CVC4 IEEE-754 implementation

Current status and benchmarks

Florian Schanda

July 28, 2017

ALTRAN

1.

CVC4 Progress

ALTRAN

FP progress in CVC4

Solved

Benchmark	07-04	07-27
Schanda	27.5%	67.4%
Griggio	16.4%	17.3%
Heizmann	4.8%	69.1%
Industrial_1	0.0%	71.3%
Industrial_1_qf	0.0%	96.3%
NyxBrain	99.8%	99.8%
PyMPF	73.9%	97.4%
SPARK 2014	0.0%	79.0%
SPARK QF	0.0%	86.2%
Wintersteiger	85.0%	99.9%

FP progress in CVC4

Error

Benchmark	07-04	07-27
Schanda	86	2
Griggio	✓	✓
Heizmann	195	2
Industrial_1	268	3
Industrial_1_qf	268	✓
NyxBrain	40	4
PyMPF	13920	45
SPARK 2014	2468	15
SPARK QF	2468	52
Wintersteiger	5668	✓

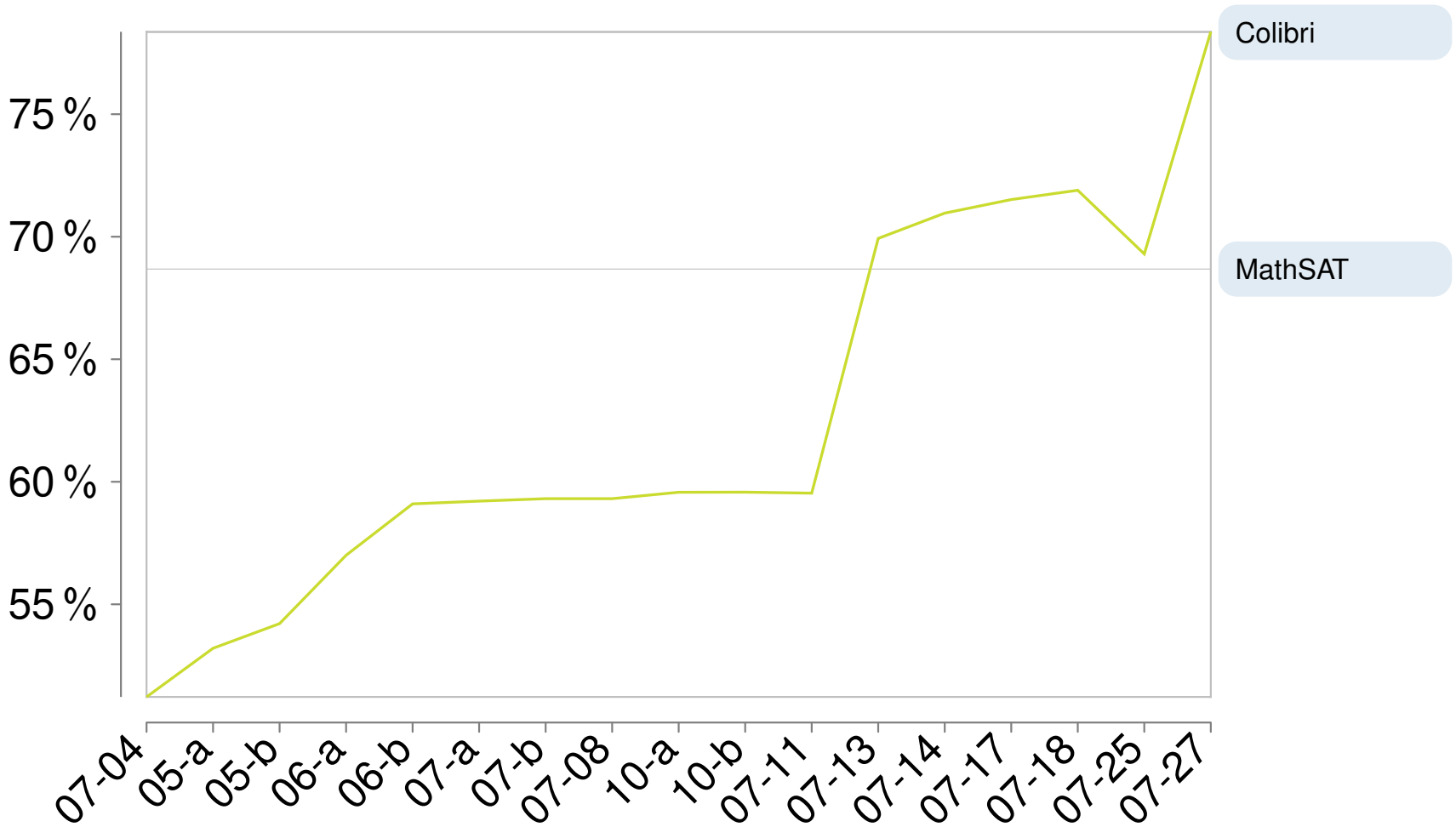
FP progress in CVC4

Unsound

Benchmark	07-04	07-27
Schanda	8	1
Heizmann	✓	✓
NyxBrain	✓	✓
PyMPF	818	1427
SPARK 2014	✓	✓
Wintersteiger	351	28

FP progress in CVC4

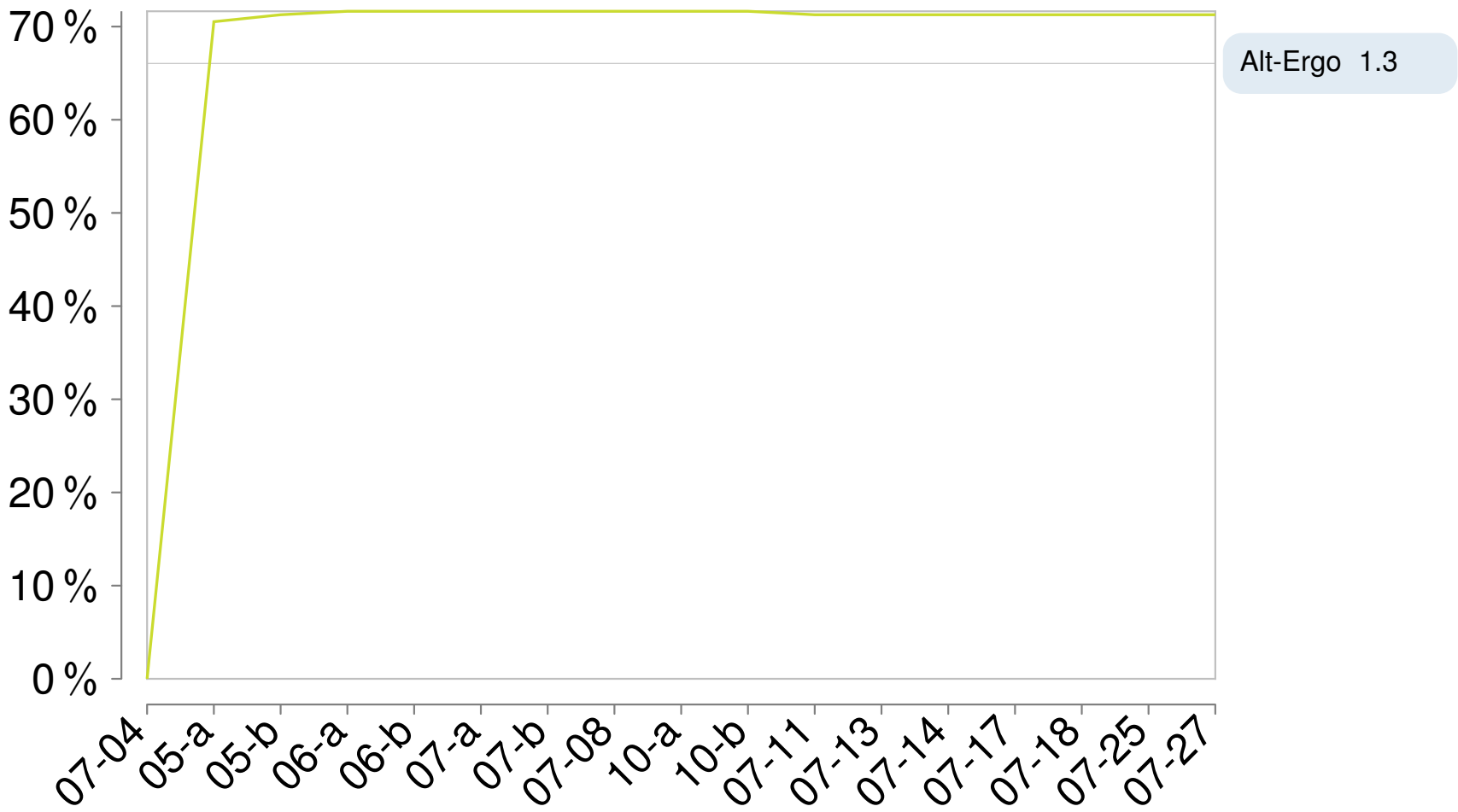
Solved over time (average of averages)



