
# *The Rise and Fall of Machine Learning for Computing and Optimization*

Cunxi Yu

*University of Maryland, College Park*

*Nvidia Research*







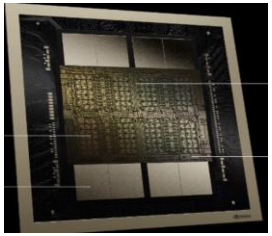

A. JAMES CLARK SCHOOL OF ENGINEERING

# Roadmap of Design Automation

19 billion


35 billion


208 billion



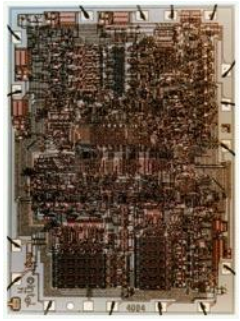
## Designed By Hand

Earliest EDA tools:
"*Berkeley VLSI Tools Tarball*"

$10^3$

***"Quine-McCluskey", 1956***

**1980s**



*2,300*

***"Espresso, SIS, Magic"***

**1990s**

$10^5$

## Birth of Commercial EDA

Cadence, Synopsys, Mentor, etc.

## Growing

***"ABC, Yosys, TDS, VPR, LegUp"***
***"Xilinx, Intel(Altera)"***

$10^{7\sim9}$

**2000s**

**Today**

$10^9$

## A primary barrier of hardware innovation

~~**Automation**~~

**Increasing manual efforts of automations**

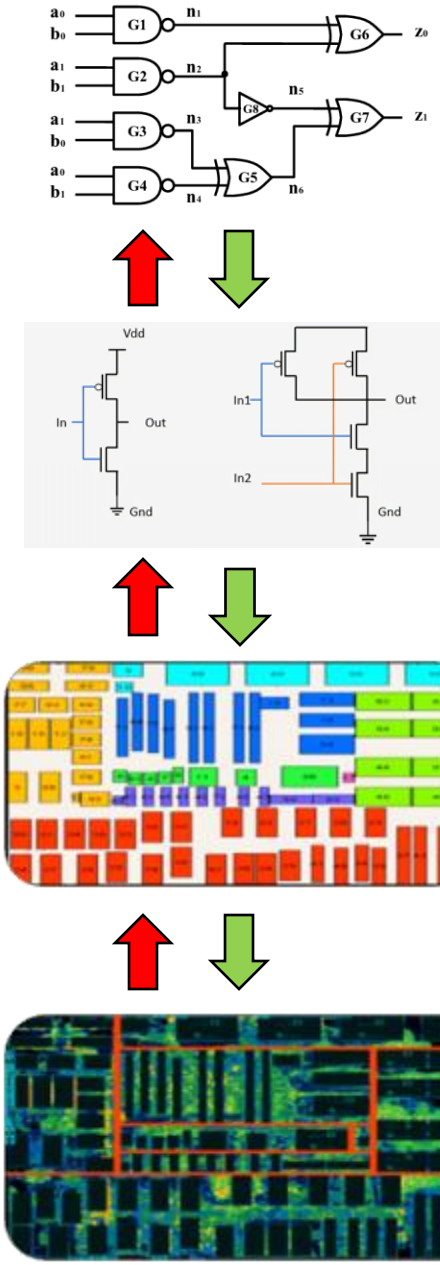
## Let's automate it, again!

# Focuses of EDA

Abstraction ↑

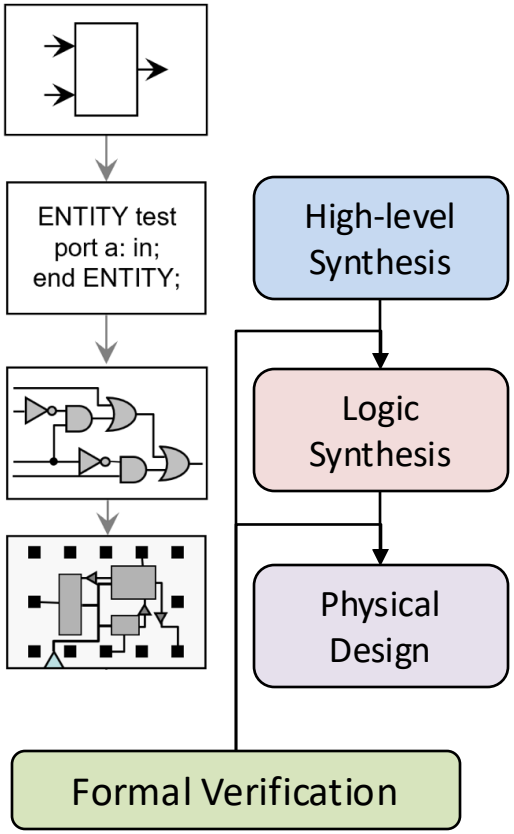
Design flow ↓

- Methodologies
  - Bottom-up abstraction
  - Top-down design flow
  - (Combinatorial) optimizations
- Advantages
  - ~~Easier R&D process~~
  - ~~Scalability and Runtime~~
- **Challenges**
  - Miscorrelations due to abstractions
  - Runtime complexity
  - Lack of algorithmic parallelism



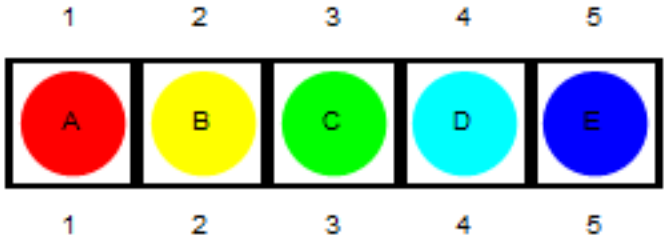
# Revisit the "Failed" Internship

- Identifying the challenges
  - Miscorrelation due to abstraction
  - Expensive turnaround
  - Massive design space (design and tools)




**Task**: Sequential Design Flow Exploration

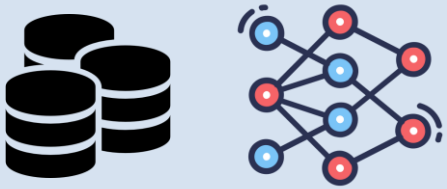
- Varies more than **150%** (e.g., 4.0 GHz vs. 2.7 GHz)
- Hard to predict (e.g., logic=4.0 GHz vs. physical=3.0 GHz)
- One evaluation on IBM PowerPC9: **7*20 hours**
  - **One year**: 52/400k = **0.013 %**



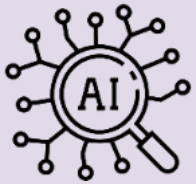
# Machine Learning in EDA and Optimizations

