29	148.92
Graph2NS-BFS	ENAS	40.8	100	100	0.806	0.595	0.746	0.746	0.745
	BN	2.2	100	100	0.591	0.819	-11601	-11892	-11950
	CKT	0.1%	100	100	0.676	0.751	-	-	-
Graph2NS-Random	ENAS	0%	-	-	0.859	0.508	-	-	-
	BN	8.4	100	100	0.535	0.857	-11523	-11624	-11909
	CKT	0%	-	-	0.680	0.760	-	-	-
DIGGED	ENAS	100	98.7	99.9	0.912	0.386	0.749	0.748	0.748
	BN	100	97.6	100	0.953	0.712	-11110	-11250	-11293
	CKT	100	100	78.8	0.627	0.787	306.32	296.82	265.53