ALTRAN

FP progress in CVC4

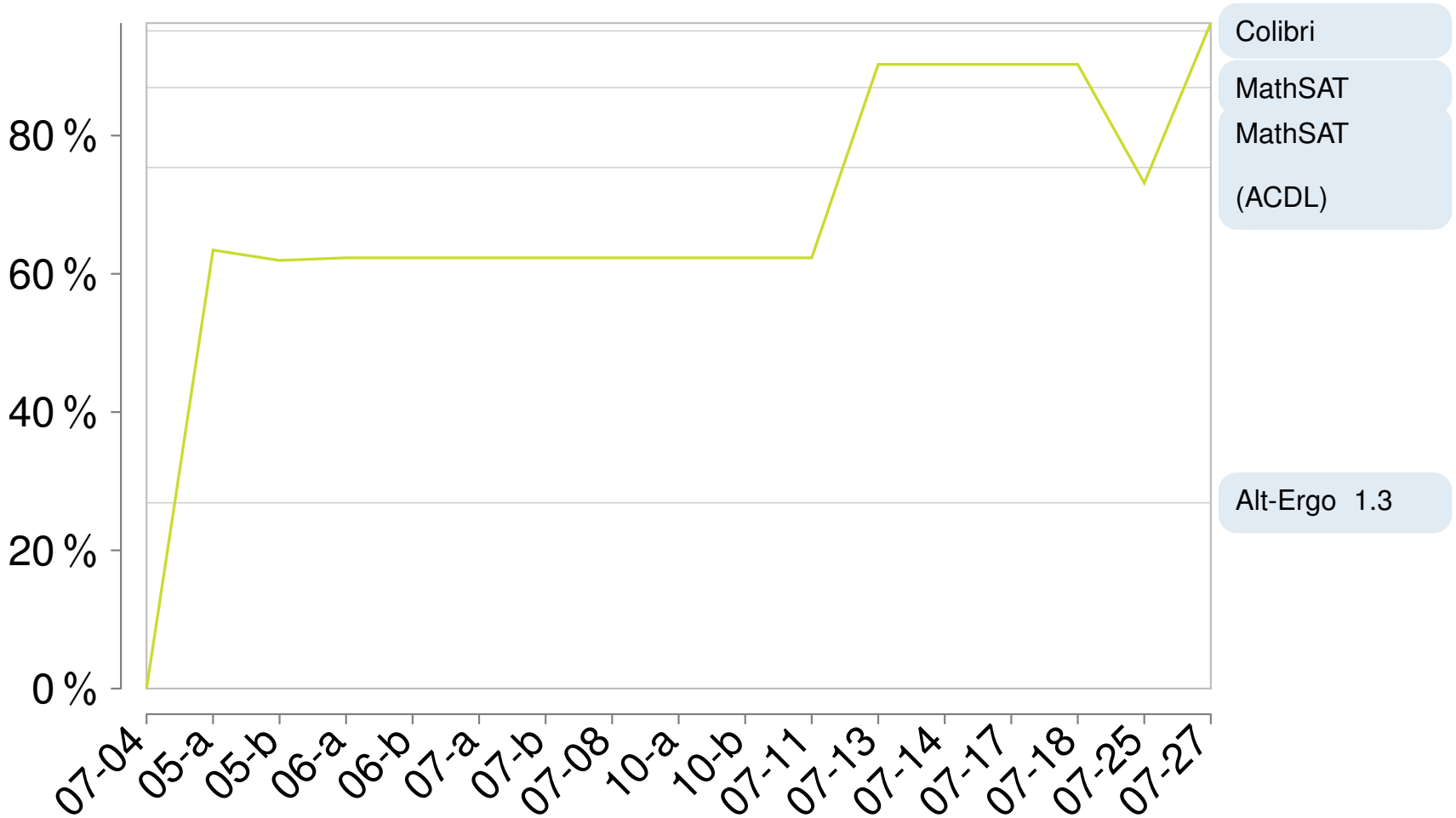
VCs solved on Industrial_1



ALTRAN

FP progress in CVC4

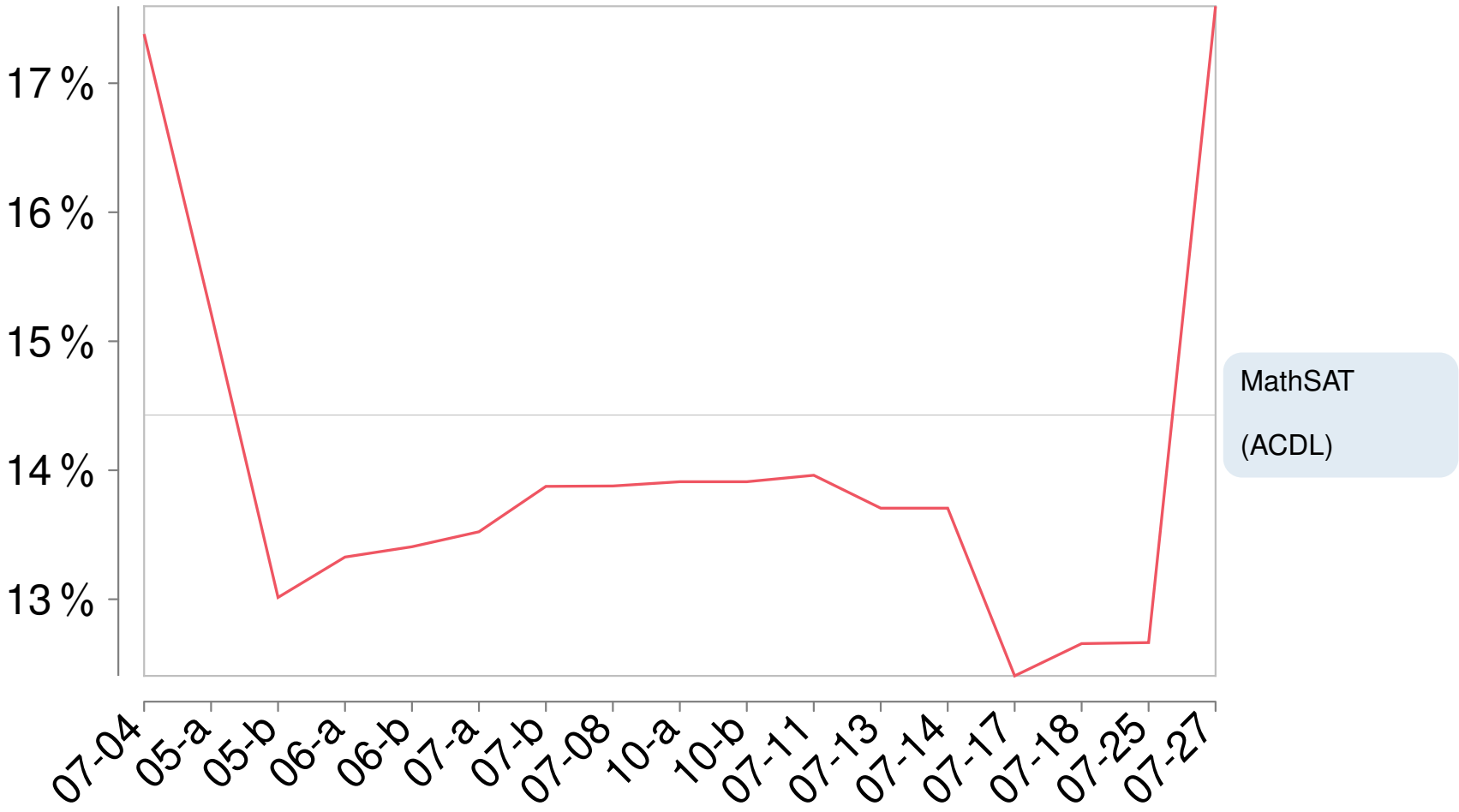
VCs solved on Industrial_1_qf



altran

