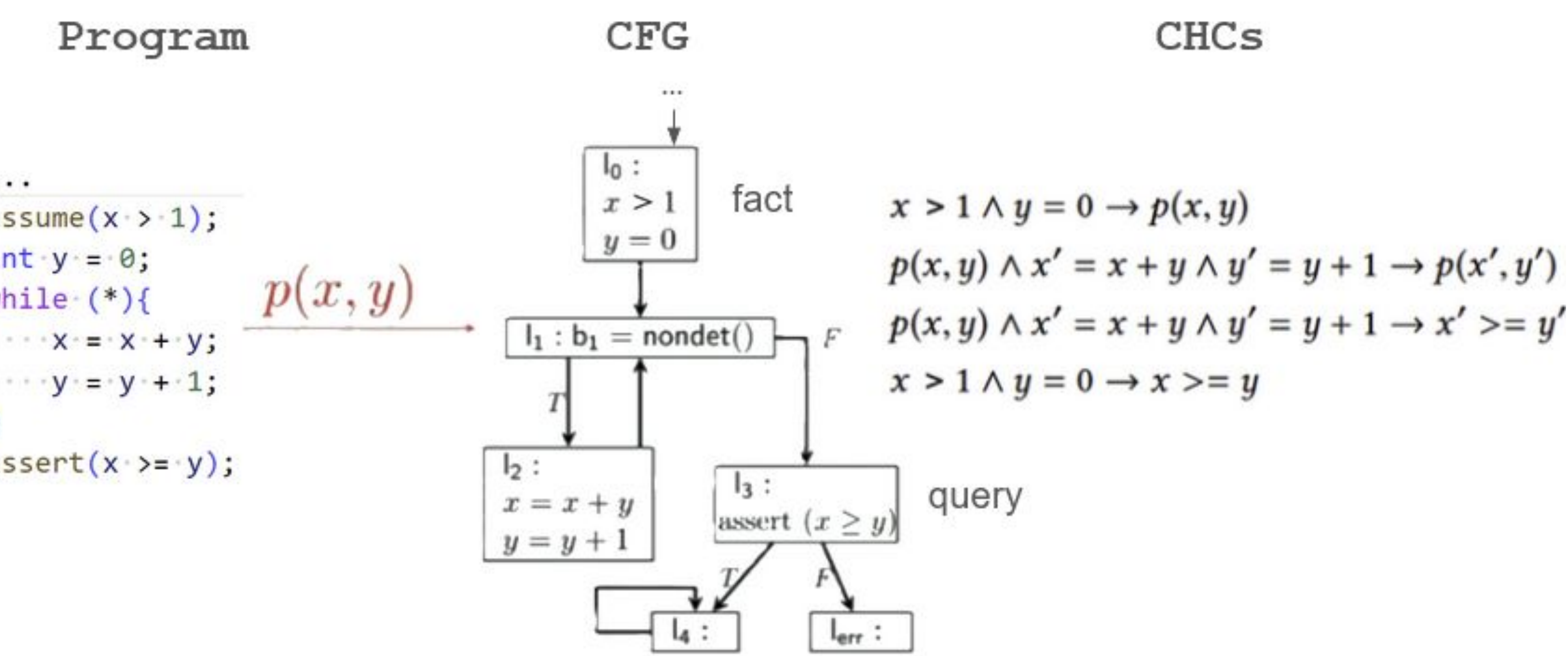# Chronosymbolic Learning

Ziyan "Ray" Luo[12], Xujie Si[13]

[1] Mila, [2] McGill University,
[3] University of Toronto

BACKGROUND

- Program verification => CHC solving (a fragment of FOL)
- Black-box data-driven methods are sample-inefficient and agnostic to the symbolic information within the CHCs
- White-box symbolic approaches struggles to identify essential patterns from data samples

CONTRIBUTION

- Design a modular framework, Chronosymbolic Learning, to efficiently leverage symbolic information (*zones, representing a set of data samples*) and data *samples*
- Propose a minimal instance (SVM+DT+BMC), exemplifying how components interact in our framework (artifact available)
- A new perspective of unifying approaches (see paper for more)
- Promising experimental result though simplistic instantiation

