

Abstract Saturation-Based Inference

Nachum Dershowitz*
School of Computer Science
Tel-Aviv University
P.O. Box 39040
Ramat Aviv, Tel-Aviv 69978
Israel
Email: Nachumd@tau.ac.il

Claude Kirchner
LORIA & INRIA
615, rue du Jardin Botanique
B.P. 101
54602 Villers-lès-Nancy Cedex
France
Claude.Kirchner@loria.fr

Abstract

Solving goals—like deciding word problems or resolving constraints—is much easier in some theory presentations than in others. What have been called “completion processes”, in particular in the study of equational logic, involve finding appropriate presentations of a given theory to solve easily a given class of problems.

We provide a general proof-theoretic setting within which completion-like processes can be modelled and studied. This framework centers around well-founded orderings of proofs. It allows for abstract definitions and very general characterizations of saturation processes and redundancy criteria.

1. Introduction

Good axiomatizations

It is common when defining a theory axiomatically to ask whether the chosen axioms (like Euclid’s axiom of parallels) are independent. Dependent axioms are superfluous from the point of view of the theory (set of theorems); such redundancies can be removed. Similarly, one speaks of independent sets of equations, or of alternative presentations of algebras. In these cases, one is comparing sets of formulae based simply on number or total size.

One also speaks of solving equations, or, more generally, sets of constraints. In such a context, one cares about the form of formulae. The process of solving transforms a defining set for the problem into formulae in *solved form*; see [11]. In Gaussian elimination, for example, one begins with a set of linear equalities involving unknowns, and infers solved forms assigning numerical values to each un-

known, or most general relations between variables. This corresponds to the point of view that arithmetic is a cheap form of inference, while equation solving is relatively hard. Thus, once one has derived a solved form, it is easy to check whether other linear equalities follow.

In these examples, as in many others, one is given an axiomatic presentation, and sets up a goal of deriving certain formulae: theorems in Euclidian geometry, in one case; solutions of equations, in the other. In both cases, some presentations of the underlying theory are better suited for solving the given problem than others. The goal of the work described here is to *define* the “best” axiomatic presentation to solve a given problem.

We compare presentations in terms of the quality of proofs they allow. Our theory is, therefore, based on a concept of “good” proofs. The archetypical instance of this paradigm is finding decision procedures for the uniform word problem in an equational theory: the best proofs are rewrite (“valley”) proofs; the best presentation is a terminating Church-Rosser rewrite system. For more on rewriting, see [2, 17, 36].

Good proofs

What makes for a good proof? Consider a naïve example: Suppose we have an equational theory defined by the axioms $a = b$ and $b = c$. Then $a = b = c = b = c$ and $a = b = c$ are two proofs of $a = c$, but, clearly, the second is better than the first, as it is shorter. More generally, in proof theory, one assigns ordinals to proofs and shows that under certain circumstances there exists a formula in a proof that can be replaced in a way that reduces the ordinal of the proof. These proof-theoretical concepts can be extended to dynamically changing proof systems (see [16]).

In this paper, proof quality is measured via a well-founded proof ordering on the set of proofs: the smaller in the ordering the better. In this way, we generalize the

*Supported in part by the Israel Science Foundation.

proof-ordering method used in term-rewriting for establishing properties of rewrite-system completion procedures [5] to the abstract setting of arbitrary proof systems, supplied with an ordering of proofs.

The best presentation

Completion processes have been devised in various different contexts, but in a rather similar fashion. These include: standard Knuth-Bendix completion [26], equational completion [30], completion in specific algebras (like order-sorted ones [19]), inductionless induction (initiated by Musser; see [25]), ordered completion [7, 20, 27], completion for semantic unification [13, 18], to mention a few. The formalization of the completion mechanism, as well as its correctness and completeness, has been intensively studied, beginning with the seminal work of Gérard Huet [21] and especially since the introduction of proof orderings in [6]. The universality of “completion” in automated deduction is further evident in syntheses of the completion procedure and Gröbner basis generation [9], as in, for example, [24].

An interesting feature of the complete set of reductions produced by completion and the Gröbner bases produced by Buchberger’s algorithm is that they are unique, regardless of nondeterministic choices made along the way [15, 28]. In other words, “best presentations” are unique for a given an ordering of proofs.

Starting from a simple, abstract and universal setting, we define a number of abstract properties of *presentations*, that is, of arbitrary sets of formulæ. Fixing inference and the ordering, we characterize the unique *canonical presentation* in several ways, which are boxed like this in the following sections:

1. Lemmata that can appear as assumptions in minimal proofs (Definition 9)
2. Smallest saturated set (Theorem 3)
3. Simplest presentation (Theorem 4)
4. Non-redundant formulæ (Corollary 1)
5. Conclusions of trivial proofs (Corollary 2)
6. Limit of a fair completion process (Theorem 9)

Abstract formal definitions of redundancy, saturation, canonicity, completeness, simplicity, triviality, and fairness will be provided.