FP progress in CVC4

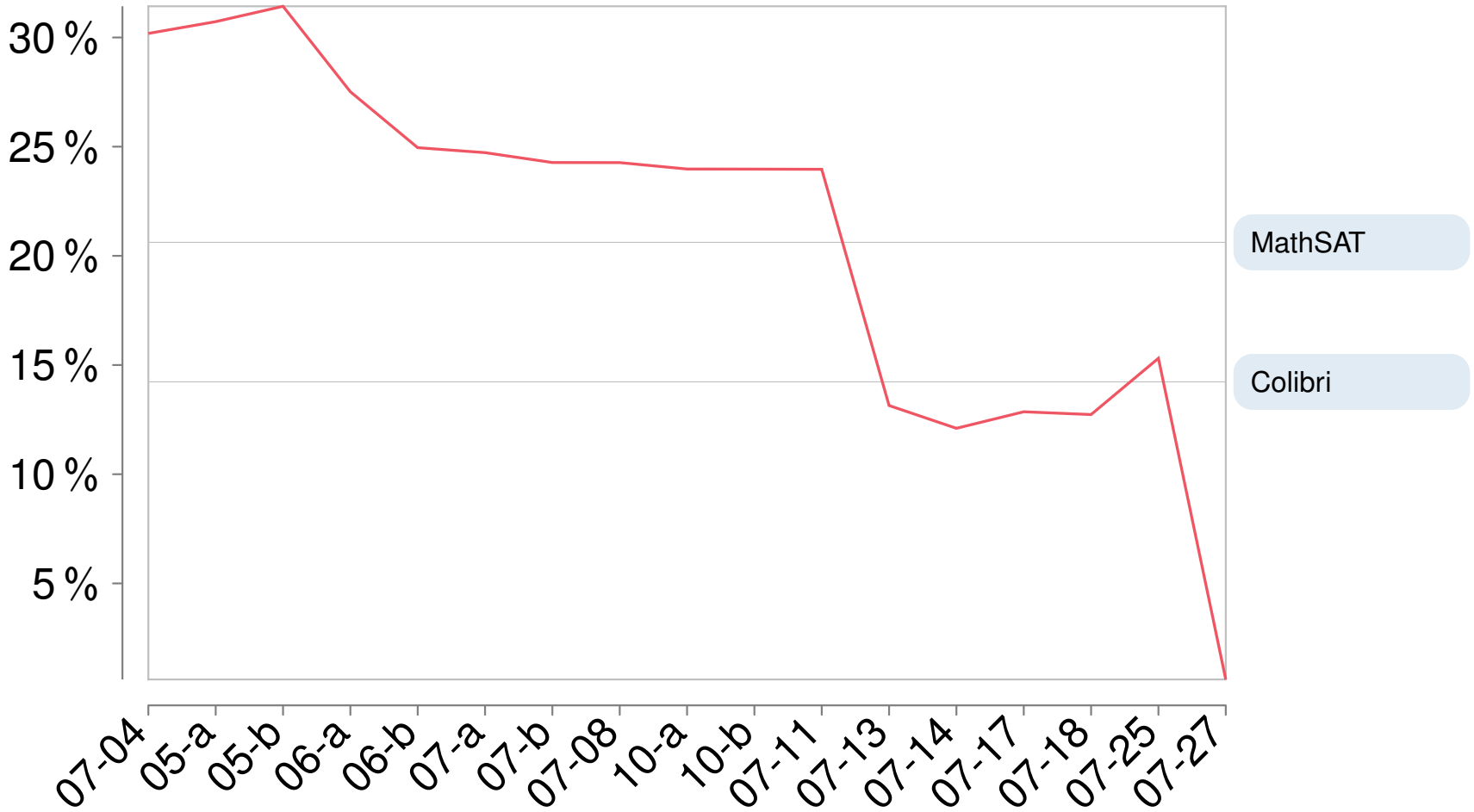
Timeout over time (average of averages)



altran

FP progress in CVC4

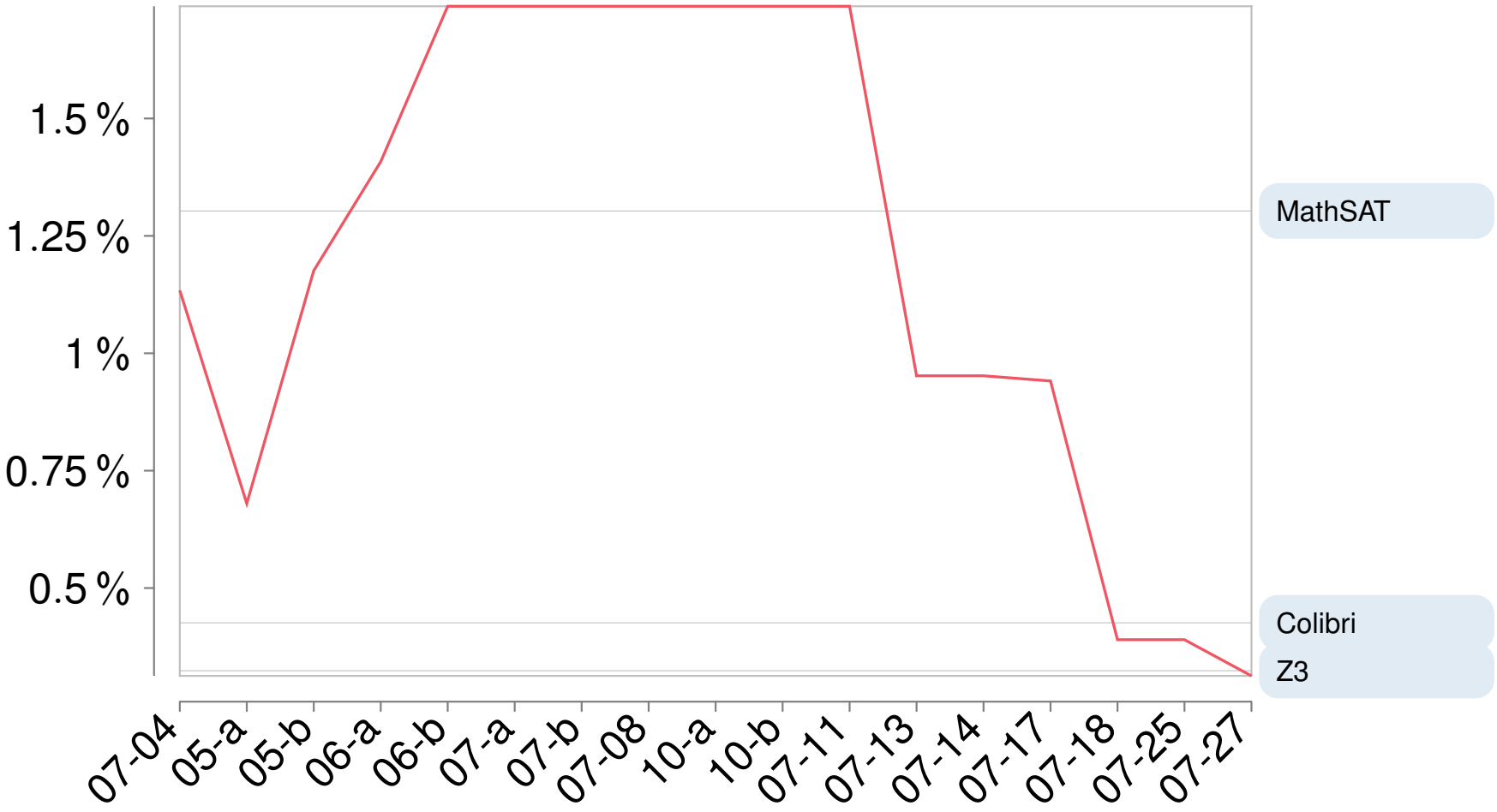
Error over time (average of averages)