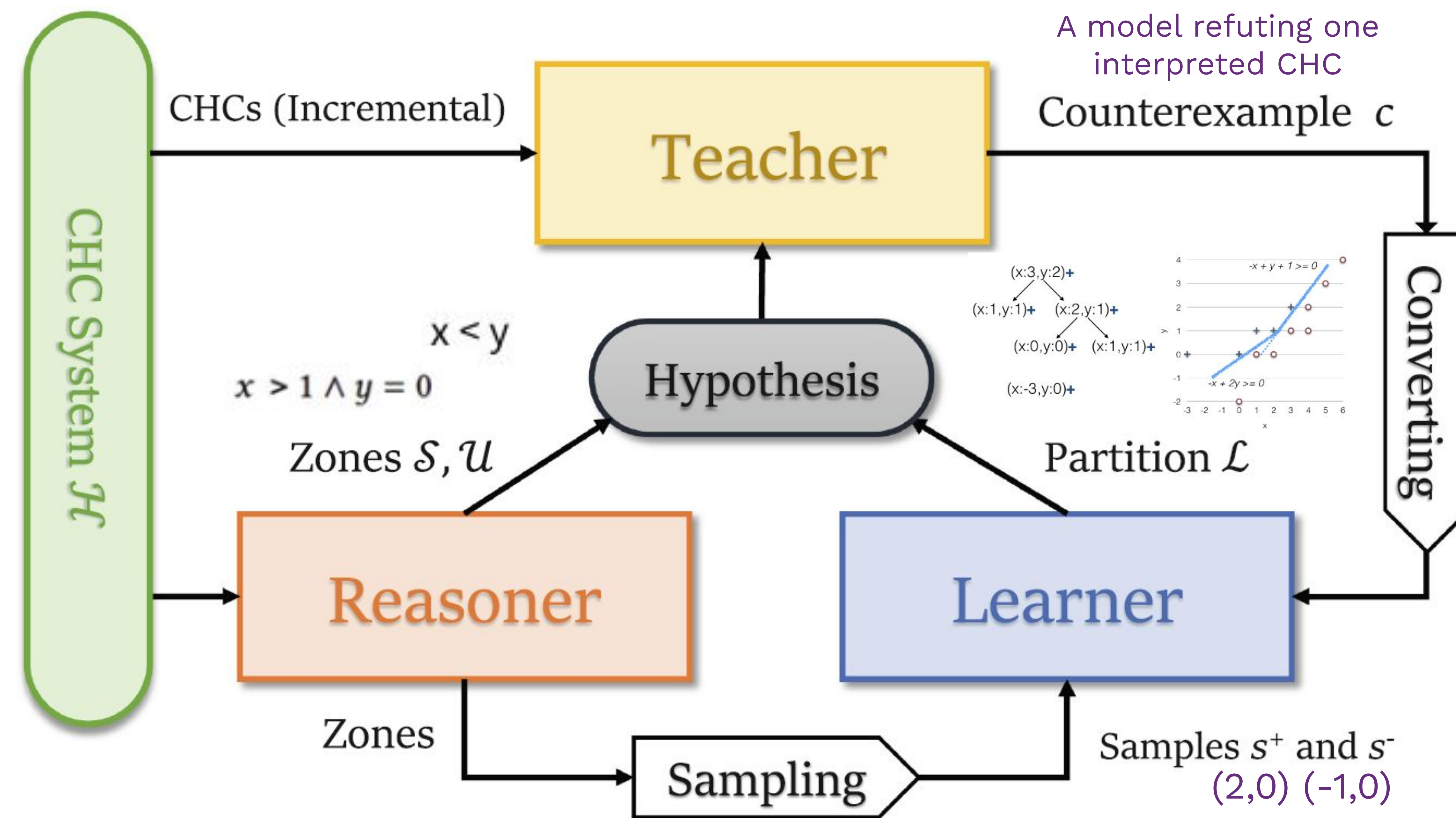Overview of different approaches through the lens of learning from samples and zones.
