# WAIT 2018

## Cyclic Superposition and Induction

Jannik Vierling

June 29, 2018

Joint work with Stefan Hetzl

# Background

Induction in Computer Science and Mathematics

Computer Science:

- Automation of proof by induction
- Interest in various inductive datatypes
- Focus on algorithms and efficiency

Mathematics:

- Interest in theories of arithmetic of natural numbers
- Consistency proofs
- Unprovability results

$\rightsquigarrow$ Independent views of induction, little interaction

# Motivation: Research Program

Automation of proof by induction

- Variety of approaches: Cyclic proofs, Rippling, ...
- Heuristics
- Empirical analysis
- Few general, formal results

Central questions:

1. What can('t)/should(n't) a given method prove?
2. Why does(n't) a method work well?
3. How do methods relate to each other?

$\rightsquigarrow$ Rigorous analysis of calculi

$\rightsquigarrow$ Apply techniques and results from mathematical logic

$\rightsquigarrow$ Developement of formal foundations

# Goals: This Talk

Analysis of semantic clause cycles inspired by n-clause calculus [1]

- analysis tool
- abstracts concrete calculi
- refutational
- clausal
- induction cycles

Goals

- Describe the induction captured by the system

  $\rightsquigarrow$ Quantifier complexity of induction invariants
- Necessity

  Is the captured type of induction necessary?
- Completeness

[1] Kersani and Peltier 2013

# Outline

# Semantic Clause Set Cycles

$\Theta$ ... set of sorts,
$\mathcal{F}$ ... set of f.s. containing $0 : \omega$, $\mathsf{s} : \omega \to \omega$ over $\Theta \cup \{\omega\}$,
$\mathcal{P}$ ... set of p.s. over $\Theta \cup \{\omega\}$.

Consider clauses over $\mathsf{n} : \omega, =/2, \mathcal{F}$ and $\mathcal{P}$

## Definition

A semantic clause set cycle (SCSC) for $S(\mathsf{n})$ is a triple $\langle S'(\mathsf{n}), i, j \rangle$ with $i, j \in \mathbb{N}, j > 0$ s.t.

$$S(\mathsf{n}) \models_{\mathrm{FOL}} S'(\mathsf{n}), \quad (1)$$

$$S'(\mathsf{s}^j \mathsf{n}) \models_{\mathrm{FOL}} S'(\mathsf{n}), \quad (2)$$

$$S'(\overline{i+k}) \models_{\mathrm{FOL}}, \text{for } k = 0, \ldots, j-1. \quad (3)$$

## Definition

$S(\mathsf{n})$ is refutable with semantic clause set cycles if $S(\mathsf{n})$ has an SCSC $(S'(\mathsf{n}), i, j)$ and $S(\overline{k}) \models_{\mathrm{FOL}}$ for all $0 \leq k < i$ (4).

# Semantic Clause Set Cycles: Soundness

Std. Semantics ... FOL Semantics restricted to $\mathcal{M}$ with $|\mathcal{M}|_\omega = \mathbb{N}$

## Proposition (Soundness)

*If $S(\mathsf{n})$ is refutable with SCSCs, then $S(\mathsf{n})$ is unsatisfiable in std. semantics.*

## Proof Sketch.

Proceed by infinite descent

$$
\begin{array}{lll}
(\mathsf{n}^{\mathcal{M}} < i) & : \mathcal{M} \models_{\text{STD}} S(\overline{k}) \models_{\text{FOL}}, & (\bot) \\
(i \leq \mathsf{n}^{\mathcal{M}} < i + j) & : \mathcal{M} \models_{\text{STD}} S(\mathsf{n}^{\mathcal{M}}) \models_{\text{FOL}} S'(\mathsf{n}^{\mathcal{M}}) \models_{\text{FOL}} & (\bot) \\
(i + j \leq \mathsf{n}^{\mathcal{M}}) & : \mathcal{M} \models_{\text{STD}} S(\mathsf{n}) \models_{\text{FOL}} S'(\mathsf{n}), & \\
 & \rightsquigarrow \mathcal{M} \cup \{\mathsf{n} \mapsto \mathsf{n}^{\mathcal{M}} - j\} \models S'(\mathsf{s}^j \mathsf{n}) \models S'(\mathsf{n}). & (\circlearrowright)
\end{array}
$$

