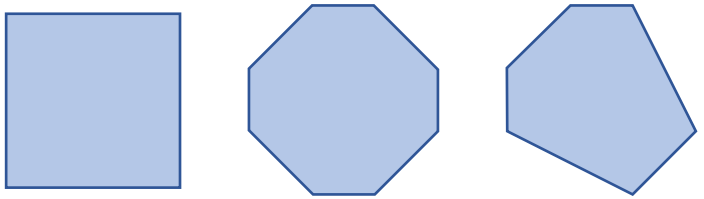
Learning Fast and Precise Numerical Analysis

Jingxuan He Gagandeep Singh Markus Püschel Martin Vechev

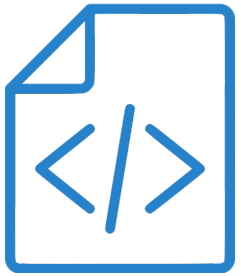
Department of Computer Science

ETH Zürich

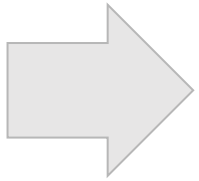
Numerical program analysis



Abstract elements



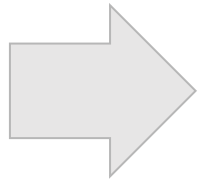
Program

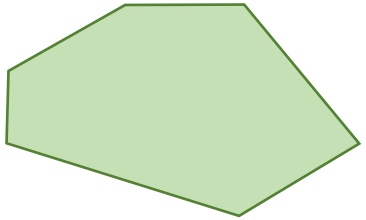