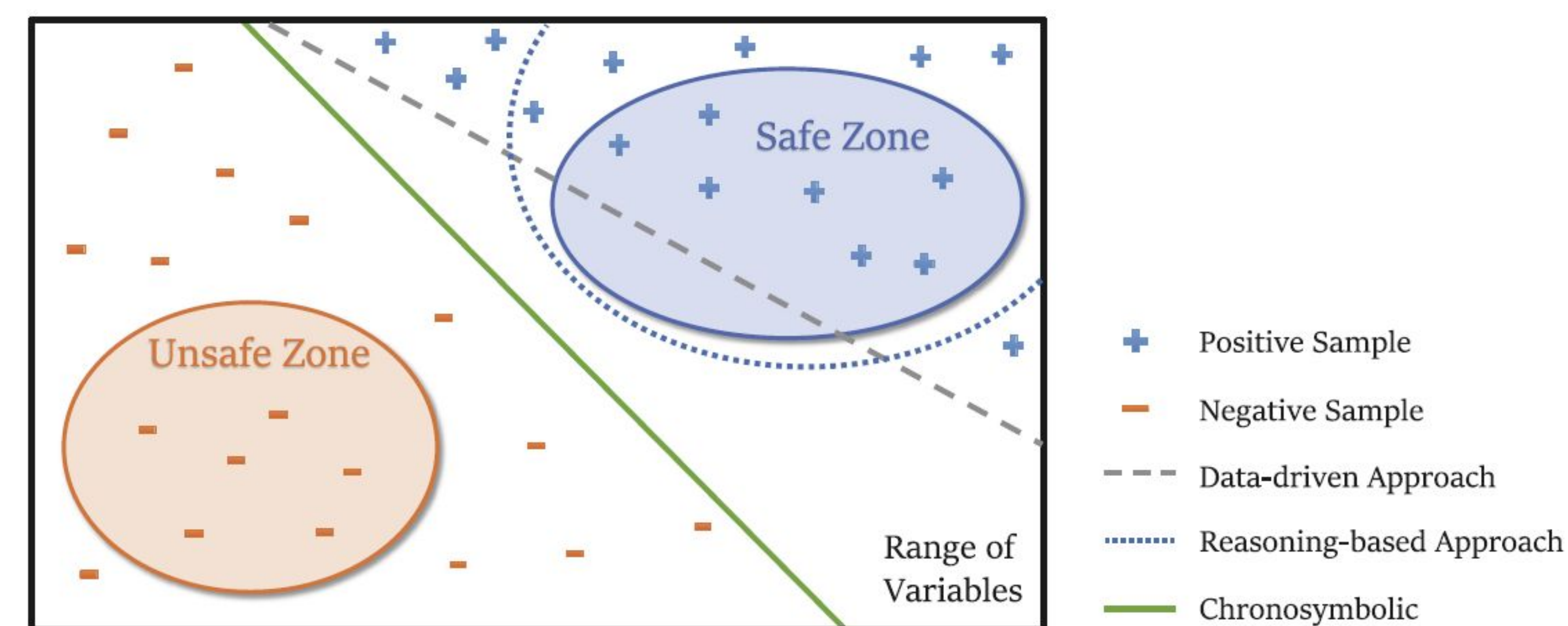
"<u>CH</u>ro*no*symboli<u>C</u>":
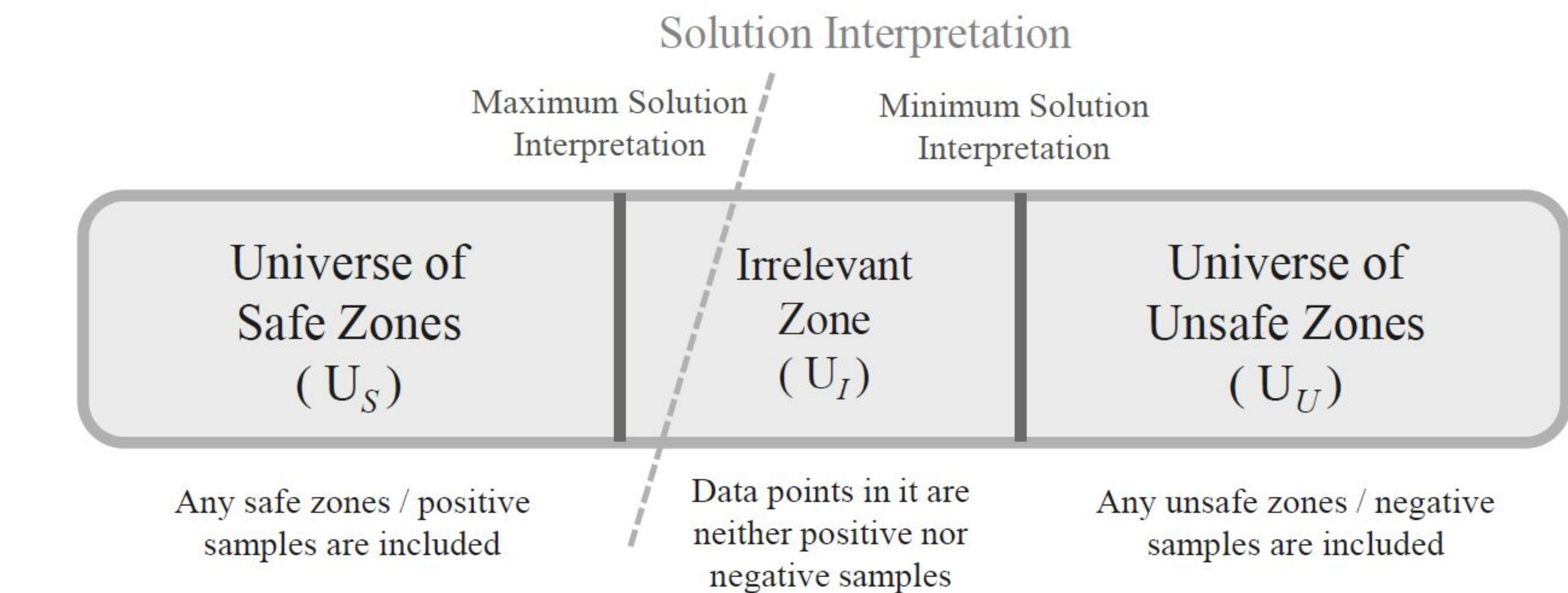## Constrained Horn Clause