- My personal roadmap

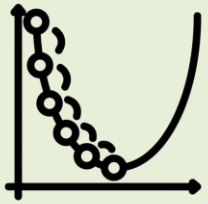
**Proxy-Model**



**Search**



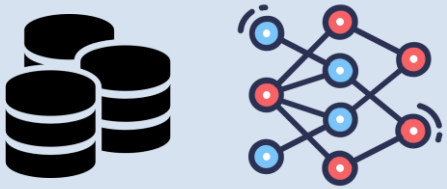

**Differentiable Methods**



# Machine Learning in EDA and Optimizations

- My personal roadmap

**Proxy-Model**



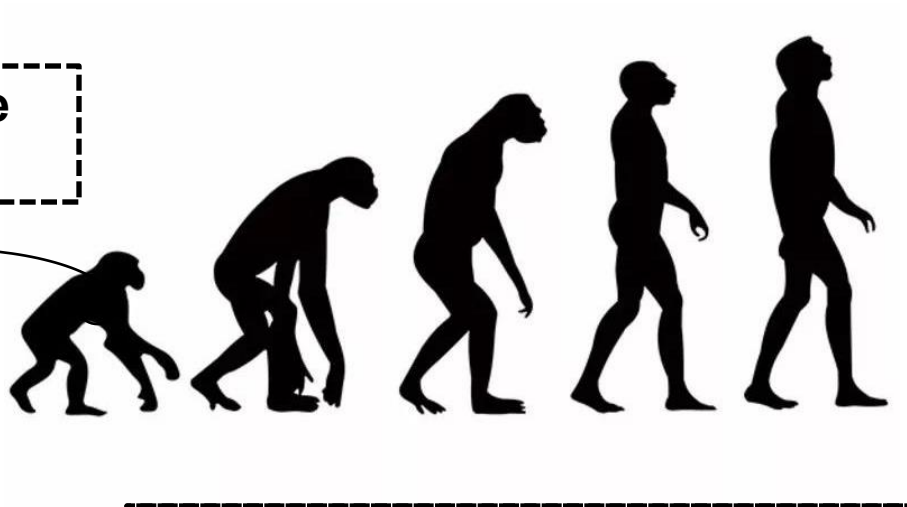
Search



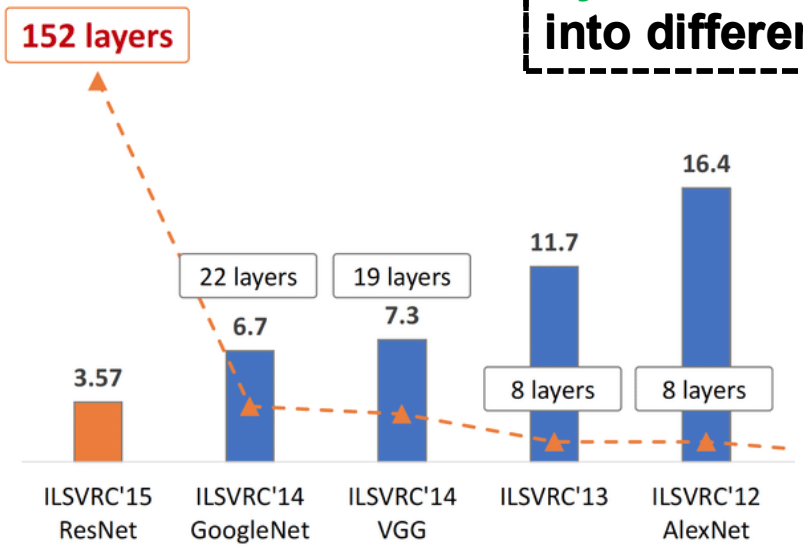

Differentiable Methods

# (My Personal) Roadmap of ML for Chip Design

Use ML to solve one synthesis problem?



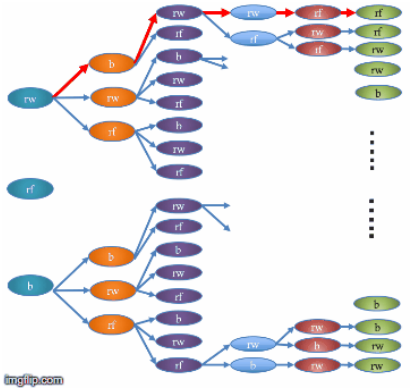
Synthesis flows ~~Images~~ to be classified into different ~~classes~~ QoR labels

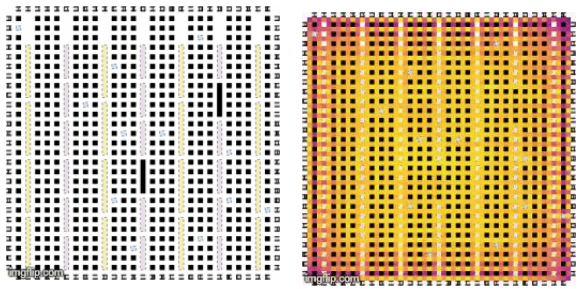



# Synthesis Design Space Exploration via Proxy Models

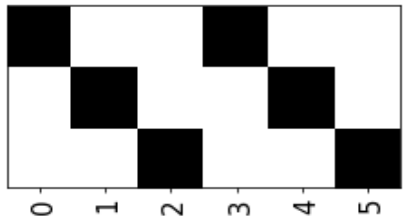
- Static and online learning
  - Imaging, graph, time-series, incremental, RL, etc.
- Model-inference guided synthesis DSE
  - *FlowGen* [DAC'18, MLCAD'20], *BoolGebra*[DATE'24]

**Open-source:**
https://github.com/orgs/Yu-Maryland/



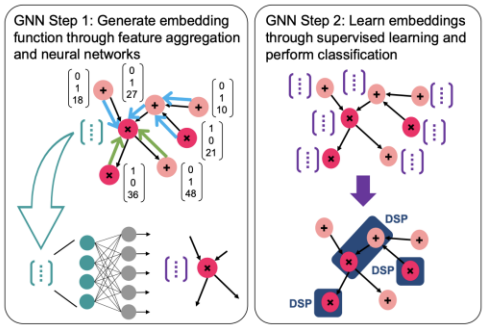


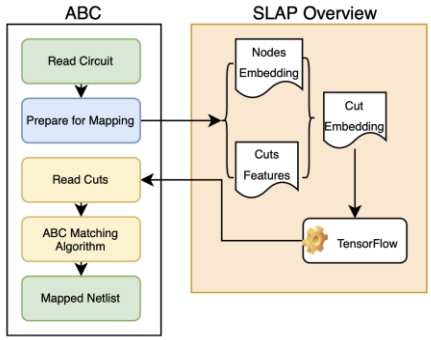
*Painting-on-Place* [DAC'19]




Flow embedded as a *"binarized image"*

- High-level Synthesis [FCCM'21] `*code-evolve*`` *in* synthesis [DAC'21, ASP-DAC'21]



