

# Stronger Higher-order Automation

## A Report on the Ongoing Matryoshka Project

Jasmin Blanchette, Pascal Fontaine,  
Stephan Schulz, Sophie Tourret, Uwe Waldmann

VU Amsterdam, Loria, DHBW Stuttgart, MPII Saarbrücken

ARCADE 2019, August 26<sup>th</sup>



# Matryoshka

<http://matryoshka.gforge.inria.fr/>

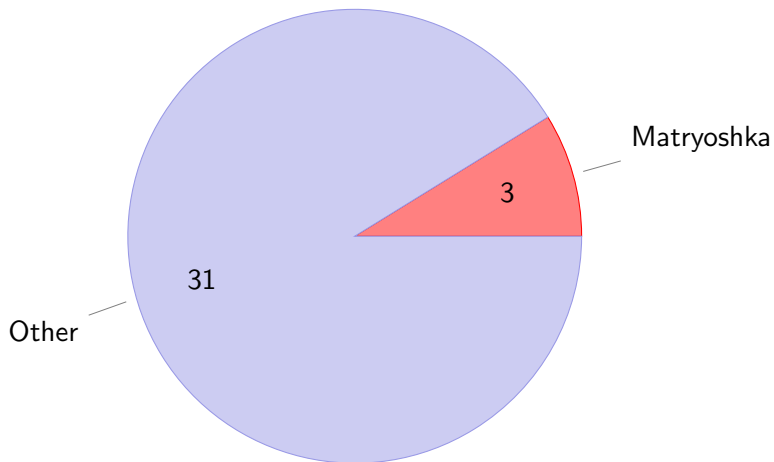


- ▶ Jasmin Blanchette's European Research Council (ERC) Starting Grant 2016  
Grant agreement No. 713999
- ▶ March 2017 – February 2022

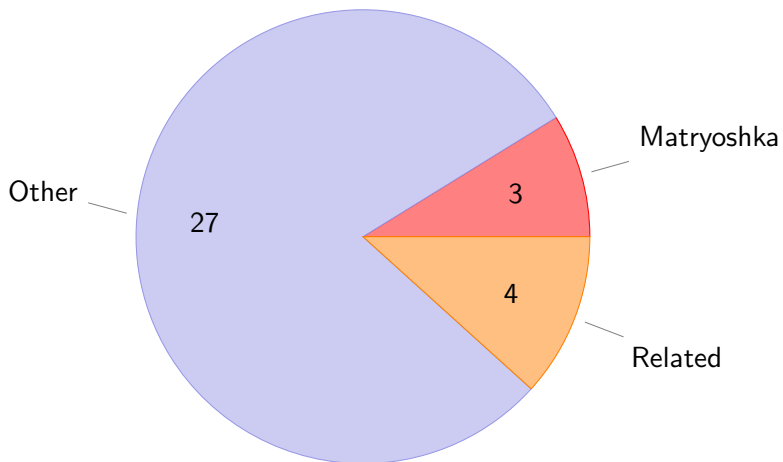
- ▶ 5 PhD students + 2 postdocs
- ▶ strong collaboration with the CVC4 team (Iowa/Stanford)
- ▶ Goal: **Bridge the gap between ATP and ITP**



# Matryoshka in CADE 2019



# Matryoshka in CADE 2019



# CADE-27: 27TH INTERNATIONAL CONFERENCE ON AUTOMATED DEDUCTION

[PROGRAM](#) [AUTHORS](#) [KEYWORDS](#) [SLIDES](#)

## PROGRAM

Days: [Tuesday, August 27th](#) [Wednesday, August 28th](#) [Thursday, August 29th](#) [Friday, August 30th](#)

Tuesday, August 27th

View this program: [with abstracts](#) [session overview](#) [talk overview](#)

08:30-10:00 Session 1

08:30 [Assia Mahboubi](#)

**Invited Talk: Computer Deduction and (Formal) Proofs in Mathematics** ([abstract](#))


09:30 [Andrei Popescu](#) and [Dmitriy Traytel](#)

**A Formally Verified Abstract Account of Gödel's Incompleteness Theorems** ([abstract](#))

10:00-10:30 Coffee break

10:30-12:30 Session 2

10:30 [Giles Regeer](#) and [Andrei Voronkov](#)