$\sqcup \quad \sqcap \quad \sqsubseteq \quad \nabla$

$[x := e] \quad [x < e]$

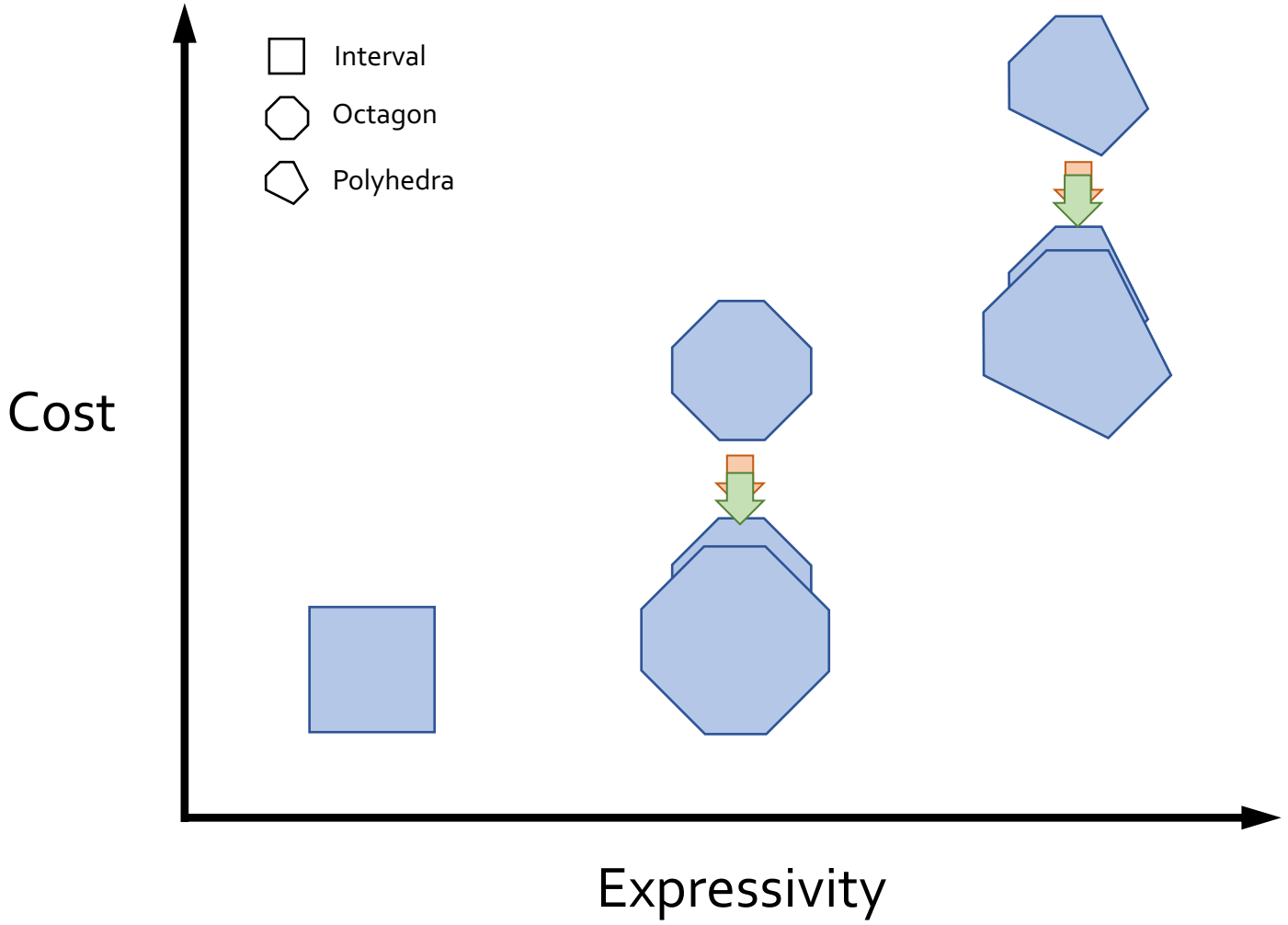
Abstract transformers





Invariants

Tradeoff for numerical analysis



Online decomposition

[PLDI 2015, POPL 2017, POPL 2018]

Drawback: redundant computation during inference of invariants

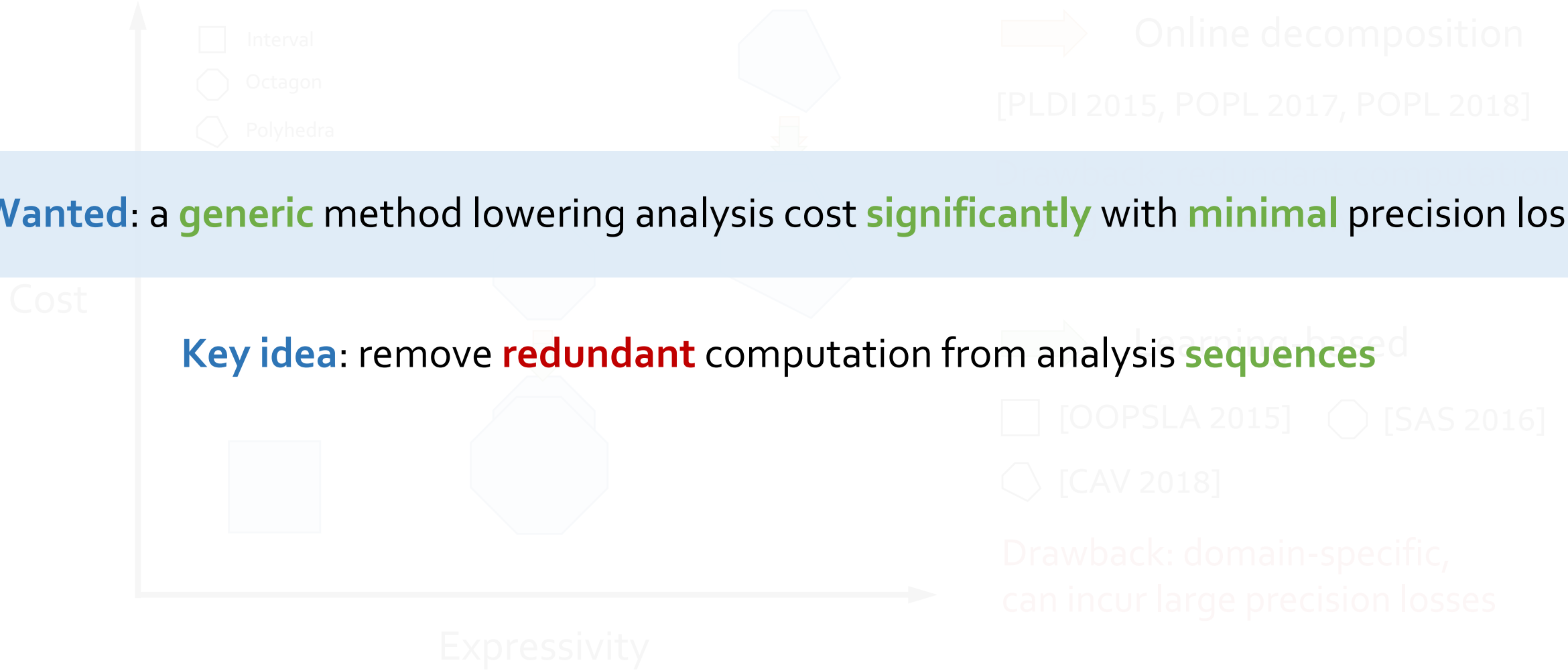
Learning-based

[OOPSLA 2015] [SAS 2016]

[CAV 2018]