□

# Semantic Clause Set Cycles: $\Sigma_1$-Bound

### Theorem ($\Sigma_1$-Bound)

*If $S(\mathrm{n})$ is refutable with SCSCs, then $S(\mathrm{n})$ is refutable in* $\mathbf{LK} + \Sigma_1$*-induction.*

### Proof Sketch.

- Derive the formula $\forall x \exists y\ (x = 0 \lor x = \mathrm{s}(y))$
- Generate subgoals $S(0), \ldots, S(\overline{i-1})$, and $S(\mathrm{s}^i\alpha)$
- Subgoals $S(0), \ldots, S(\overline{i-1})$ are trivial (FOL-unsat)
- Reduce $S(\mathrm{s}^i\alpha)$ to $S'(\mathrm{s}^i\alpha)$
- Refute $S'(\mathrm{s}^i\alpha)$ by induction on $\neg S'(\mathrm{s}^i\alpha)$. $\square$

$\longrightarrow$ Restriction to $\Sigma_1$-induction due to clause normal form.

# Outline

# $\Sigma_1$-Necessity

## Theorem

*There exists a clause set $S(\mathsf{n})$ which is refutable with SCSCs, and is not refutable in* **LK** $+$ *quantifier-free induction.*

## Definition

Let $\mathrm{opt}_{\Sigma_1}$ consist of the following clauses

$$\Gamma := \begin{cases} 0 \neq \mathsf{s}(x), \\ \mathsf{s}(x) = \mathsf{s}(y) \rightarrow x = y, \\ x{+}0 = x, \\ x{+}\mathsf{s}(y) = \mathsf{s}(x{+}y), \\ \mathsf{p}(0, \bar{1}), \\ \mathsf{p}(x, y) \rightarrow \mathsf{p}(\mathsf{s}(x), y{+}y), \end{cases}$$

$$\neg\mathsf{p}(\mathsf{n}, y).$$

# $\Sigma_1$-Necessity ctd.

**Lemma**

*The clause set* $\mathsf{opt}_{\Sigma_1}$ *is refutable with SCSCs.*

**Proof Sketch.**

(1)

$$\mathsf{opt}_{\Sigma_1}(\mathsf{n}) \models_{\mathrm{FOL}} \mathsf{opt}_{\Sigma_1}(\mathsf{n})$$

(2)

$$\mathsf{p}(x,y) \rightarrow \mathsf{p}(\mathsf{s}x, y+y) \models_{\mathrm{FOL}} \neg\mathsf{p}(\mathsf{s}x, y+y) \rightarrow \neg\mathsf{p}(x,y)$$
$$\neg\mathsf{p}(\mathsf{s}(\mathsf{n}), y) \models_{\mathrm{FOL}} \neg\mathsf{p}(\mathsf{s}(\mathsf{n}), y+y)$$
$$\mathsf{opt}_{\Sigma_1}(\mathsf{s}(\mathsf{n})) \models_{\mathrm{FOL}} \neg\mathsf{p}(\mathsf{n}, y) \models_{\mathrm{FOL}} \mathsf{opt}_{\Sigma_1}(\mathsf{n})$$

(3)

$$\mathsf{opt}_{\Sigma_1}(0) \models_{\mathrm{FOL}} \neg\mathsf{p}(0,y), \mathsf{p}(0,\overline{1}) \models_{\mathrm{FOL}}$$

□

# $\Sigma_1$-Necessity ctd.

### Lemma

*The sequent* $\Gamma \Rightarrow \forall x \exists y \mathrm{p}(x, y)$ *is not provable in* **LK** $+$ *quantifier-free induction.*

### Proof Sketch.

Proceed indirectly and assume that there exists a proof $\pi$ of $\Gamma \Rightarrow \forall x \exists y \mathrm{p}(x, y)$ in **LK** $+$ quantifier-free induction.
Since the induction in $\pi$ is quantifier-free we can transform $\pi$ into a proof of the form

$$\dfrac{\begin{array}{c}(\pi_1(\alpha)) \\ \Gamma \Rightarrow \exists y \mathrm{p}(\alpha, y)\end{array}}{\Gamma \Rightarrow \forall x \exists y \mathrm{p}(x, y),} \ \forall_r$$

where $\pi_1$ is an **LK** $+$ quantifier-free induction proof and contains no strong quantifier inferences, by eliminating free-cuts and permuting strong quantifier inferences downwards.