ALTRAN

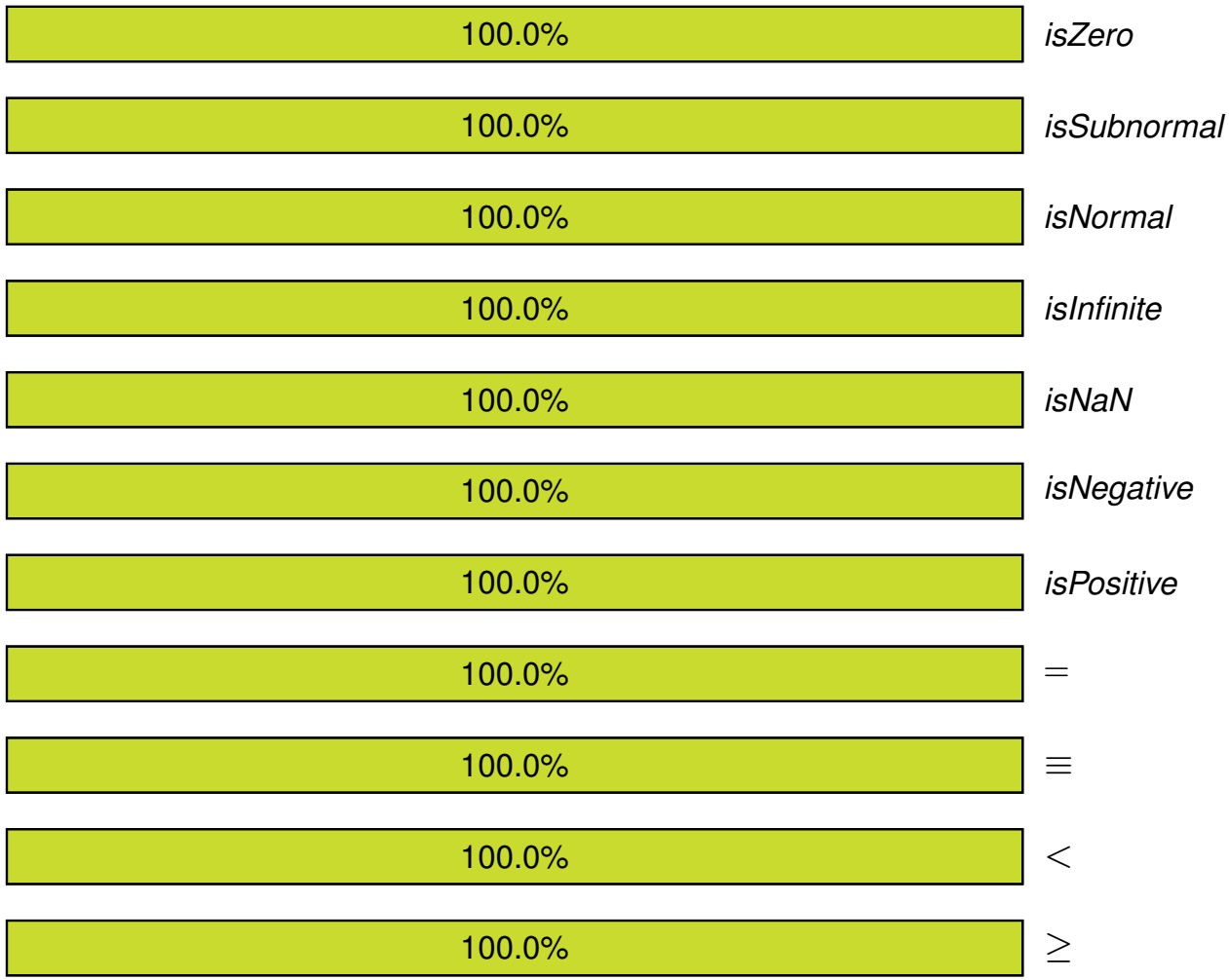
FP progress in CVC4

Unsound over time (average of averages)