Drawback: domain-specific, can incur large precision losses

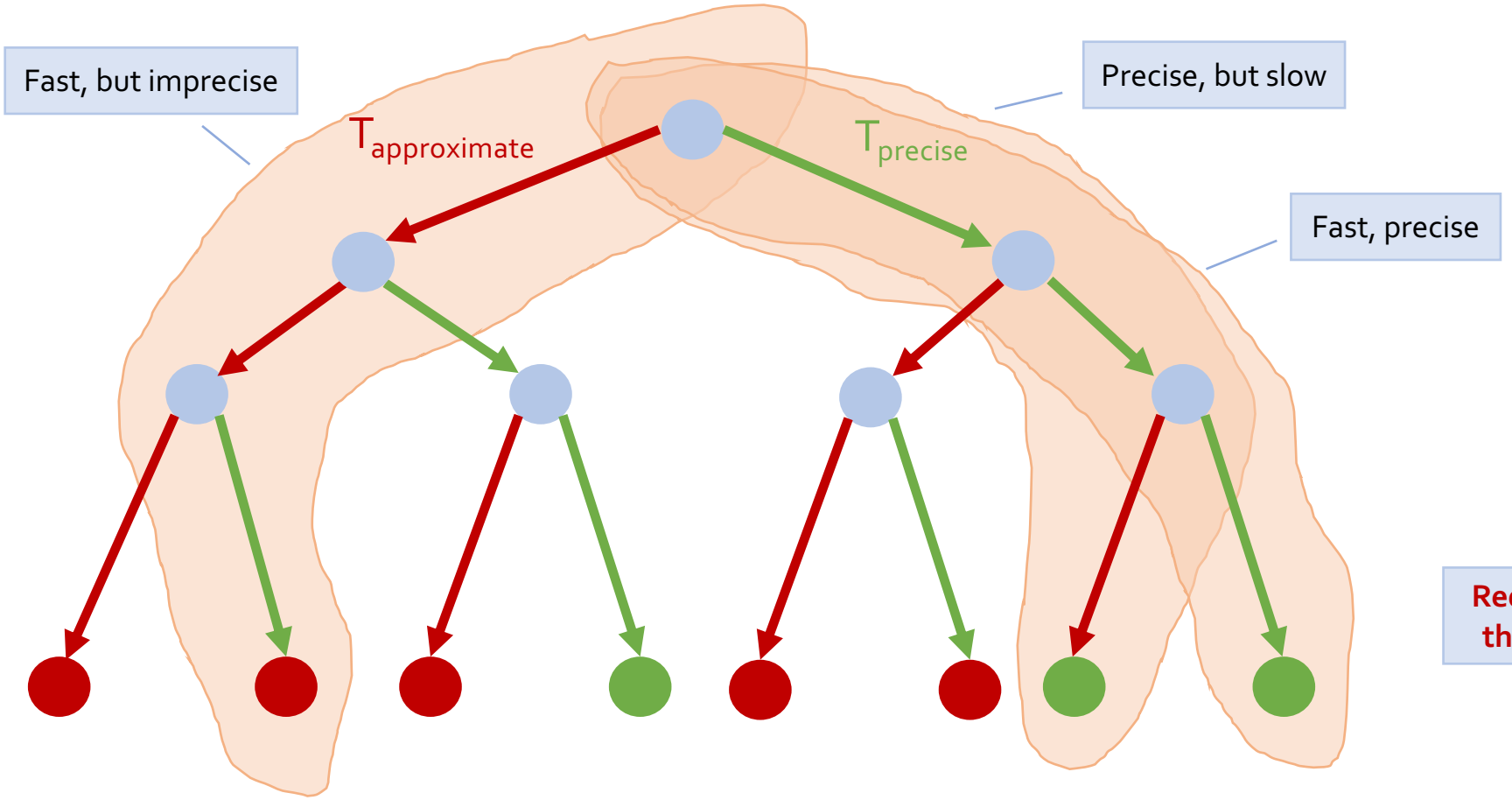
Tradeoff for numerical analysis

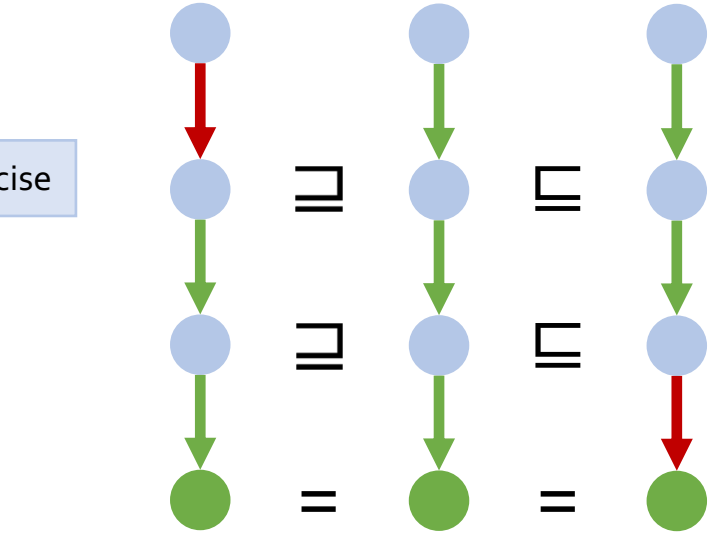


Wanted: a **generic** method lowering analysis cost **significantly** with **minimal** precision loss

Key idea: remove **redundant** computation from analysis **sequences**

Key observation: redundancy in analysis sequences





Redundancy in the sequence

This work:

> 100X speedup with **precise** invariants for large linux device driver programs by structured redundancy removal