Efficient CHC solver =

Symbolic Reasoner + Data-driven Learner
Zones   Synchronously   Samples

Paper: Github:
(with running examples)

SOLUTION SPACE

MAIN RESULT

- 288 arithmetic instances collected by FreqHorn
- Timeout = 360s
- Chronosymbolic-single: one config for all instances
- Chronosymbolic-cover: for each instance, best result over 13 configs
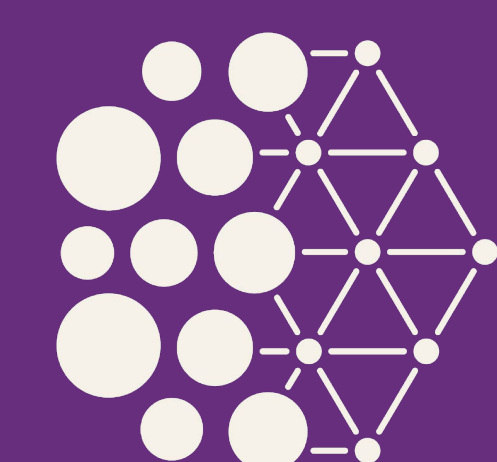- parallel: learner and reasoner running individually and simultaneously for 360s

Method	#total	percentage	#safe	#unsafe
LinearArbitrary	187	64.93%	148	39
FreqHorn	191	66.32%	191	0
FreqHorn-expl	50	17.36%	0	50
Spacer	184	63.89%	132	52
GSpacer	220	76.39%	174	46
Chronosymbolic-single	**237**	**82.29%**	**189**	**48**
Chronosymbolic-cover	**252**	**87.50%**	**204**	**48**

Configuration	#total	percentage	#safe	#unsafe
without safe zones	228	79.17%	183	45
without unsafe zones	218	75.69%	173	45
without both zones	211	73.26%	166	45
without learner	131	45.49%	96	35
parallel	216	75.00%	180	36
Chronosymbolic-single	**237**	**82.29%**	**189**	**48**