# $\Sigma_1$-Necessity ctd.

**Lemma**

*The sequent* $\Gamma \Rightarrow \forall x \exists y \mathsf{p}(x, y)$ *is not provable in* **LK** + *quantifier-free induction.*

**Proof Sketch.**

Since $\pi_1$ contains only quantifier-free induction, no free-cuts and no strong quantifiers we can transform $\pi_1$ into a proof of the form

$$\cfrac{\begin{array}{c}(\pi_2(\alpha)) \\ \Gamma \Rightarrow \mathsf{p}(\alpha, t_1(\alpha)), \ldots, \mathsf{p}(\alpha, t_k(\alpha))\end{array}}{\Gamma \Rightarrow \exists y \mathsf{p}(\alpha, y),} \exists_r$$

by shifting weak quantifier inferences downwards.

# $\Sigma_1$-Necessity ctd.

## Lemma

*The sequent* $\Gamma \Rightarrow \forall x \exists y \mathrm{p}(x, y)$ *is not provable in* **LK** $+$ *quantifier-free induction.*

## Proof Sketch.

Let $m \in \mathbb{N}$. By unfolding induction in $\pi_2(\overline{m})$, we obtain a proof $\rho_m$ of the sequent

$$\Gamma \Rightarrow \mathrm{p}(\overline{m}, t_1(\overline{m})), \ldots, \mathrm{p}(\overline{m}, t_k(\overline{m})).$$

Let $\mathcal{M} = (\mathbb{N}, I)$ where $\mathrm{p}^I = \{(n, n^2) : n \in \mathbb{N}\}$, and $I$ interprets $\mathrm{s}, 0, +$ naturally. Observe that $\mathcal{M} \models \Gamma$ and

$$[\![t_i(\overline{m})]\!]^{\mathcal{M}} = |t_i|_\alpha m + |t_i|_s.$$

There is $j \in \mathbb{N}$ s.t. $j^2 > [\![t_i(j)]\!]^{\mathcal{M}}$, i.e. $\mathcal{M} \not\models \mathrm{p}(\bar{j}, t_i(\bar{j}))$, $1 \leq i \leq k$. $\square$

# Outline

# $\Sigma_1$-Completeness

## Definition

Let the clause set $S(\mathrm{n})$ consist of the clauses

$$x + 0 = x$$
$$x + \mathrm{s}(y) = \mathrm{s}(x + y)$$
$$\mathrm{n} + (\mathrm{n} + \mathrm{n}) \neq (\mathrm{n} + \mathrm{n}) + \mathrm{n}.$$

## Lemma

*The set $S(\mathrm{n})$ is refutable in* **LK** $+$ *quantifier-free induction.*

## Sketch.

Use $x + (x + y) = (x + x) + y$ as induction invariant. □

## Conjecture

*The clause set $S(\mathrm{n})$ is not refutable with SCSCs.*

# Outline

# n-Clause Calculus

Introduced by Kersani and Peltier 2013

Superposition calculus + Cycle detection mechanism

- Sorts $\omega$, $\iota_1$, $\iota_2$, $\ldots \iota_l$
- Signature $\Sigma$ containing 0: $\omega$ and s: $\omega \to \omega$
- Parameter n ($\approx$ Skolem constant)
- Constraint clauses

$$[\underbrace{r_1 \bowtie_1 s_1, \ldots, r_k \bowtie_k s_k}_{\text{Clause part}} \underbrace{\mid}_{\leftarrow} \underbrace{\mathsf{n} \simeq t_1, \ldots, \mathsf{n} \simeq t_m}_{\text{Constraint part}}],$$

for $i = 1, \ldots, n$, $r_i, s_i$ are $\iota_{j_i}$-terms with $j_i \in \{1, \ldots, l\}$, $\bowtie_i \in \{\simeq, \not\simeq\}$, and $t_1, \ldots, t_m$ are $\omega$-terms.