Challenge: define and apply approximate transformers for redundancy removal

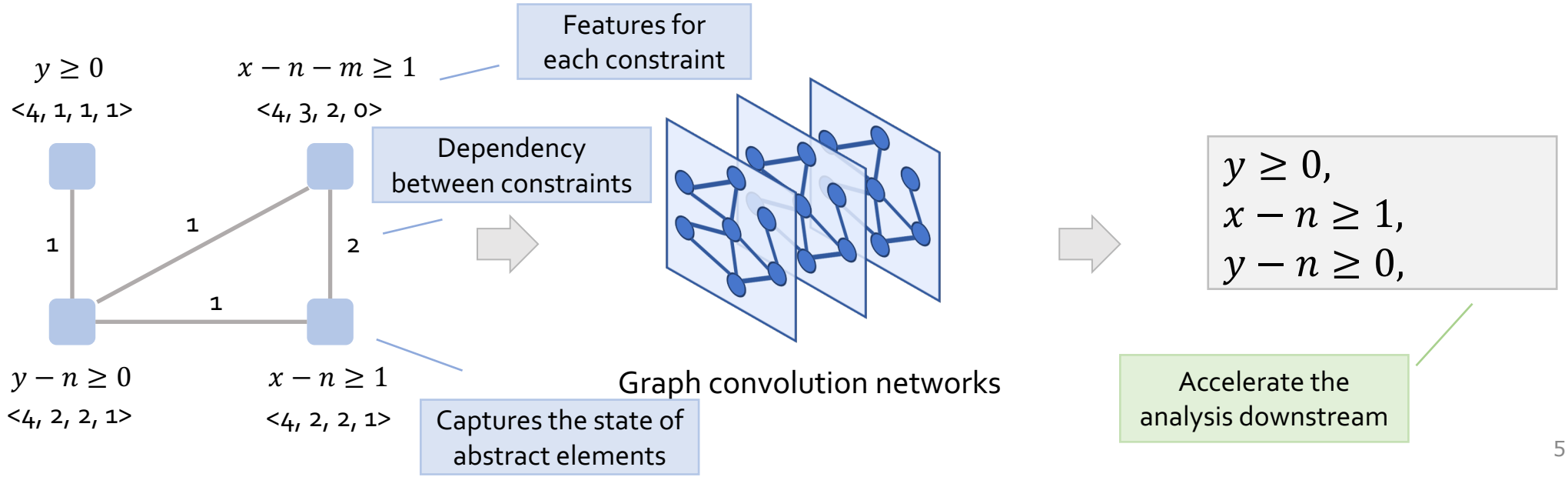
Lait: approximate join transformer

Lait: learning-based constraint removal from the expensive join transformer

Join:

$$\begin{aligned} &y \geq 0,\\ &x - n - m = 1,\\ &y - n \geq 0,\\ &m \geq 0, \end{aligned} \quad \sqcup \quad \begin{aligned} &y \geq 0,\\ &x - n = 2,\\ &y - n \geq 0,\\ &m \leq 0, \end{aligned} \quad = \quad \begin{aligned} &y \geq 0,\\ &x - n - m \geq 1,\\ &x - n \geq 1,\\ &y - n \geq 0, \end{aligned}$$

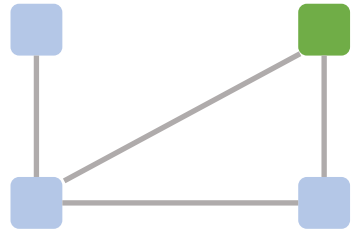
Lait:



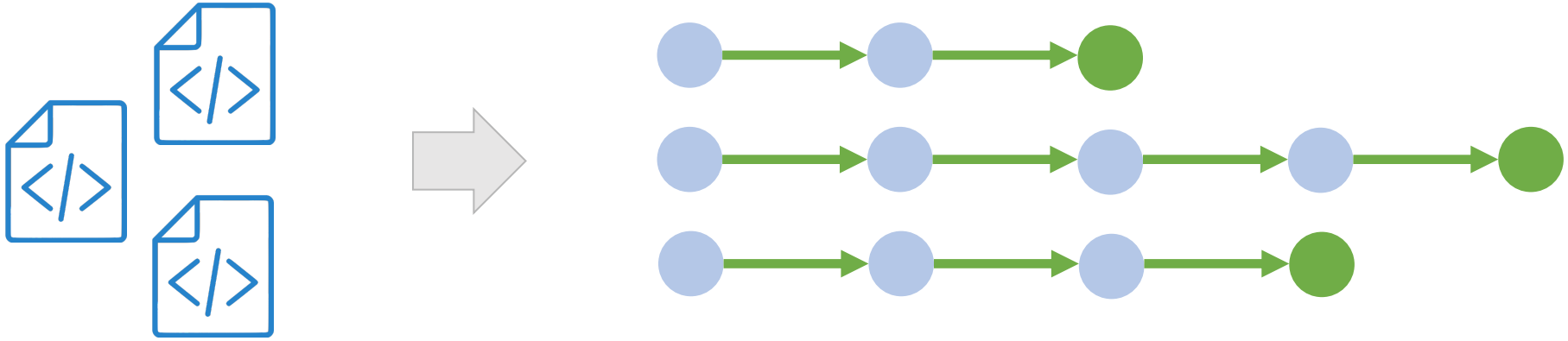
Our learning algorithm

To train  , we need a **supervised** dataset of graphs where removed constraints are labelled:

- **True**, removing the constraint does not affect the analysis precision, or
- **False**, removing the constraint loses precision.

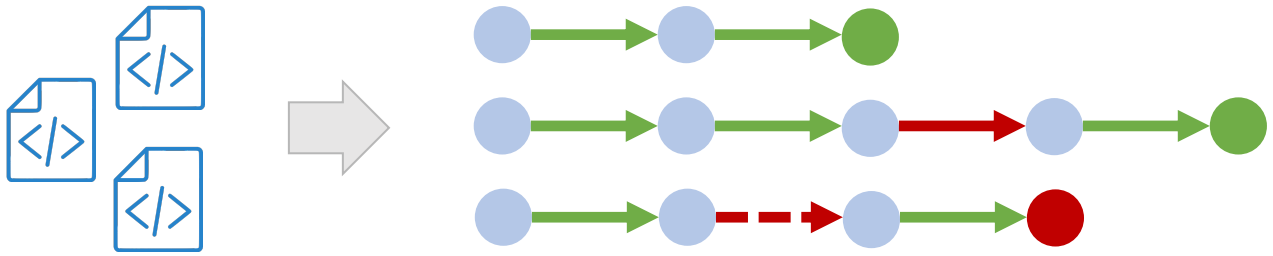


Step 1: running precise analysis on training programs to obtain **ground truth** for precision



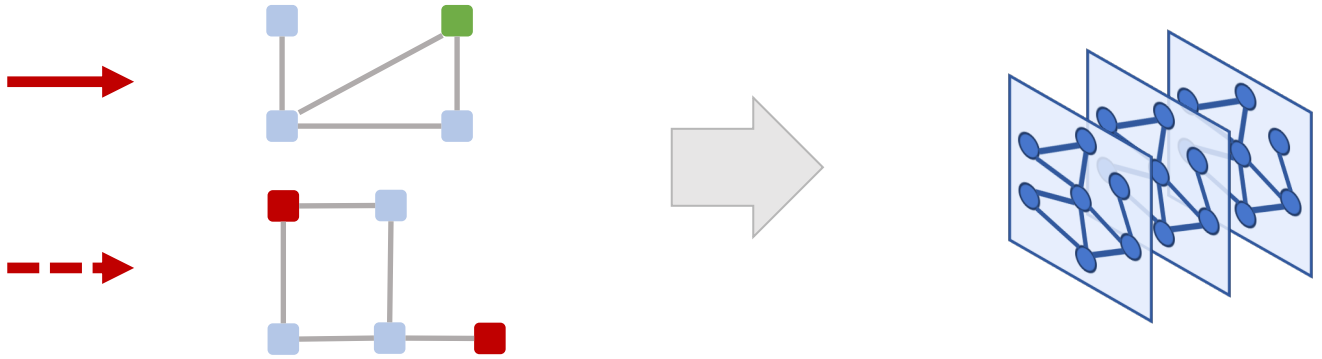
Our learning algorithm

Step 2: running approximate analysis with constraint removal on training programs