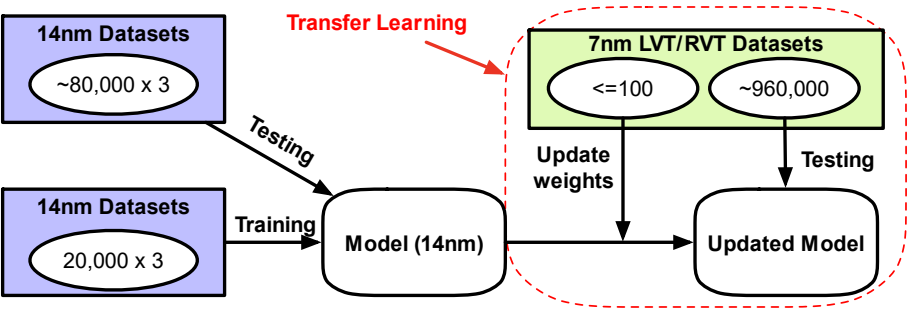



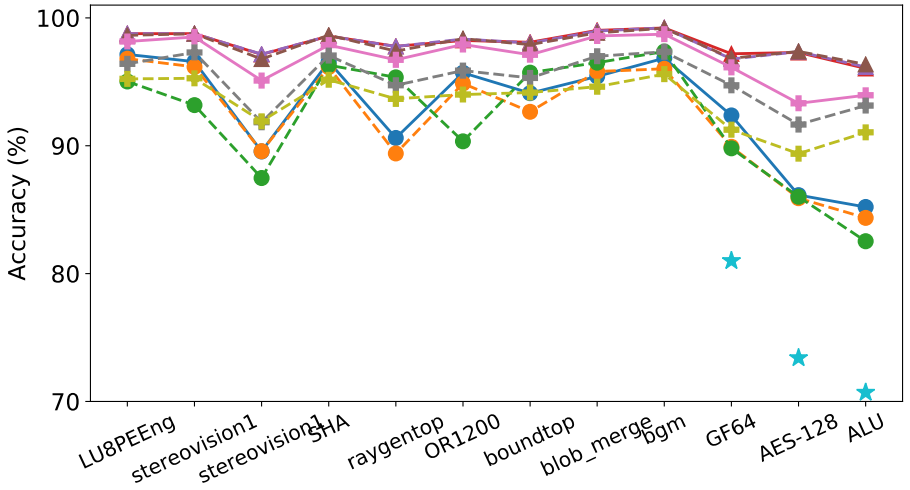
***Is life always this beautiful?***

# Limitations of the First Generation

- Case studies of logic synthesis

| Transferability | Data Availability | Integration |
|---|---|---|






Decision Making in Synthesis cross Technologies using LSTMs and Transfer Learning. Cunxi Yu, Wang Zhou. MLCAD'20

# Limitations of the First Generation

- Case studies of logic synthesis

| Transferability | **Data Availability** | Integration |
|---|---|---|

| Design | 14nm | 7nm RVT | 7nm LVT |
|---|---|---|---|
| 64-bit Montgomery | 99,997 | 99,997 | 99,997 |
| 64-bit ALU | 100,000 | 100,000 | 100,000 |
| 128-bit AES core | 99,737 | 99,737 | 99,737 |
| LU8PEng | - | 20,000 | 20,000 |
| Stereovison0 | - | 20,000 | 20,000 |
| Stereovison1 | - | 20,000 | 20,000 |
| SHA | - | 20,000 | 20,000 |
| raygentop | - | 20,000 | 20,000 |
| OR1200 | - | 20,000 | 20,000 |
| Boundtop | - | 20,000 | 20,000 |
| blob_merge | - | 20,000 | 20,000 |
| bgm | - | 20,000 | 20,000 |

**Another example of "lack of predictability"!!**



Decision Making in Synthesis cross Technologies using LSTMs and Transfer Learning. Cunxi Yu, Wang Zhou. MLCAD'20

# Limitations of the First Generation

- Case studies of logic synthesis

| Transferability | Data Availability | **Integration** |
|---|---|---|

```
(base) [cunxi@yulab1 ~]$ abc
UC Berkeley, ABC 1.01
abc 01> flowtune
```

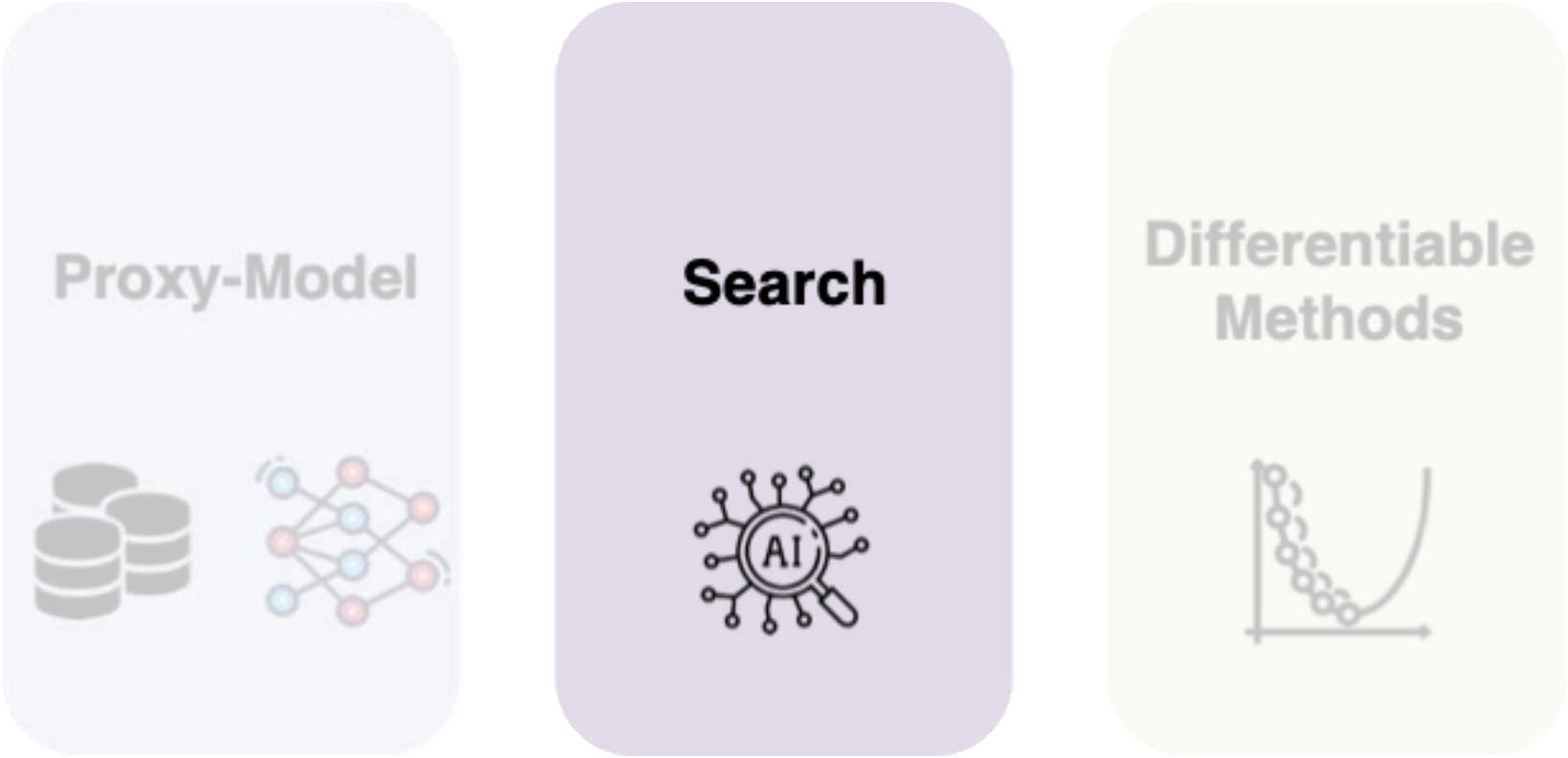
```
(base) [cunxi@yulab1 ~]$ abc
UC Berkeley, ABC
abc 01> resyn
```