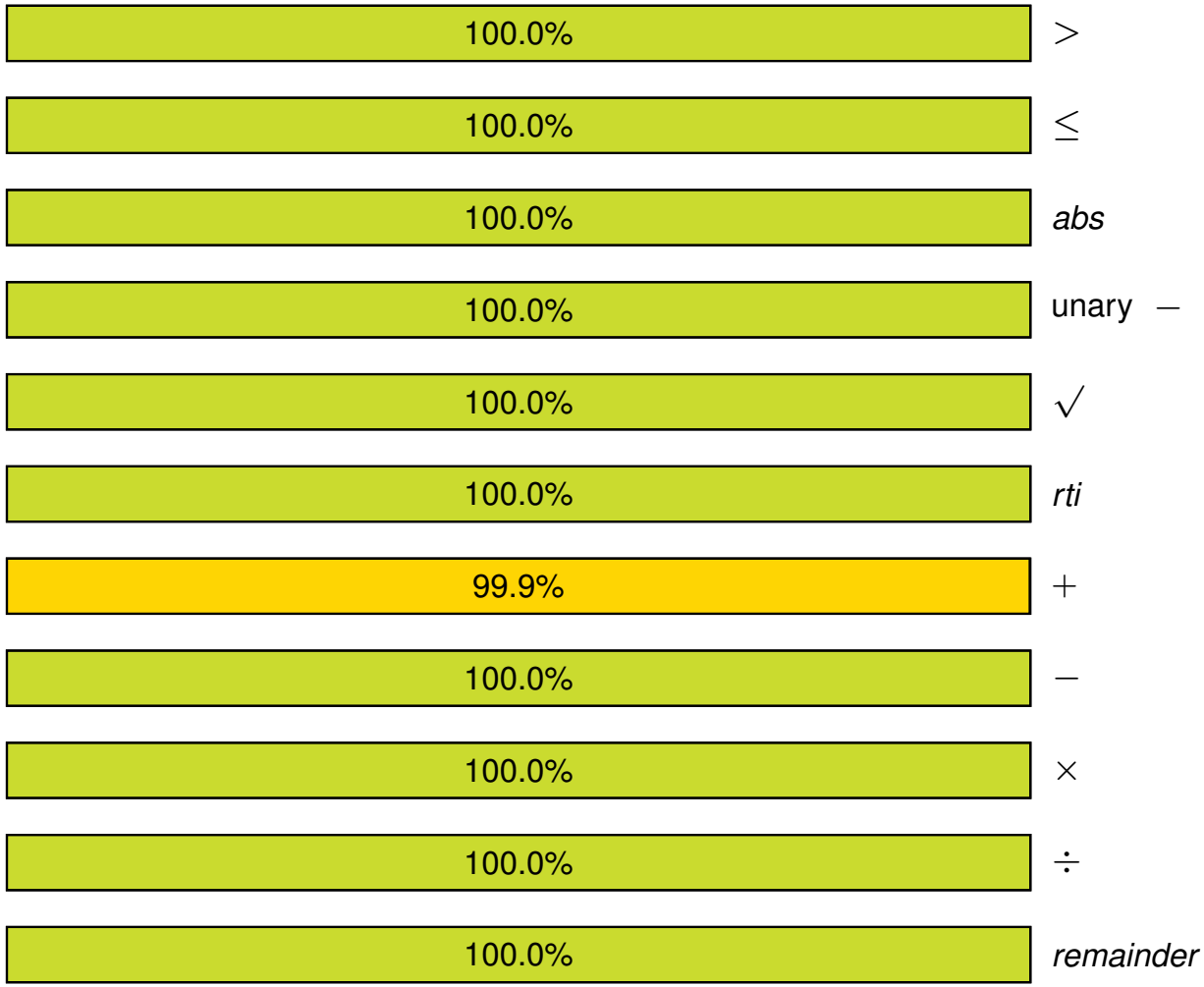


ALTRAN

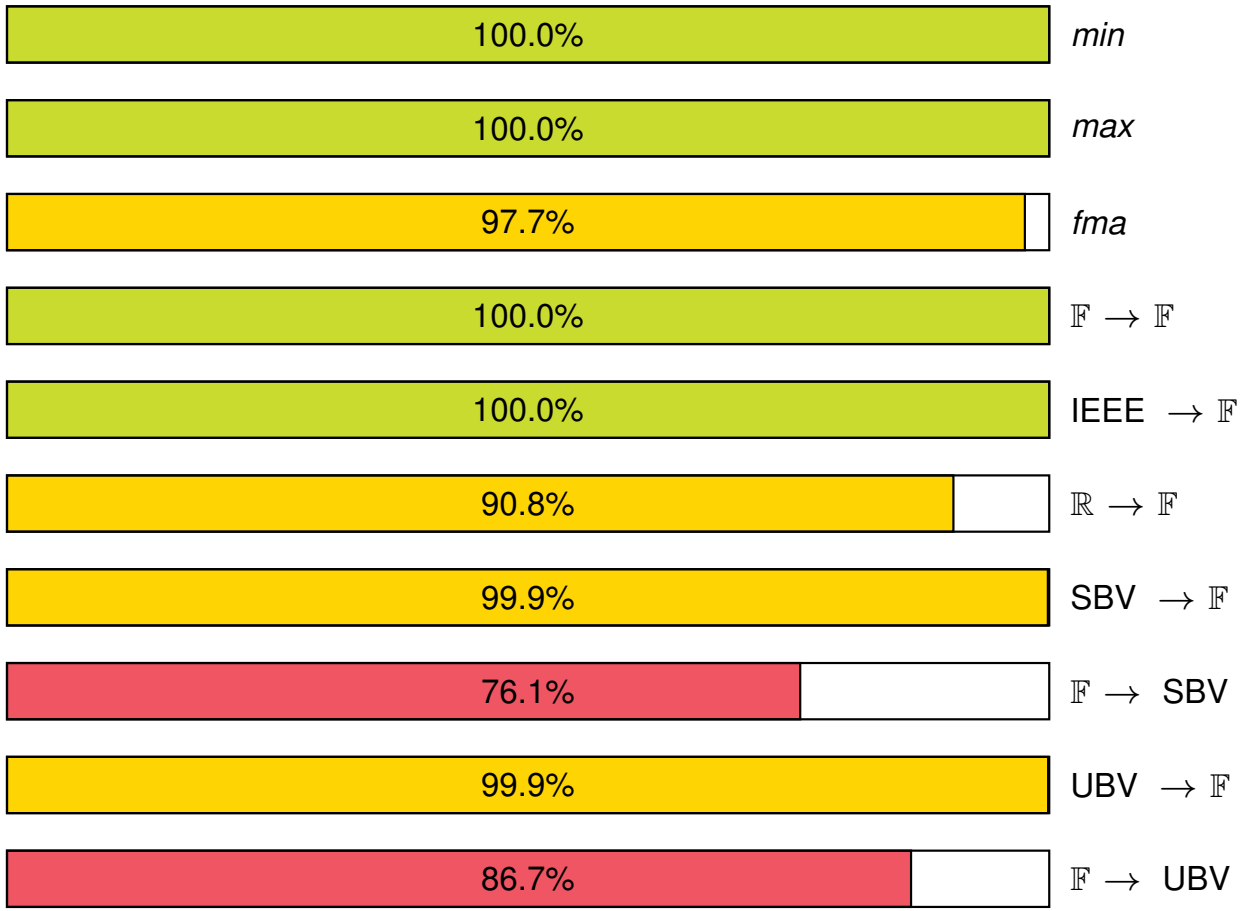
Critical Success Factors for CVC4



Critical Success Factors for CVC4



Critical Success Factors for CVC4



2.

Comparisons

Benchmarks

With status ‘solved’

Benchmark	Alt-Ergo 1.3	Colibri	CVC4	MathSAT	MathSAT (ACDL)	Z3
Schanda		78.1%	67.4%	72.5%*	28.7%*	84.3%
Griggio		63.6%	17.3%	39.7%	22.9%	30.4%
Heizmann		12.1%	69.1%	58.5%	26.6%	31.9%
Industrial_1	66.0%*		71.3%			74.3%
Industrial_1_qf	26.9%*	95.1%	96.3%	86.9%	75.4%	97.0%
NyxBrain		99.2%	99.8%	95.4%	95.0%	99.9%
PyMPF		83.5%	97.4%	30.9%	26.4%	98.9%
SPARK 2014	54.3%*		79.0%		0.0%	82.5%*
SPARK QF		95.1%	86.2%	79.7%	70.2%	90.3%
Wintersteiger		✓	99.9%	85.8%	85.8%	✓
Summary	49%	78%	78%	69%	48%	79%

*) uses different VCs

Benchmarks

With status ‘timeout’

Benchmark	Alt-Ergo 1.3	Colibri	CVC4	MathSAT	MathSAT (ACDL)	Z3
Schanda		21	55	23*	51*	23
Griggio		77	177	129	165	149
Heizmann		0	14	0	2	141
Industrial_1	91*		73			69
Industrial_1_qf	0*	7	10	0	28	8
NyxBrain		0	96	20	128	56
PyMPF		0	6	0	0	0
SPARK 2014	1127*		314		1	357*
SPARK QF		65	289	48	306	202
Wintersteiger		0	0	0	0	0
Summary	27%	7%	18%	9%	14%	20%

*) uses different VCs

Benchmarks

With status ‘error’

Benchmark	Alt-Ergo 1.3	Colibri	CVC4	MathSAT	MathSAT (ACDL)	Z3
Schanda		10	2	18*	42*	✓
Griggio		1	✓	✓	✓	✓
Heizmann		182	2	86	150	✓
Industrial_1	✓*		3			✓
Industrial_1_qf	✓*	5	✓	35	38	✓
NyxBrain		388	4	64	120	4
PyMPF		8758	45	38671	38671	✓
SPARK 2014	✓*		15		2467	✓*
SPARK QF		47	52	452	429	✓
Wintersteiger		✓	✓	5668	5668	✓
Summary	0%	14%	1%	21%	34%	0%

*) uses different VCs

Benchmarks

With status ‘unsound’

Benchmark	Alt-Ergo 1.3	Colibri	CVC4	MathSAT	MathSAT (ACDL)	Z3
Schanda		4	1	8*	34*	4
Heizmann		✓	✓	✓	✓	✓
NyxBrain		6	✓	2318	2366	12
PyMPF		658	1427	866	3408	553
SPARK 2014	✓*		✓		✓	✓*
Wintersteiger		✓	28	✓	1	✓
Summary	0%	0%	0%	1%	3%	0%