**Induction in Saturation-Based Proof Search** ([abstract](#)) 

11:00 [Karel Chvalovský](#), [Jan Jakubuv](#), [Martin Suda](#) and [Josef Urban](#)

**ENIGMA-NG: Efficient Neural and Gradient-Boosted Inference Guidance for E** ([abstract](#))

11:30 [Ulrich Furbach](#), [Teresa Krämer](#) and [Claudia Schön](#)

**Names are not just Sound and Smoke: Word Embeddings for Axiom Selection.** ([abstract](#))

12:00 [Michael Rawson](#) and [Giles Regeer](#)

**Old or Heavy? Decaying Gracefully with Age/Weight Shapes.** ([abstract](#))

12:30-14:00 Lunch

14:00-16:00 Session 3

14:00 [David Plaisted](#)

**The Aspect Calculus** ([abstract](#)) 

14:30 [Valentin Cassano](#), [Raul Fervari](#), [Guillaume Hoffmann](#), [Carlos Areces](#) and [Pablo Castro](#)

**A Tableaux Calculus for Default Intuitionistic Logic** ([abstract](#))

15:00 [Yizheng Zhao](#) and [Renate A. Schmidt](#)

**FAME 2.0 – A Semantic Forgetting Tool for Description Logics with Qualified Number Restrictions (System Description)** ([abstract](#))

15:20 [Tanel Tammet](#)

**GKC: a Reasoning System for Large Knowledge Bases (System Description)** ([abstract](#))

15:40 [Raúl Gutiérrez](#) and [Salvador Lucas](#)

**Automatic Generation of Logical Models with AGES (System Description)** ([abstract](#))

16:00-16:30 Coffee break

16:30-17:10 Session 4

16:30 [Stephan Schulz](#), [Simon Cruanes](#) and [Petar Vukmirović](#)

**Faster, Higher, Stronger: E 2.3 (System Description)** ([abstract](#))

16:50 [Geoff Sutcliffe](#) and [Francis Jeffrey Pelletier](#)

**JGXYZ - An ATP System for Gap and Glut Logics (System Description)** ([abstract](#))

17:10-18:40 Session 5: Awards session

Wednesday, August 28th

View this program: [with abstracts](#) [session overview](#) [talk overview](#)

08:30-09:50 Session 6

08:30 [Cas Cremers](#)

**Invited Talk: Automated Reasoning for Security Protocols** ([abstract](#))

09:30 [Di Long Li](#) and [Alwen Tiu](#)

**Combining ProVerif and Automated Theorem Provers for Security Protocol Verification (System Description)** ([abstract](#))

10:00-10:30 Coffee break

10:30-12:30 Session 7

10:30 [Christian Sternagel](#) and [Sarah Winkler](#)

**Certified Equational Reasoning via Ordered Completion** ([abstract](#))

11:00 [Siva Anantharaman](#), [Peter Hibbs](#), [Paliath Narendran](#) and [Michael Rusinowitch](#)

**Unification modulo Lists with Reverse - Relation with Certain Word Equations** ([abstract](#))

11:30 [Nao Hirokawa](#), [Julian Nagele](#), [Vincent Van Oostrom](#) and [Michio Oyamauchi](#)

**Confluence by Critical Pair Analysis Revisited** ([abstract](#))

12:00 [Christina Kohl](#) and [Aart Middeldorp](#)

**Composing Proof Terms** ([abstract](#))

12:30-13:30 Session 8: Lunch

13:30-18:00 Excursion

View this program: [with abstracts](#) [session overview](#) [talk overview](#)

09:30-10:00 Session 14

09:30 [Geoff Sutcliffe](#)

**The CADE-27 ATP System Competition - CASC-27** ([abstract](#))

10:00-10:30 Coffee break

10:30-12:30 Session 15

10:30 [Vojtěch Havlena](#), [Lukas Holik](#), [Ondrej Lengal](#) and [Tomas Vojnar](#)

**Automata Terms in a Lazy WSks Decision Procedure** ([abstract](#))

11:00 [Mateus De Oliveira Oliveira](#) and [Alexsander Andrade de Melo](#)

**On the Width of Regular Classes of Finite Structures** ([abstract](#))

11:30 [Alberto Fiori](#) and [Christoph Weidenbach](#)