- Close-loop system integration is challenging

Yu, Cunxi. "Flowtune: Practical multi-armed bandits in boolean optimization." ICCAD'20
Neto, Walter Lau, Yingjie Li, Pierre-Emmanuel Gaillardon, and Cunxi Yu. "FlowTune: End-to-end Automatic Logic Optimization Exploration via Domain-specific Multi-armed Bandit." *TCAD'22*

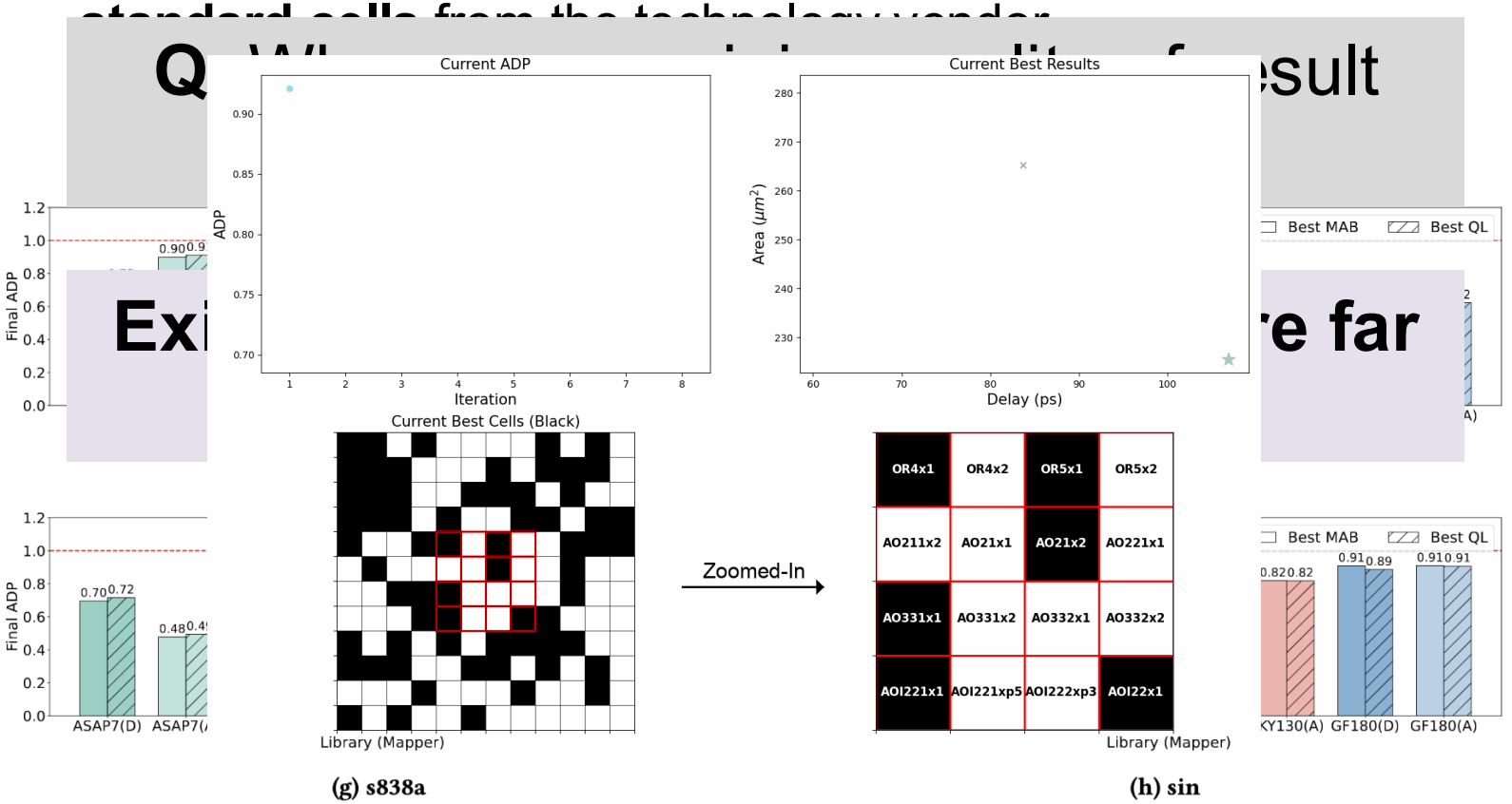
# Machine Learning in EDA and Optimizations