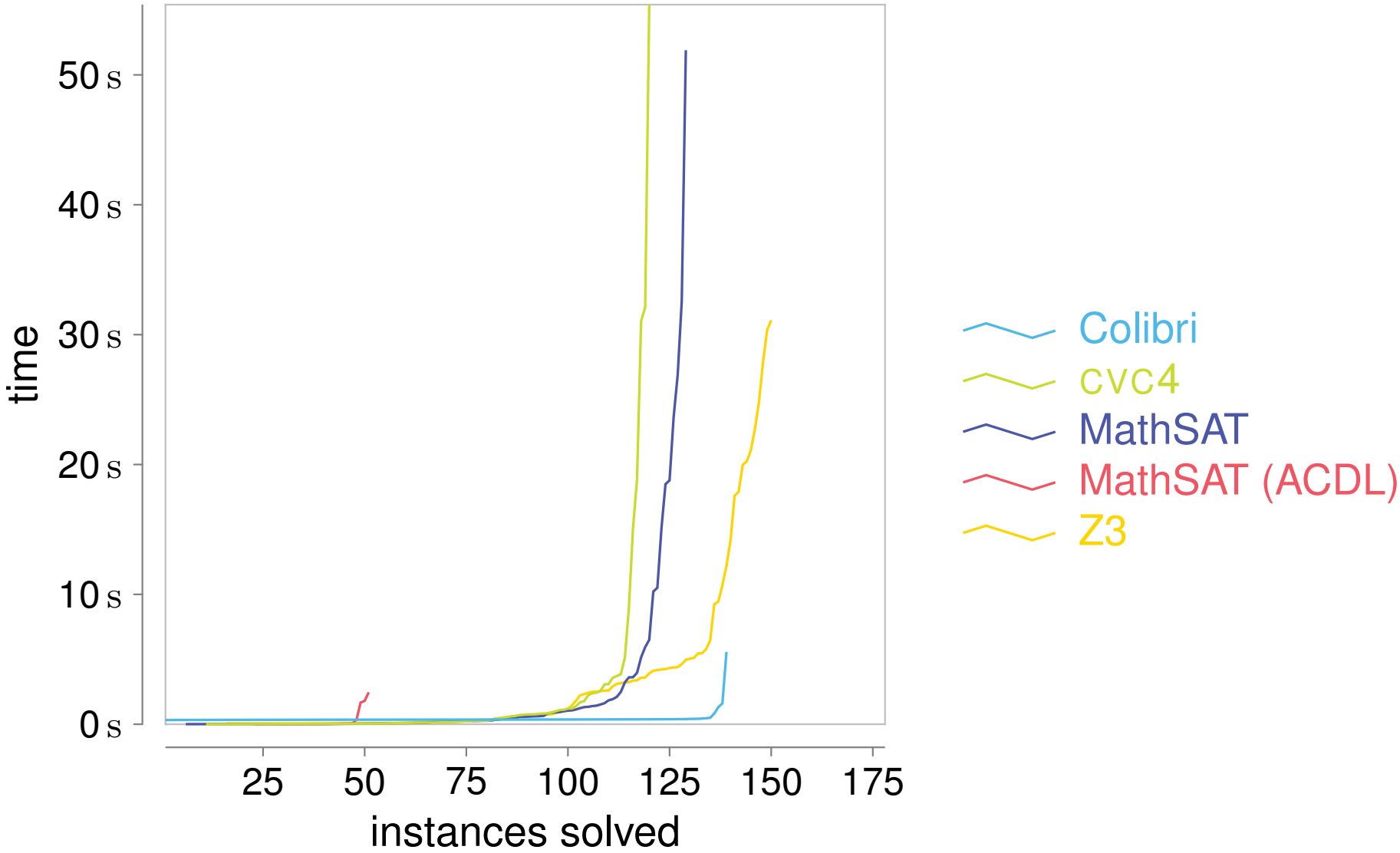
*) uses different VCs

3.