**SCL – Clause Learning from Simple Models** ([abstract](#))

12:30-14:00 Lunch

14:00-16:00 Session 16

14:00 [Chad Brown](#), [Thibault Gauthier](#), [Cezary Kaliszyk](#), [Geoff Sutcliffe](#) and [Josef Urban](#)

**GRUNGE: The Grand Unified ATP Challenge** ([abstract](#))

14:30 [Alexander Bentkamp](#), [Jasmin Christian Blanchette](#), [Sophie Tournet](#), [Petar Vukmirović](#) and [Uwe Waldmann](#)

**Superposition with Lambdas** ([abstract](#))

15:00 [Ahmed Bhayat](#) and [Giles Reger](#)

**Restricted Combinatory Unification** ([abstract](#))

15:30 [Haniel Barbosa](#), [Andrew Reynolds](#), [Daniel El Ouaoui](#), [Cesare Tinelli](#) and [Clark Barrett](#)

**Extending SMT solvers to Higher-Order Logic** ([abstract](#))

# Structure of the Matryoshka Contributions

AR for Proof Assistants

Proof Assistants for AR



# Structure of the Matryoshka Contributions

## AR for Proof Assistants

- ▶ Superposition for HOL  
[IJCAR18], [TACAS19], [CADE19]
- ▶ SMT for HOL  
[SMT18], [CADE19]
- ▶ SMT proof reconstruction  
[CADE17], [AITP19], [PxTP19]

## Proof Assistants for AR

# Structure of the Matryoshka Contributions

## AR for Proof Assistants

- ▶ Superposition for HOL  
[IJCAR18], [TACAS19], [CADE19]
- ▶ SMT for HOL  
[SMT18], [CADE19]
- ▶ SMT proof reconstruction  
[CADE17], [AITP19], [PxTP19]

## Proof Assistants for AR

- ▶ SAT solver  
[CPP18], [NFM19]
- ▶ Ordered resolution prover  
[IJCAR18], [CPP19]
- ▶ Framework for saturation  
provers [work in progress]

# Structure of the Matryoshka Contributions

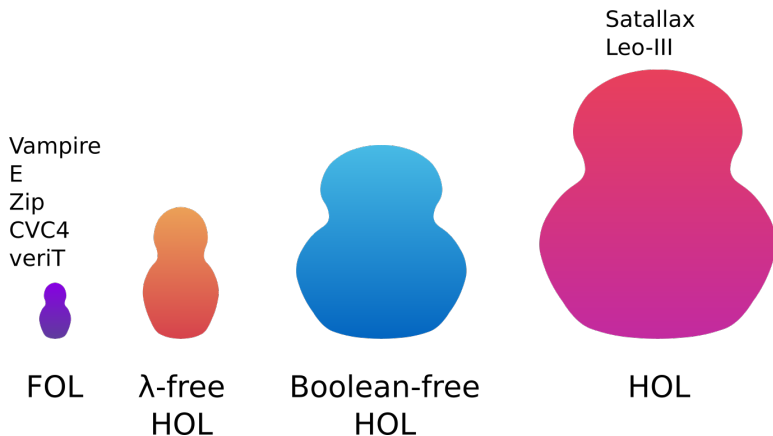
## AR for Proof Assistants

- ▶ **Superposition for HOL**  
[IJCAR18], [TACAS19], [CADE19]
- ▶ **SMT for HOL**  
[SMT18], [CADE19]
- ▶ **SMT proof reconstruction**  
[CADE17], [AITP19], [PxTP19]

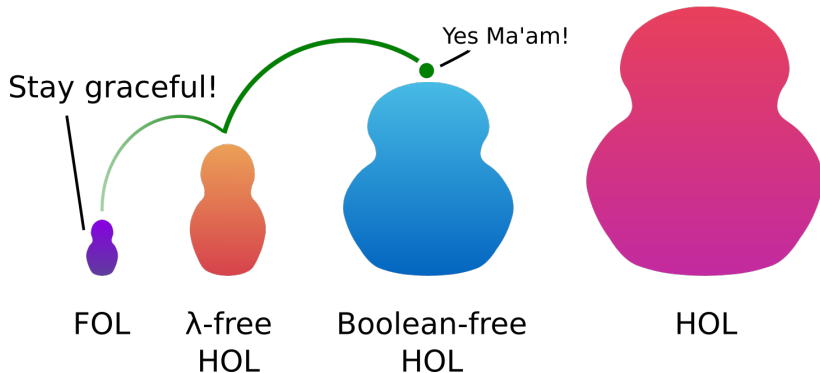
## Proof Assistants for AR

- ▶ **SAT solver**  
[CPP18], [NFM19]
- ▶ **Ordered resolution prover**  
[IJCAR18], [CPP19]
- ▶ **Framework for saturation provers** [work in progress]