- My personal roadmap



# Example in *Search* – Technology Mapping

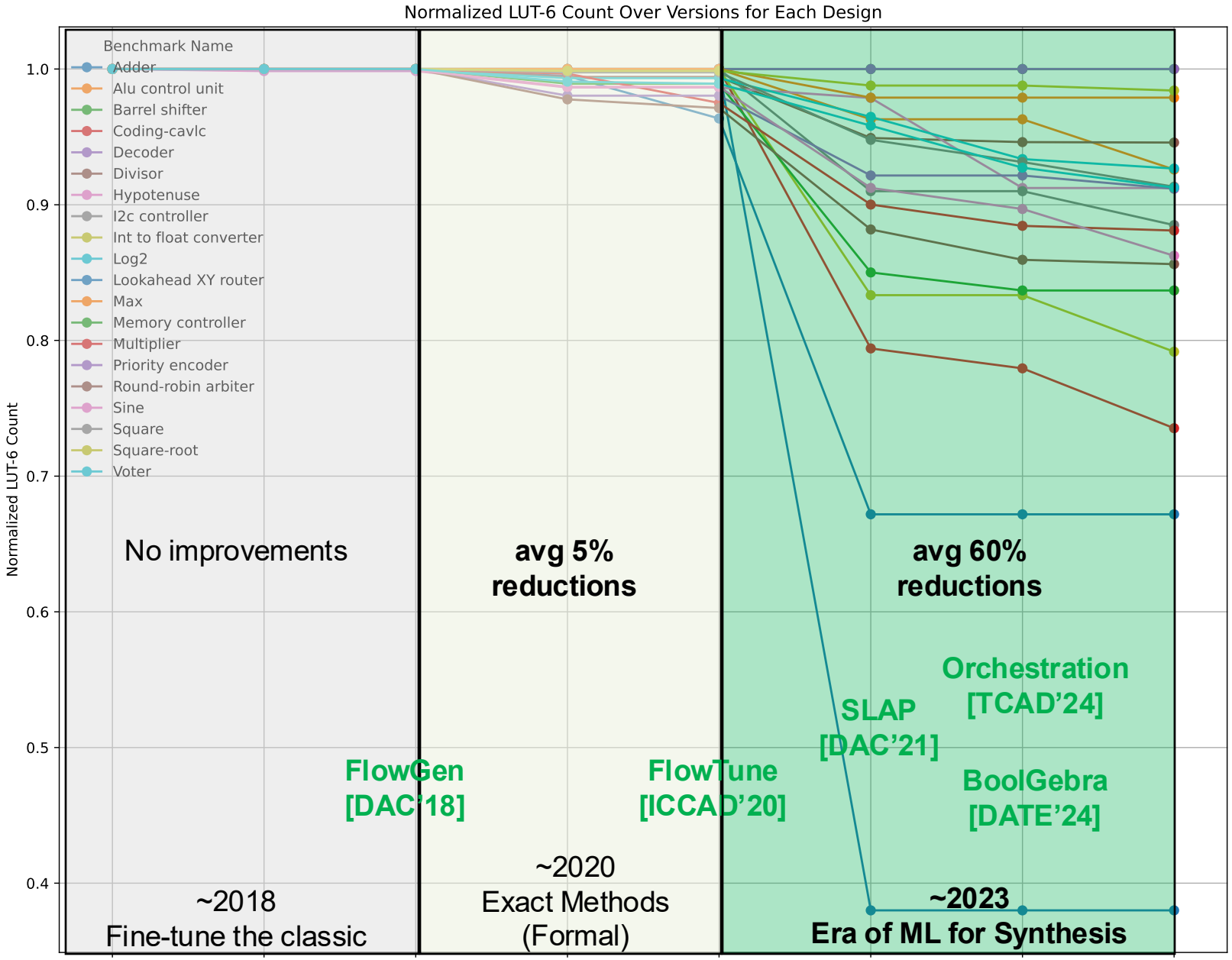
- Technology mapping
  - *SLAP* [DAC'21], *FlowTune* [ICCAD'20, TCAD'22], *MapTune* [ICCAD'24]
  - Counter-intuitive takeaway: mapping with **partially selected**



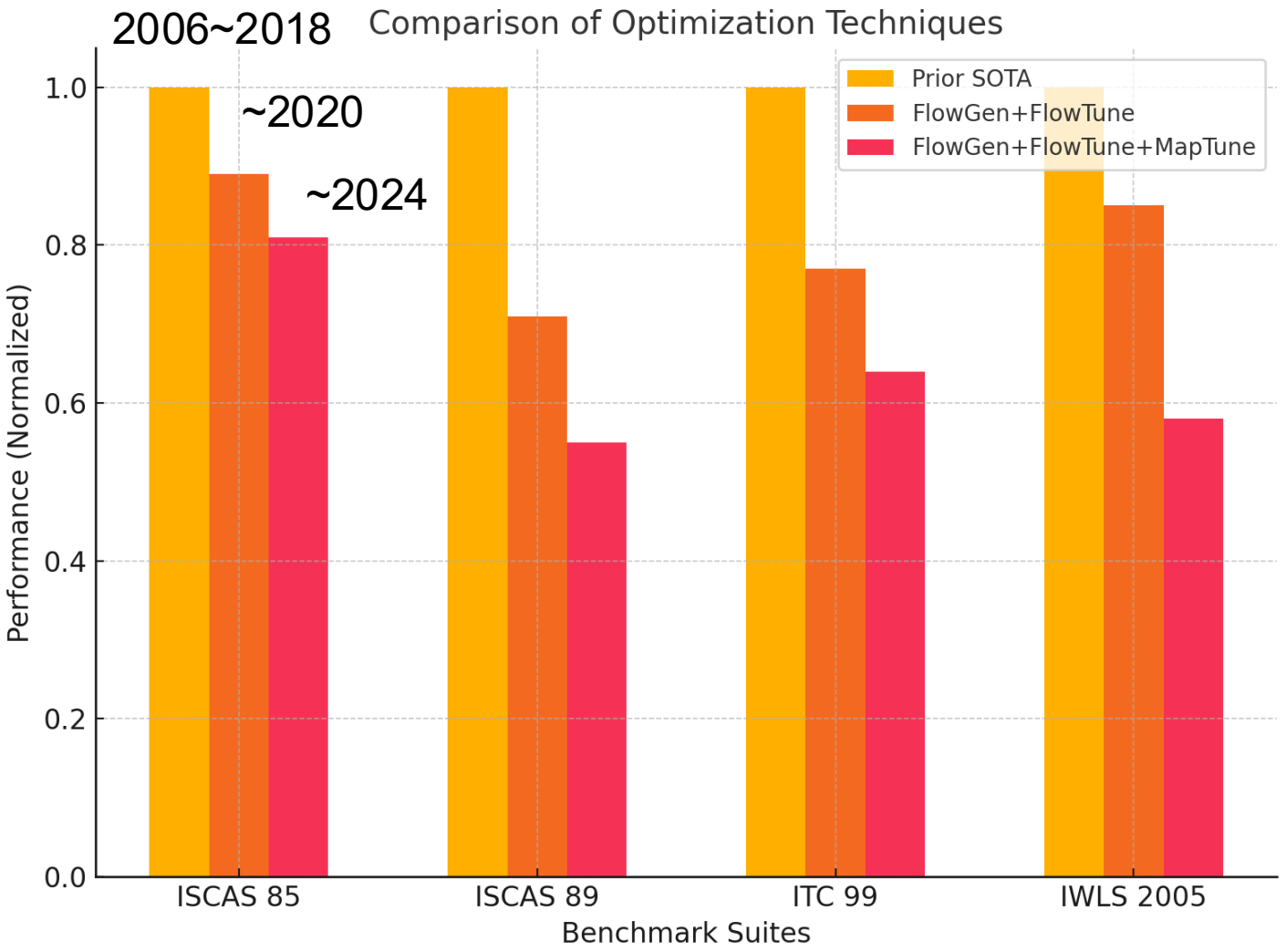

(g) s838a (h) sin

MapTune: Advancing ASIC Technology Mapping via Reinforcement Learning Guided Library Tuning
Mingju Liu, Daniel Robinson, Yingjie Li, Cunxi Yu. IEEE/ACM International Conference On Computer Aided Design (ICCAD'24)