Cactus plots

Cactus plot

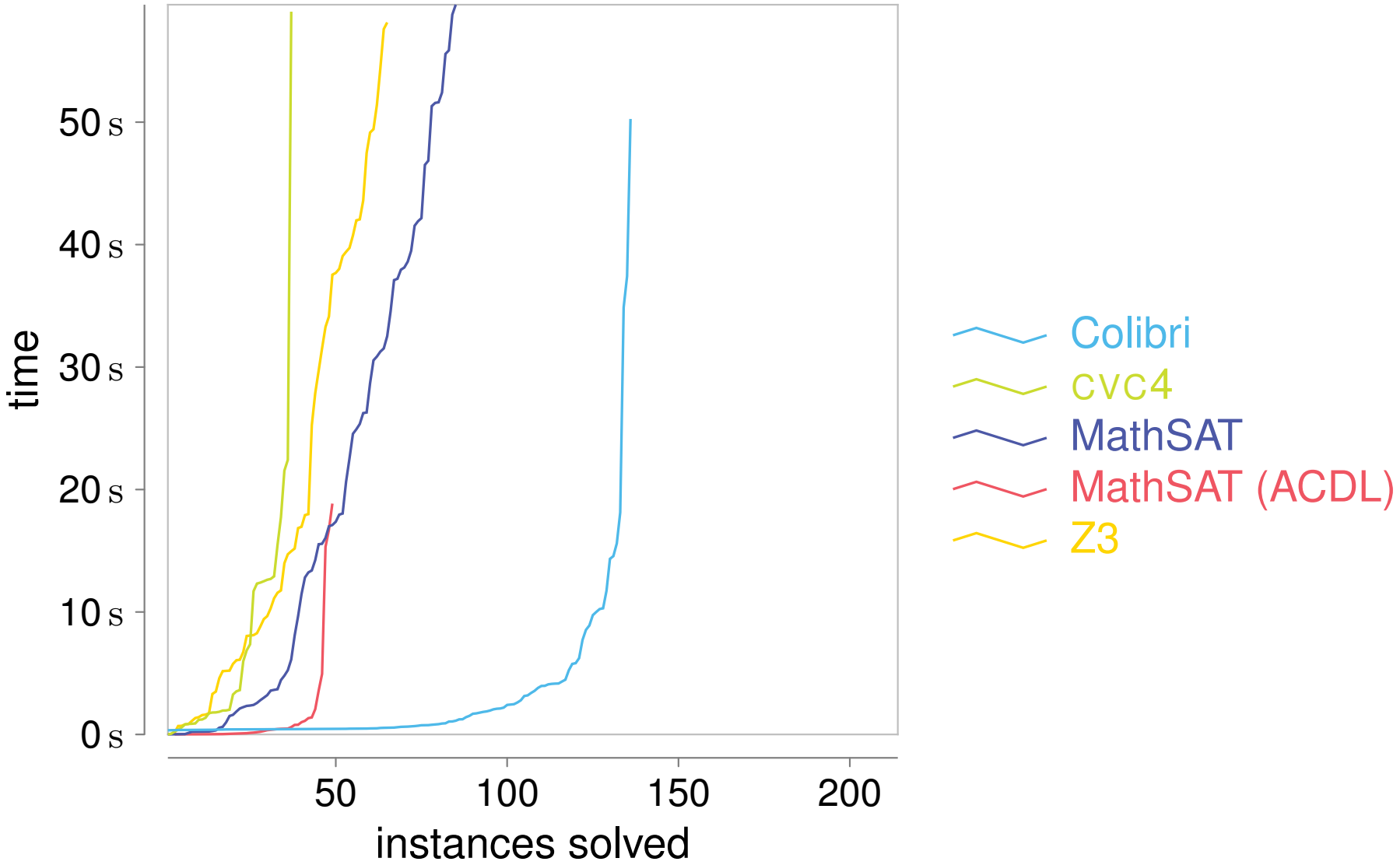
Schanda



ALTRAN

Cactus plot

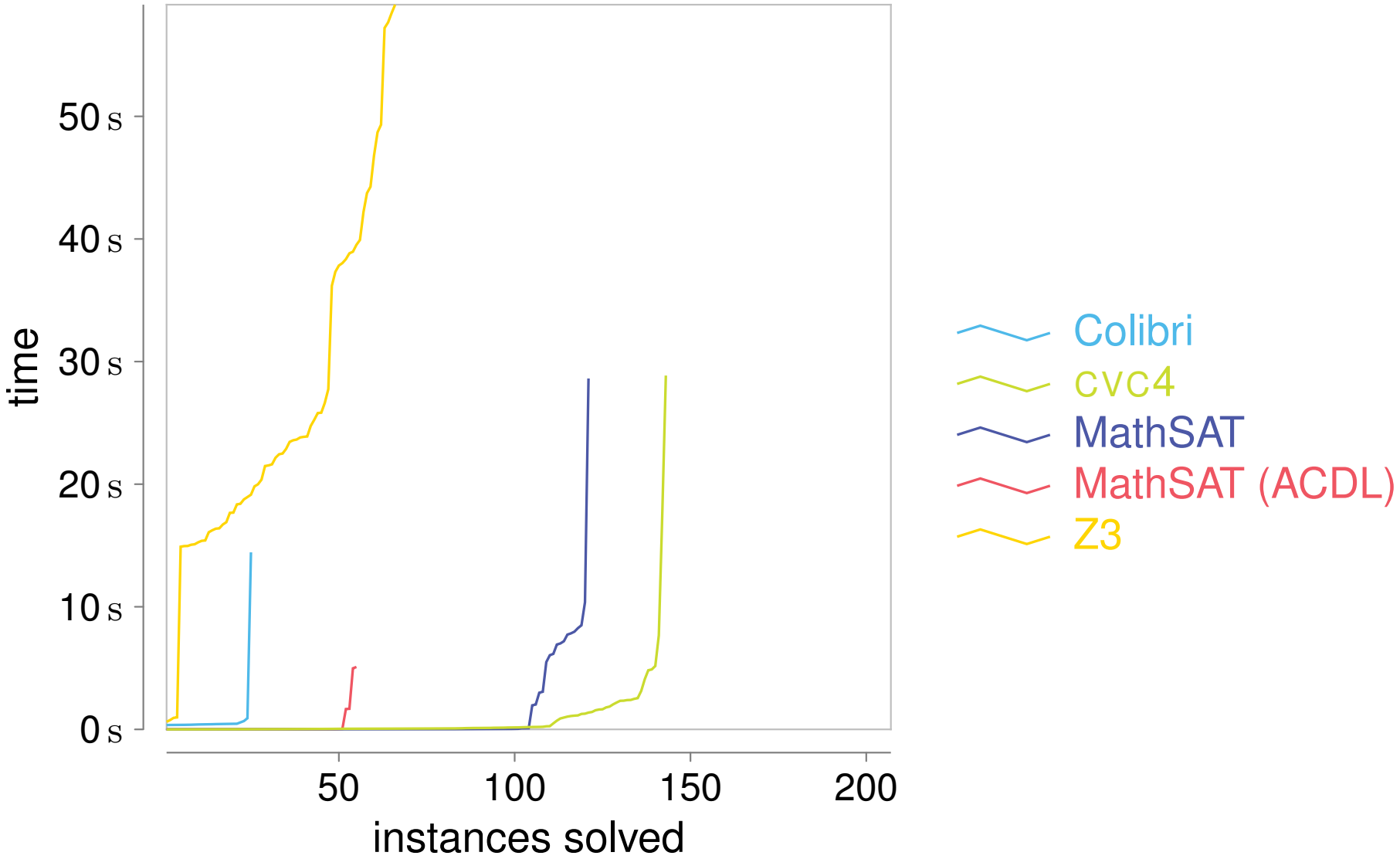
Griggio



altran

Cactus plot

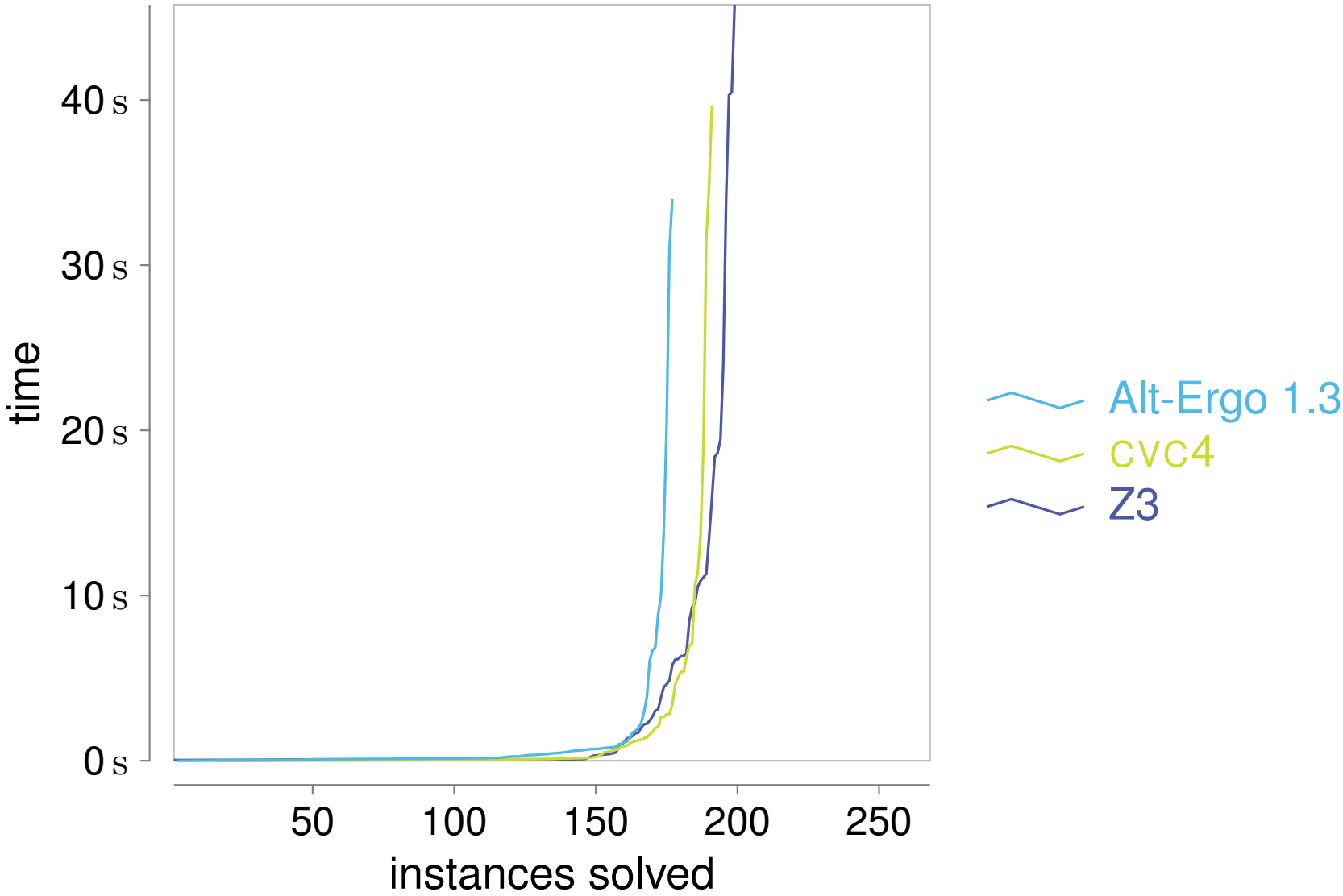
Heizmann



ALTRAN

Cactus plot