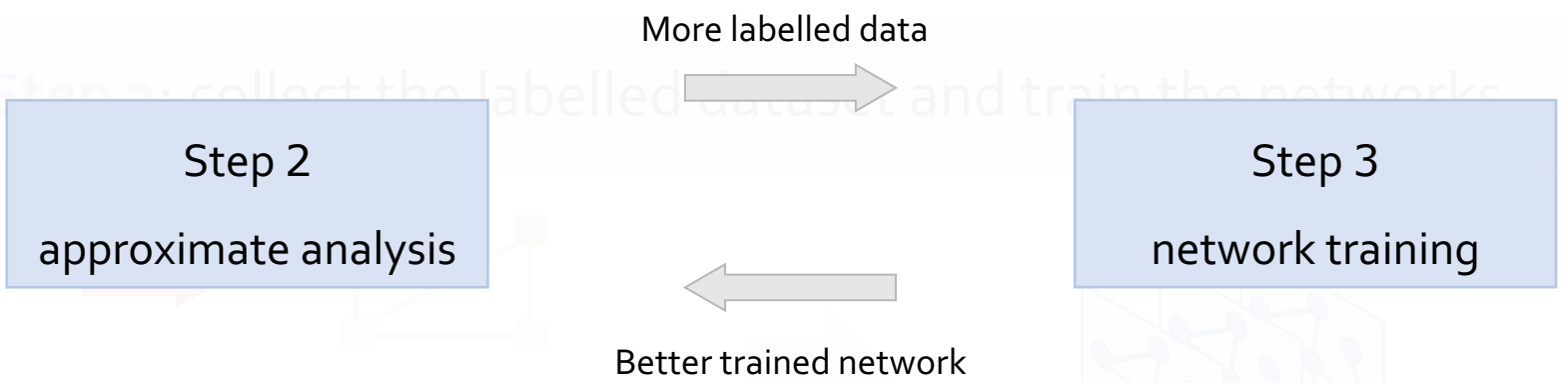
ε-greedy : calls Lait with $1 - \varepsilon$ probablity,
or a random removal policy with ε probability.

Step 3: collect the labelled dataset and train the networks.



Our learning algorithm

Iterative training: running steps 2 and 3 for multiple iterations



Evaluation setup

Instantiation for online decomposed Polyhedra and Octagon analysis

- Implementation incorporated in http://elina.ethz.ch/ ELINA

SV-COMP benchmarks and crab-llvm analyzer

Lait v.s.

- ELINA: a state-of-the-art library for numerical domains (**ground truth** for precision)
- Poly-RL: an approximate Polyhedra analysis with reinforcement learning [CAV 2018]
- HC: a hand-created heuristics for redundancy removal

Precision = % of program points with the same invariants as ELINA

Results for Polyhedra analysis

Training on 30 programs

Time limit: 2h, Memory limit: 50GB

Benchmark	Number of program points	ELINA	HC		Poly-RL		Lait	
		time (s)	speedup	precision	speedup	precision	speedup	precision
qlogic_qlge	5748	3474	49x	99	4.2x	98.8	53x	100
peak_usb	1300	1919	325x	81	1.3x	95.1	315x	100
stv090x	7726	3401	4.6x	95	MO	100	6.3x	97
acenic	1359	3290	TO	65	1.1x	100	223x	100
qla3xxx	2141	2085	163x	95	210x	99.8	169x	100
cx25840	1843	56	8.8x	83	0.7x	97.9	9.9x	83
mlx4_en	6504	46	1.2x	91	1.2x	98.9	1.0x	100
advansys	3568	109	1.7x	92	crash	98.8	1.4x	99.7
i7300_edac	309	36	2.6x	83	1.2x	99.8	1.4x	99
oss_sound	2465	2428	245x	80	1.2x	100	229x	80

Results for Polyhedra analysis

Training on 30 programs

Time limit: 2h, Memory limit: 50GB

Benchmark	Number of program points	ELINA	HC		Poly-RL		Lait	
		time (s)	speedup	precision	speedup	precision	speedup	precision
qlogic_qlge	5748	3474	49x	99	4.2x	98.8	53x	100
peak_usb	1300	1919	325x	81	1.3x	95.1	315x	100
stv090x	7726	3401	4.6x	95	MO	100	6.3x	97
acenic	1359	3290	TO	65	1.1x	100	223x	100
qla3xxx	2141	2085	163x	95	210x	99.8	169x	100
cx25840	1843	56	8.8x	83	0.7x	97.9	9.9x	83
mlx4_en	6504	46	1.2x	91	1.2x	98.9	1.0x	100
advansys	3568	109	1.7x	92	crash	98.8	1.4x	99.7
i7300_edac	309	36	2.6x	83	1.2x	99.8	1.4x	99
oss_sound	2465	2428	245x	80	1.2x	100	229x	80