# The EPFL Benchmark Results (2016 – 2023)

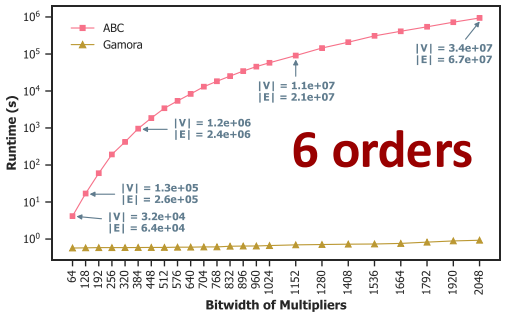


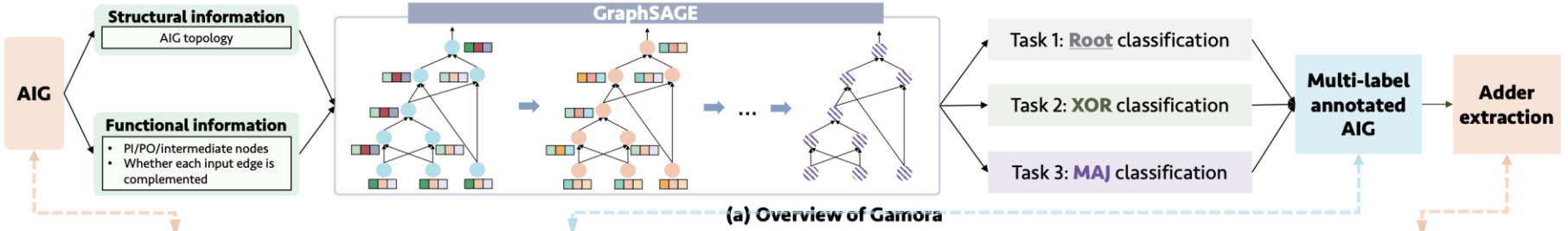
# The ISCAS 85/89/99 Benchmarks

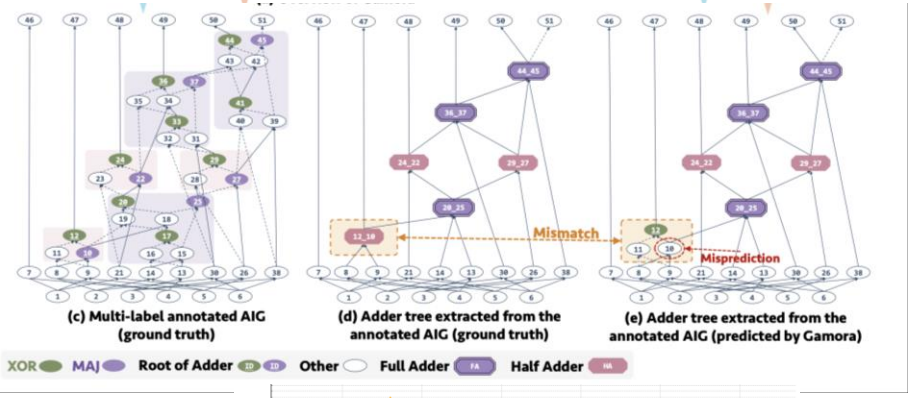


# Applications in Formal Verification

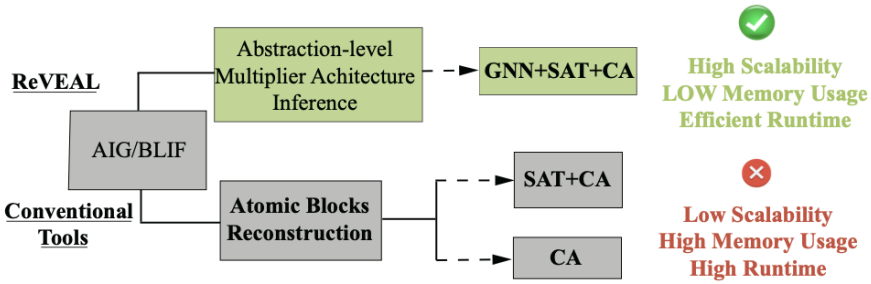
- Boolean Reasoning



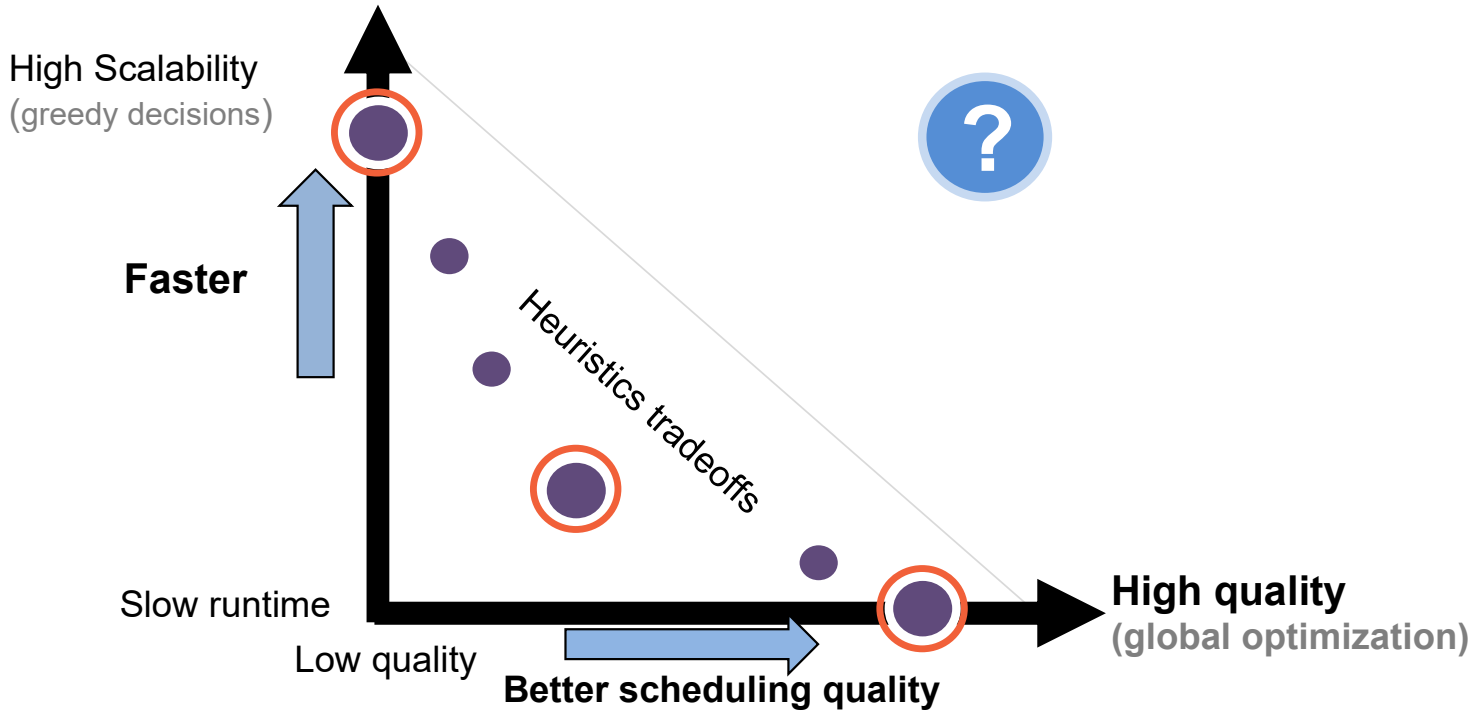

*GAMORA (**Best Paper**)* [DAC'23]
*HOGA* [DAC'24]