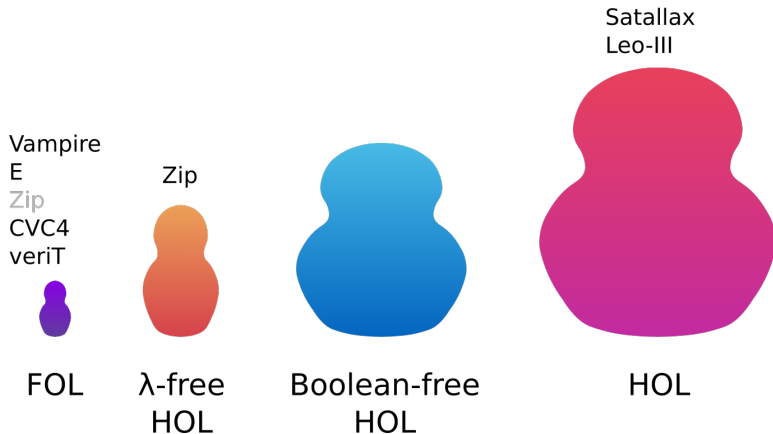
# Theorem Provers from First- to Higher-order Logic



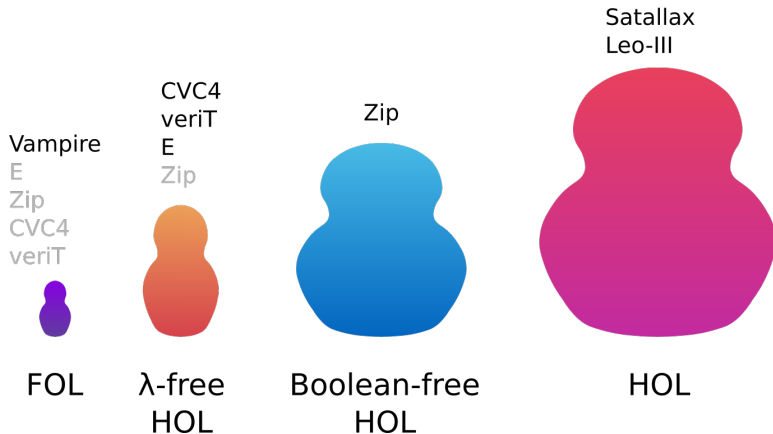
# Theorem Provers from First- to Higher-order Logic



# Theorem Provers from First- to Higher-order Logic



# Theorem Provers from First- to Higher-order Logic



# A Framework for Saturation-based Theorem Provers

static completeness of calculus



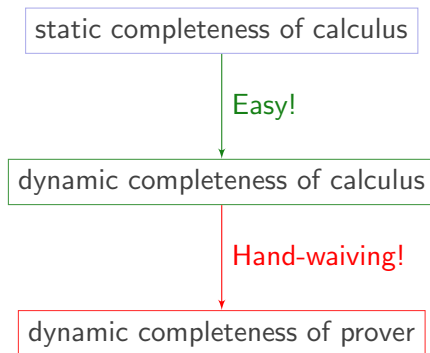
# A Framework for Saturation-based Theorem Provers

static completeness of calculus

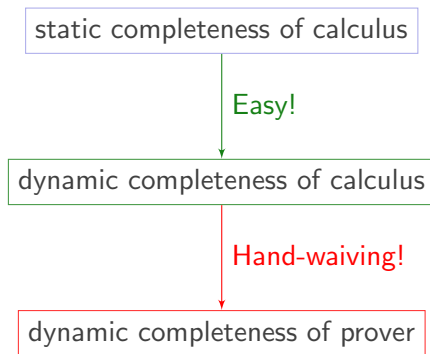
Easy!