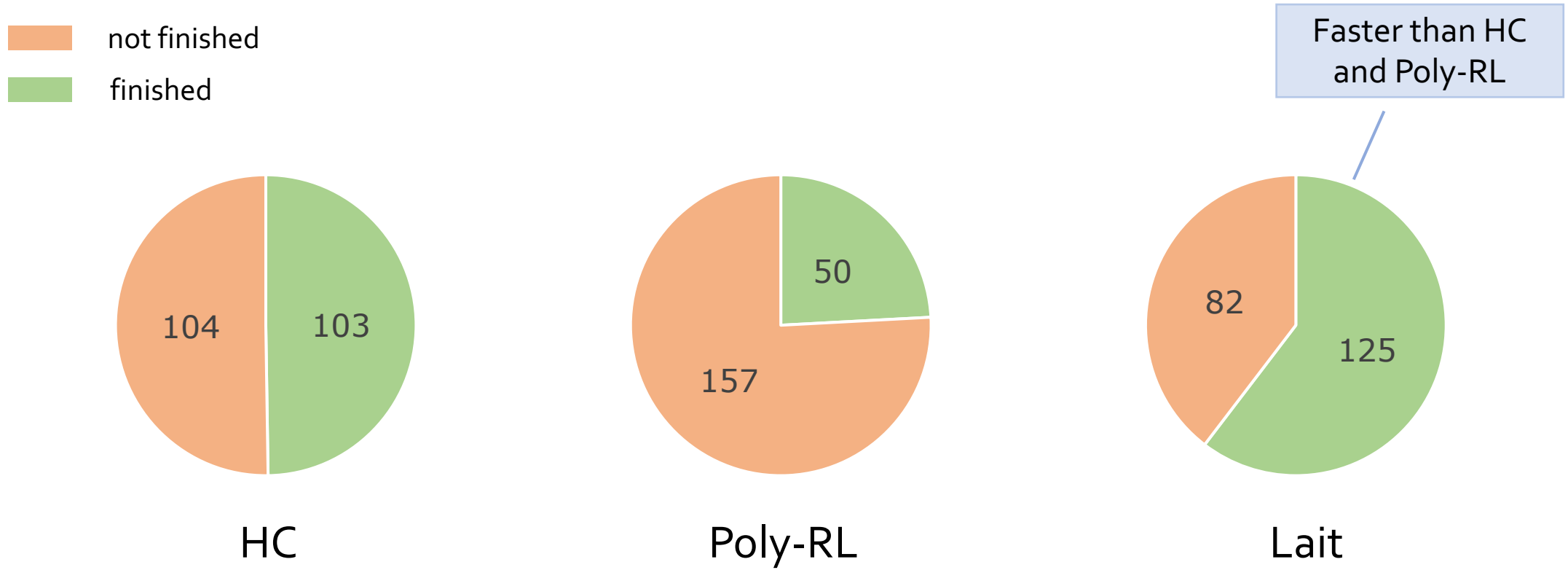
Statistics on the number of constraints

m: the number of constraints in abstract elements

Benchmark	m_{ELINA}		m_{HC}		$m_{\text{Poly-RL}}$		m_{Lait}	
	max	avg	max	avg	max	avg	max	avg
qlogic_qlge	267	6	19	4	205	5	33	4
peak_usb	48	7	17	5	48	7	24	7
stv090x	74	12	32	14	-	-	35	13
acenic	98	9	-	-	98	8	28	5
qla3xxx	284	17	30	9	218	15	19	8
cx25840	26	10	17	7	26	9	17	8
mlx4_en	56	4	53	4	54	4	56	4
advansys	38	9	37	9	-	-	38	8
i7300_edac	41	14	20	9	41	14	28	11
oss_sound	47	9	38	7	47	8	23	7

Results for Polyhedra analysis

On 207 programs for which ELINA does not finish within 2h:



Results for Octagon analysis

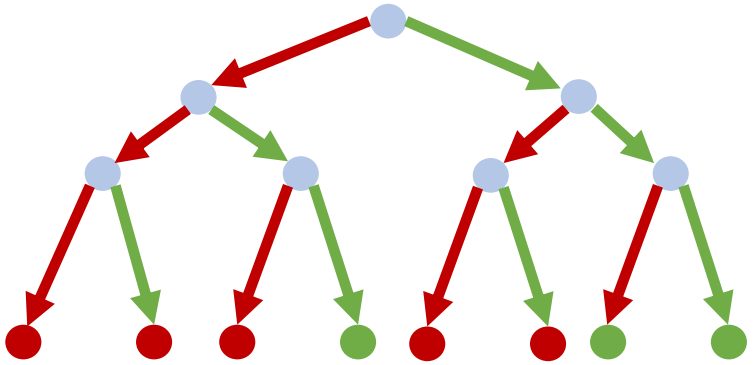
Training on 10 programs

Time limit: 2h, Memory limit: 50GB

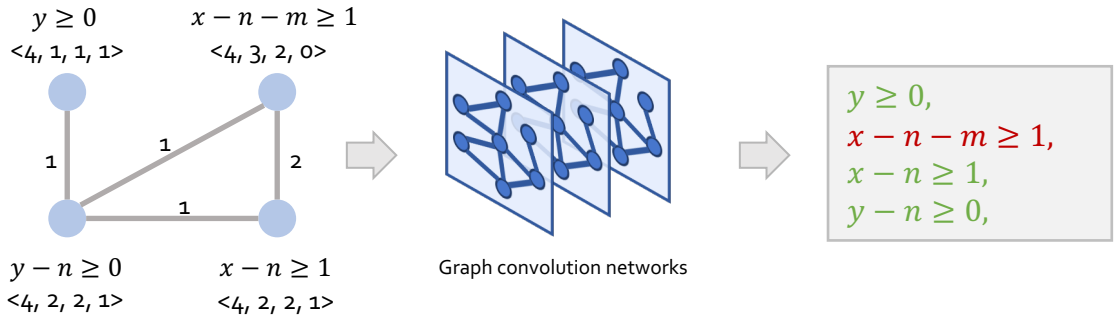
Benchmark	Number of program points	ELINA	HC		Lait	
		time (s)	speedup	precision	speedup	precision
advansys	3408	34	1.22x	99.4	1.15x	98.8
net_unix	2037	13	TO	52.5	1.45x	95.1
vmwgfx	7065	45	1.08x	100	1.24x	100
phoenix	644	26	1.55x	96.9	1.31x	100
mwl8k	4206	27	1.05x	64.2	1.55x	99.8
saa7164	6565	117	1.00x	57.8	1.54x	97.9
md_mod	8222	1309	TO	68.1	28x	98.9
block_rsxx	2426	14	1.11x	73.9	1.26x	98.8
ath_ath9k	3771	26	1.07x	65.7	1.33x	99.8
synclik_gt	2324	44	1.28x	100	1.23x	100

Summary

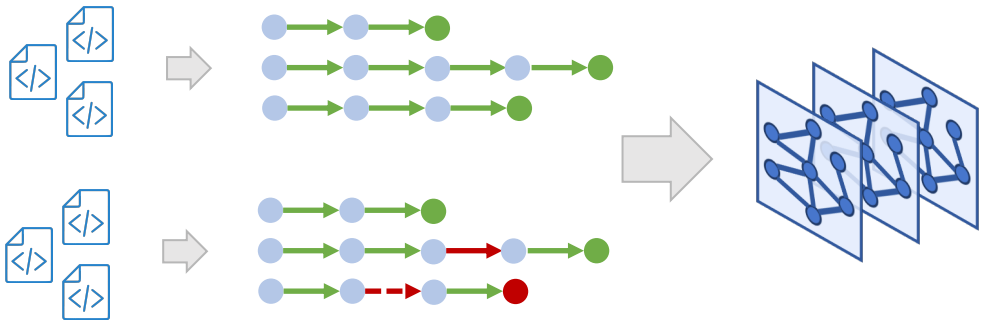
Redundancy in numerical analysis



Approximate join by constraint removal



Iterative learning algorithm



Promising results on two domains

Benchmark	Number of program points	ELINA	HC		Poly-RL		Lait	
		time (s)	speedup	precision	speedup	precision	speedup	precision
qlogic_qlge	5748	3474	49x	99	4.2x	98.8	53x	100
peak_usb	1300	1919	325x	81	1.3x	95.1	315x	100
stvo9ox	7726	3401	4.6x	95	MO	100	6.3x	97
acenic	1359	3290	TO	65	1.1x	100	223x	100
qla3xxx	2141	2085	163x	95	210x	99.8	169x	100
cx25840	1843	56	8.8x	83	0.7x	97.9	9.9x	83
mlx4_en	6504	46	1.2x	91	1.2x	98.9	1.0x	100
advansys	3568	109	1.7x	92	crash	98.8	1.4x	99.7
i73oo_edac	309	36	2.6x	83	1.2x	99.8	1.4x	99
oss_sound	2465	2428	245x	80	1.2x	100	229x	80