*REVEAL* [IWLS'25]

# Touching the Algorithms

# ML or ML System for Combinatorial Optimization



# Challenges of Prior Methods

**Q1**: How to get rid of training and data collection?

**Map the problem as an optimization form**

**Q2:** How to leverage the power of ML infrastructure?

**Construct the problem as a differentiable optimization form**

**Q3:** How to generalize the differentiable mapping ?

**Differentiable mapping of a generic formal encoding**

# Machine Learning in EDA and COs

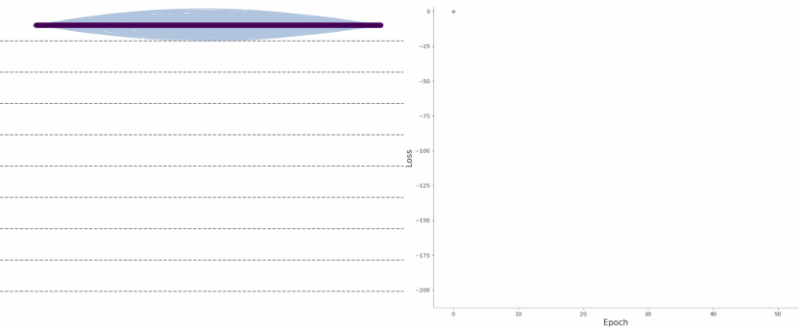
Proxy-Model

Search

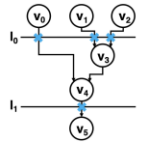



**Differentiable Methods**

# Differentiable Methods in Chip Design

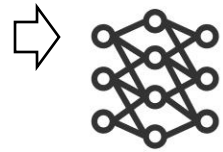


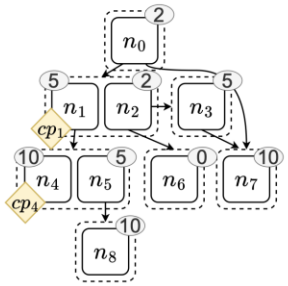

Orders of magnitude faster than IBM CPLEX, Google CP-SAT, Gurobi



Differentiate SDC/ILP

$$s_0 - s_4 \le 0$$
$$s_1 - s_3 \le 0$$
$$s_2 - s_3 \le 0$$
$$s_3 - s_4 \le 0$$
$$s_4 - s_5 \le 0$$



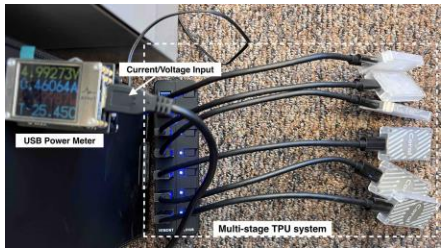


| $n$ | $cp_n$ | $p_n$ |
|---|---|---|
| 0 | 1 | 1 |
| 1 | $cp_1$ | $cp_1$ |
| 2 | $1 - cp_1$ | $1 - cp_1$ |
| 3 | 1 | $1 - cp_1$ |
| 4 | $cp_4$ | $cp_1 cp_4$ |
| 5 | $1 - cp_4$ | $cp_1 - cp_1 cp_4$ |
| 6 | 1 | $1 - cp_1$ |
| 7 | 1 | 1 |
| 8 | 1 | $cp_1 - cp_1 cp_4$ |

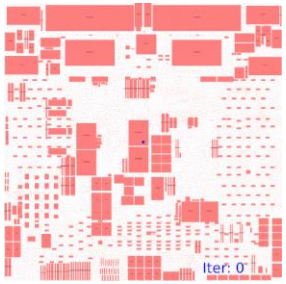
*Differentiable COs*
[ICML'24, ICML'25]

*SmoothE (**Best Paper**)*
[ASPLOS'25]

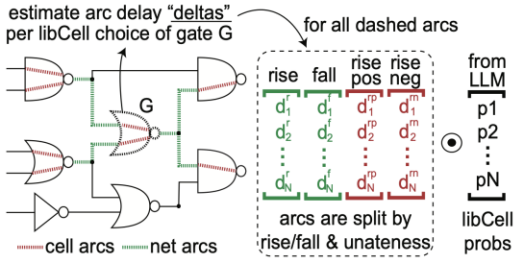
*RESPECT*
[DAC'23, ICCAD'24]



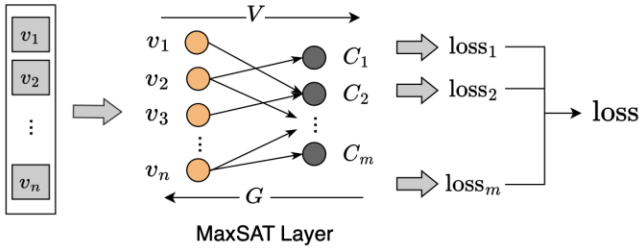

Improved energy efficiency and compilation runtime than Google EdgeTPU Compiler




*DREAMPlace*
[Lin'19, Agnesina'23]

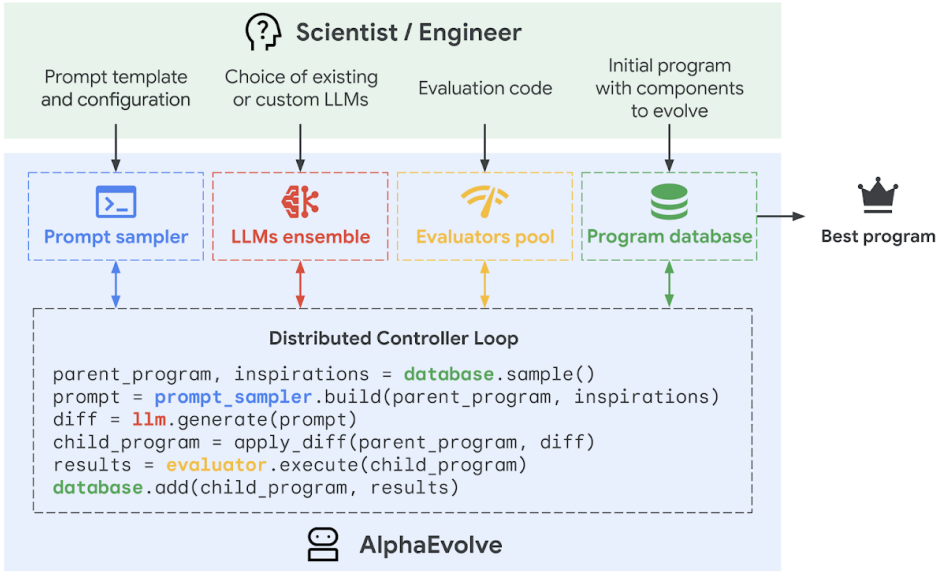



*LEGO Sizer, INSTA*
[Lu'24*2]