dynamic completeness of calculus

# A Framework for Saturation-based Theorem Provers



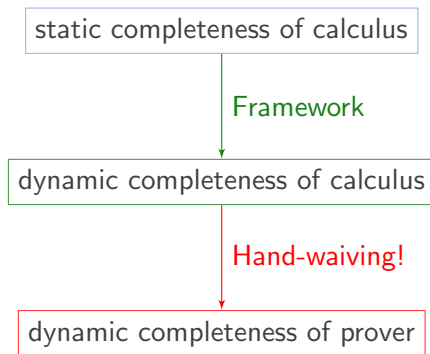
# A Framework for Saturation-based Theorem Provers



## Reason

- ▶ In **theory** only **redundant** clauses are deleted.
- ▶ In **practice** **subsumed** clauses are deleted.
- ▶ Ex:  $p(a, a) \quad ? \quad p(x, x)$

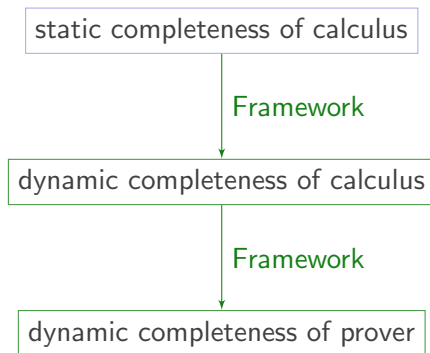
# A Framework for Saturation-based Theorem Provers



## Reason

- ▶ In **theory** only **redundant** clauses are deleted.
- ▶ In **practice** **subsumed** clauses are deleted.
- ▶ Ex:  $p(a, a) \quad ? \quad p(x, x)$

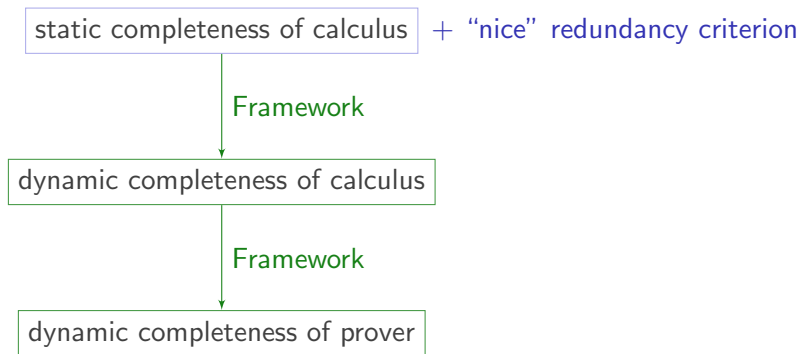
# A Framework for Saturation-based Theorem Provers



## Reason

- ▶ In **theory** only **redundant** clauses are deleted.
- ▶ In **practice** **subsumed** clauses are deleted.
- ▶ Ex:  $p(a, a) \quad ? \quad p(x, x)$

# A Framework for Saturation-based Theorem Provers



## Reason

- ▶ In **theory** only **redundant** clauses are deleted.
- ▶ In **practice** **subsumed** clauses are deleted.
- ▶ Ex:  $p(a, a) \quad ? \quad p(x, x)$

# A Framework for Saturation-based Theorem Provers

Applications:

- ▶ ordered resolution
- ▶ standard superposition
- ▶ constraint superposition

# A Framework for Saturation-based Theorem Provers

Applications:

- ▶ ordered resolution
- ▶ standard superposition
- ▶ constraint superposition
- ▶ theory superposition
- ▶ hierarchic superposition
- ▶ higher-order superposition
- ▶ Knuth-Bendix completion
- ▶ ...?



# A Framework for Saturation-based Theorem Provers

Applications:

- ▶ ordered resolution
- ▶ standard superposition
- ▶ constraint superposition
- ▶ theory superposition
- ▶ hierarchic superposition
- ▶ higher-order superposition
- ▶ Knuth-Bendix completion
- ▶ ...?

Formalization in Isabelle/HOL

# Possible Discussion Topics

How much higher-orderness do we really need in practice?

Which parts of AR would you like to see being formally verified?  
Do you know of such ongoing efforts?