# Cyclic Refutations

### Definition

Let $S$ be a set of clauses. An inductive cycle for $S$ is a 3-tuple $\langle i, j, S_{\text{init}} \rangle$ with $i, j \in \mathbb{N}$, $j > 0$, $S_{\text{init}} \subseteq S$ s.t.

- Base cases:

$$S_{\text{init}} \vdash \mathsf{n} \neq i + k, \text{for } k = 0, \dots, j - 1.$$

- Step case:

$$S_{\text{init}} \vdash S_{\text{init}}[\mathsf{n}/\mathsf{n} - j].$$

### Proposition (Kersani and Peltier 2013)

*If $S$ contains an inductive cycle $\langle i, j, S_{\text{init}} \rangle$, then $S \models_{\mathrm{KP}} \mathsf{n} \prec i$.*

$\rightsquigarrow$ $S$ is refuted if $S \vdash \mathsf{n} \not\approx 0, \dots, \mathsf{n} \not\approx i - 1$ and $S \vdash_{\text{cycle}} \mathsf{n} \prec i$.

# Reduction to Semantic Clause Set Cycles

## Lemma (Reduction)

*If $S(\mathsf{n})$ is refutable in the n-clause calculus, then $S_{\mathsf{inj}} \cup S(\mathsf{n})$ is refutable with SCSCs.*

## Proof Sketch.

- Represent n-clauses as clauses: $[C \mid \mathsf{n} \simeq t] \rightsquigarrow C \vee \mathsf{n} \neq t$
- Normalization requires injectivity
  $S_{\mathsf{inj}} = \{0 \neq \mathsf{s}(x), \mathsf{s}(x) = \mathsf{s}(y) \rightarrow x = y\}$
- Observe that $S_1 \vdash S_2$ implies $S_1 \models^{S_{\mathsf{inj}}}_{\mathrm{FOL}} S_2$. $\square$

## Corollary ($\Sigma_1$-Bound)

*If $S(\mathsf{n})$ is refutable in the n-clause calculus, then $S_{\mathsf{inj}} \cup S(\mathsf{n})$ is refutable in* **LK**$+$ *$\Sigma_1$-induction.*

# $\Sigma_1$-Necessity and $\Sigma_1$-Completeness

## Theorem

*There is a clause set $S(\mathsf{n})$ which is refutable in the n-clause calculus such that $S_{\mathsf{inj}} \cup S(\mathsf{n})$ is not refutable in* **LK** $+$ *quantifier-free induction.*

## Proof Idea.

Let $\mathsf{c} : \iota, \mathsf{f} : \iota \to \iota, \mathsf{t} : o, \mathsf{p} : \omega \to \iota \to o$ be constants. An **LK** $+$ quantifier-free induction proof of the clause set below implies an **LK** $+$ quantifier-free induction proof of $\mathsf{opt}_{\Sigma_1}$.

$$\mathsf{p}(0, \mathsf{c}) \simeq \mathsf{t}, \mathsf{p}(x, y) \not\simeq \mathsf{t} \vee \mathsf{p}(\mathsf{s}(x), \mathsf{f}(y)) \simeq \mathsf{t}, [\mathsf{p}(x, y) \not\simeq \mathsf{t} \mid \mathsf{n} \simeq x] \quad \square$$

## Conjecture

*There is a clause set $S(\mathsf{n})$ which is refutable in the n-clause calculus but is not refutable in* **LK** $+$ *quantifier-free induction.*

# Conclusion

- Two independent views of induction

  Computer science vs. mathematics, little interaction
- Analysis of approaches to autom. ind. theorem proving

  What kind of problems can a method solve?
  Improve general understanding of approaches
- Semantic Clause Set Cycles

  Abstract a family of clausal, refutational calculi
- $\Sigma_1$-bound, $\Sigma_1$-necessity, and $\Sigma_1$-completeness

  Describe the provable sentences
- Case Study: n-clause calculus

  Extend results to concrete calculi