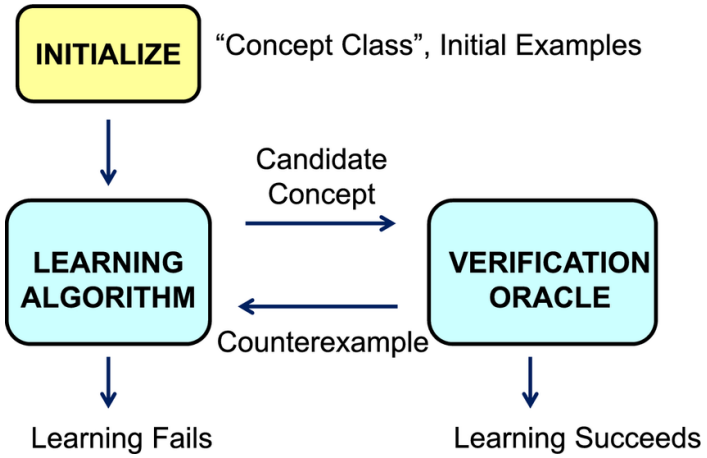



DiffSAT
[Zhang'24]

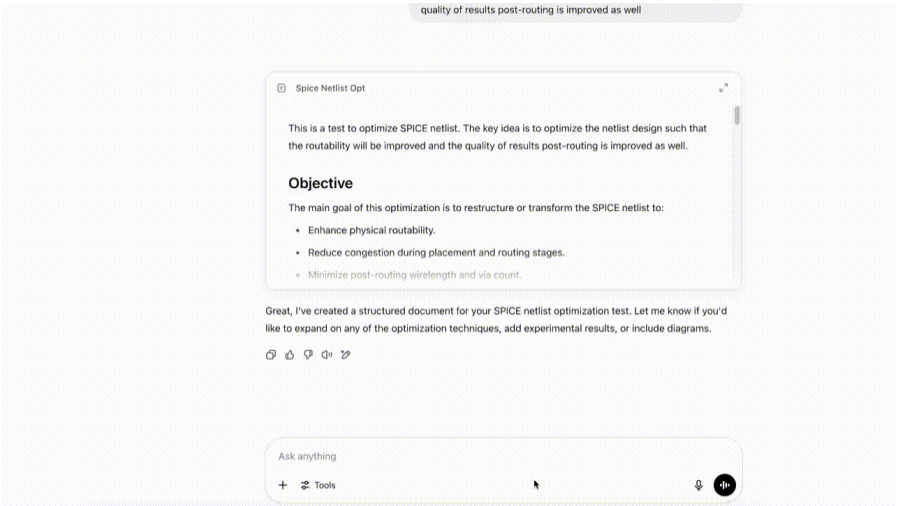
# What's Next?




Google's AlphaEvolve



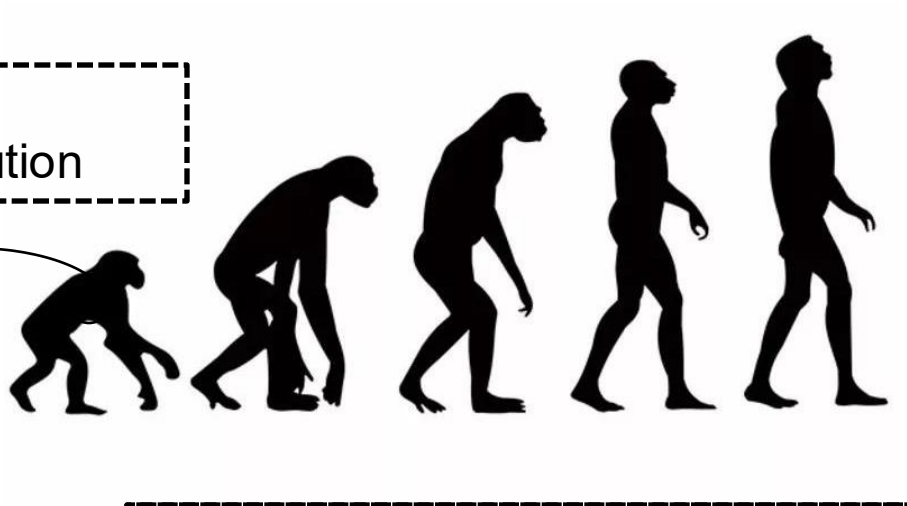

Formal Methods in the Loop (Chain of Thoughts & Proof)



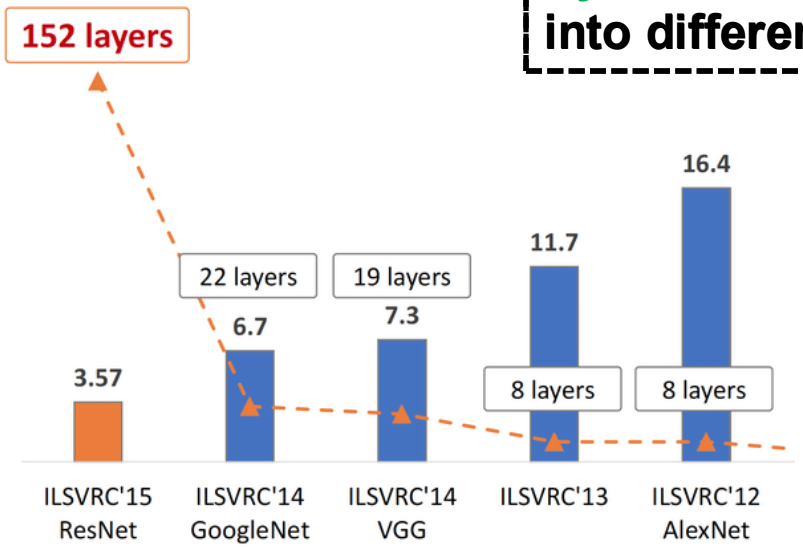

Agentic AI Assistant for X

# Conclusion

Hopefully, I am not experiencing a devolution



**Synthesis flows ~~Images~~ to be classified into different ~~classes~~ QoR classes**




# Thank you!