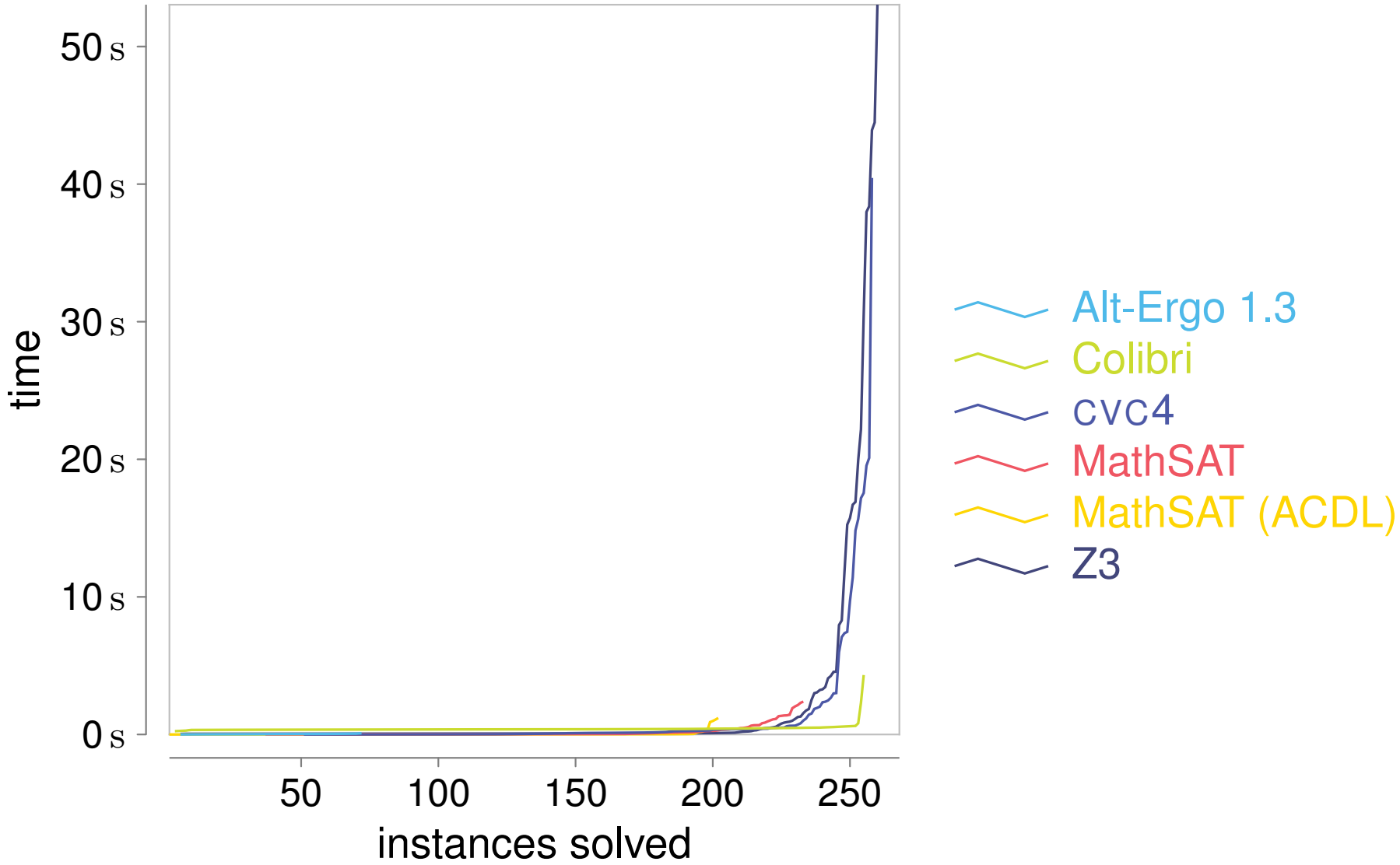
Industrial_1



ALTRAN

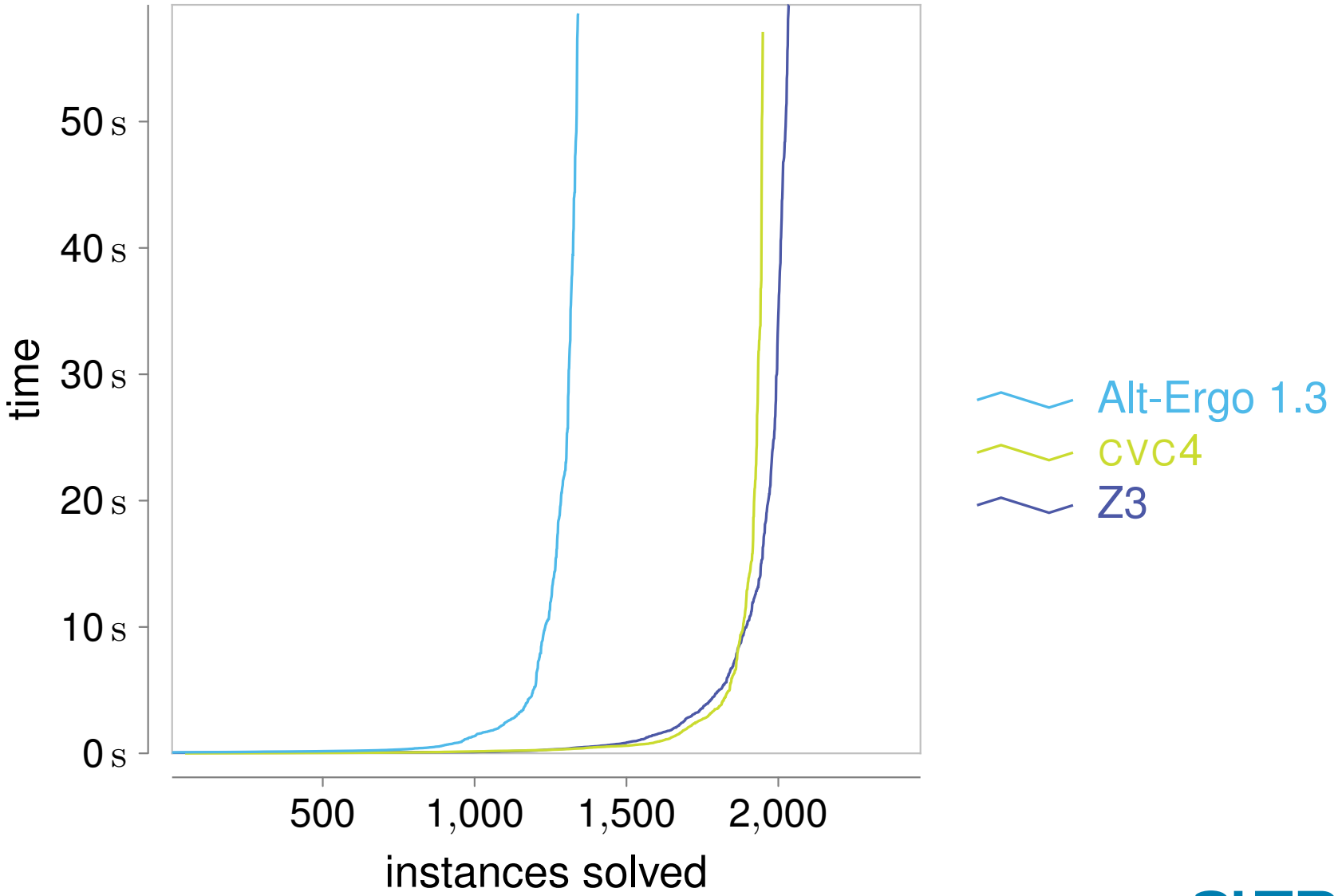
Cactus plot

Industrial_1_qf



Cactus plot

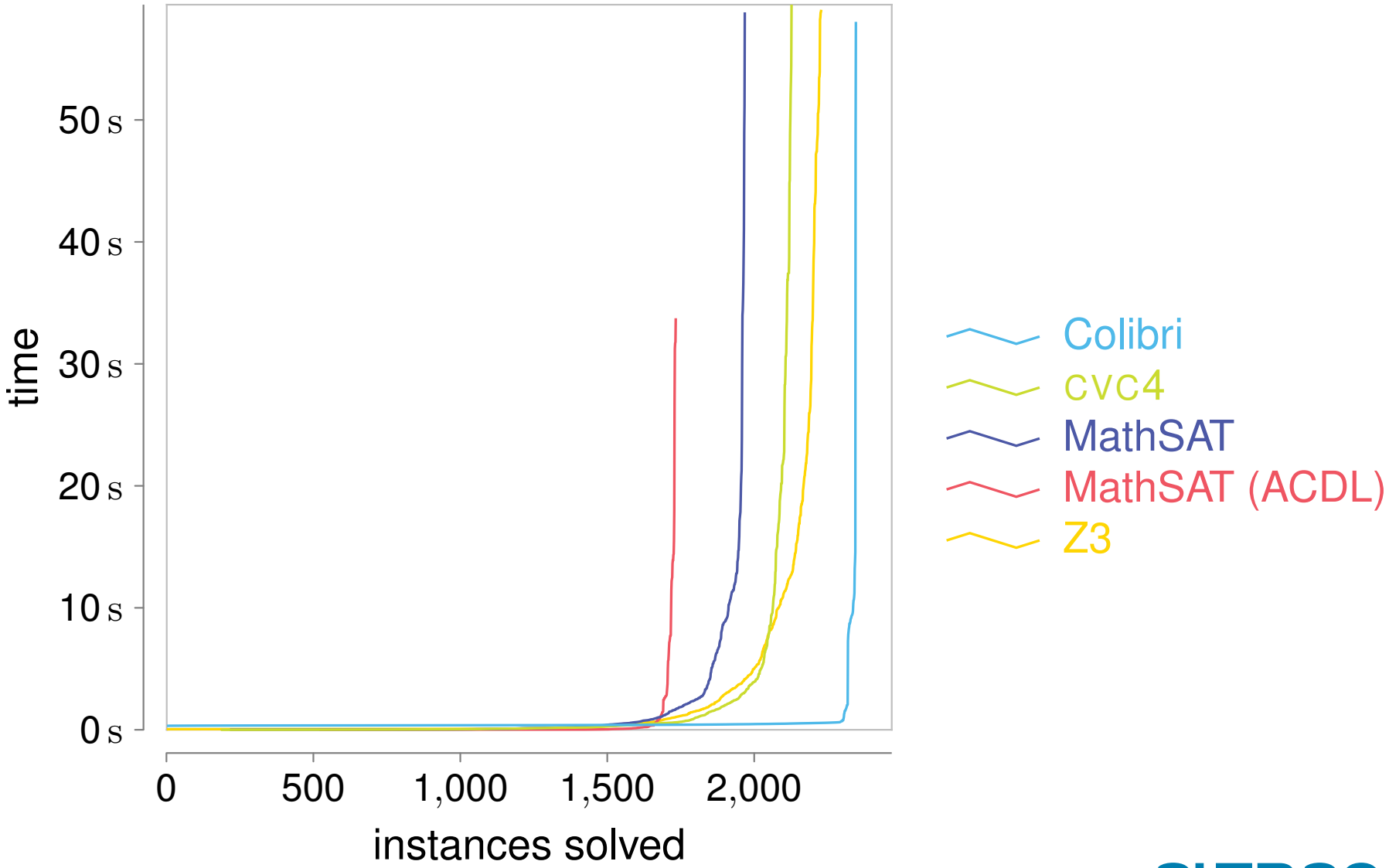
SPARK 2014



ALTRAN

Cactus plot

SPARK QF



ALTRAN