One of the main goals in our work is to demonstrate that an abstract framework can be designed, with non-trivial results, without assuming anything about the context other than the existence of a well-founded ordering of entities called proofs. We believe we have attained a high level

of abstraction together with meaningful results because of a fruitful definition of “good proofs”. Somewhat similarly motivated research, includes [1, 35]; an attempt to get an abstraction of critical pairs in category theory was proposed in [31].

2. Proof Systems

We begin with the following structure, which we will call an *ordered proof system*:

- *Proofs* \mathbb{P} ;
- *Formulæ* \mathbb{A} ;
- *Assumptions* $\Gamma : \mathbb{P} \rightarrow 2^{\mathbb{A}}$;
- *Conclusion* $\Delta : \mathbb{P} \rightarrow \mathbb{A}$;
- *Proof ordering* $\geq : \mathbb{P} \times \mathbb{P} \rightarrow 2$.

The important point is the proof ordering, which may be partial. As usual, we use $>$ for $\geq \cap \neq$. For conciseness later, we will presume that proofs of different theorems are *incomparable*: $p > q$ implies $\Delta p = \Delta q$.

We will use the term *presentation* to mean a set of formulæ, and *justification* to mean a set of proofs. We reserve the term *theory* for deductively closed presentations.

For example, consider a propositional (ordered) resolution calculus: Formulæ are finite sets of literals; proofs are finite ordered unary-binary trees, with formulæ for leaves and literals labelling internal nodes. Propositional constants are linearly ordered and proofs are compared using the corresponding recursive path ordering [12], but with all unary nodes cheaper than all binary nodes. On the concrete level, a binary node ℓ corresponds to an inference

$$\frac{\ell \vee f \quad \bar{\ell} \vee f'}{f \vee f'}$$

and a unary node ℓ corresponds to $\frac{f}{\ell \vee f}$, where ℓ is a literal, $\bar{\ell}$ is its negation, and f, f' are sets of literals. There are also maximally expensive binary nodes for projections:

$$\frac{f \quad f'}{f} \quad \frac{f' \quad f}{f}$$

Let $B = \{a \vee \bar{b}, b \vee c, \bar{a} \vee \bar{c}, a\}$; the smallest proof of b (elaborated with formulæ at internal nodes), assuming precedence $a < b < c$, is

$$\frac{\frac{b \vee c \quad \bar{a} \vee \bar{c}}{\bar{a} \vee b} \quad a}{b}$$

Extend Γ and Δ to sets (of proofs) in the standard fashion: $\Gamma P \stackrel{\text{def}}{=} \cup_{p \in P} \Gamma p$; $\Delta P \stackrel{\text{def}}{=} \{\Delta p : p \in P\}$. Trivially: $\Gamma \emptyset = \emptyset$ and $\Delta \emptyset = \emptyset$. It follows immediately from the definitions that Γ and Δ are monotonic: For all justifications P and Q , $P \subseteq Q \Rightarrow \Gamma P \subseteq \Gamma Q$ and $P \subseteq Q \Rightarrow \Delta P \subseteq \Delta Q$.

Definition 1 (Proofs) *The set of all proofs using some of the assumptions:*

$$\overline{\Pi} A \stackrel{!}{=} \{p \in \mathbb{P} : \Gamma p \subseteq A\}$$

For a specific conclusion $c \in \mathbb{A}$, we sometimes write:

$$\overline{\Pi}_c A \stackrel{!}{=} \{p \in \overline{\Pi} A : \Delta p = c\}$$

It follows from these definitions that for all presentations A, B :

$$\Gamma \overline{\Pi} A \subseteq A \quad (1)$$

$$A \subseteq B \Rightarrow \overline{\Pi} A \subseteq \overline{\Pi} B \quad (2)$$

and for all justifications P :

$$P \subseteq \overline{\Pi} \Gamma P \quad (3)$$

Note. Presentations and justifications are related by the *Galois connection* formed by $\overline{\Pi}$ and Γ with respect to \subseteq .

From the previous definitions, it is easy to see that proofs need only what they use, that is, $\overline{\Pi} \Gamma \overline{\Pi} A = \overline{\Pi} A$.

The *pre-image* Γ^{-1} of A are those proofs with exactly A as assumptions: $\Gamma^{-1} A = \{p : \Gamma p = A\}$.

Definition 2 (Theories)

1. *The theory (or deductive closure) of a presentation A :*

$$\Theta A \stackrel{!}{=} \Delta \Gamma^{-1} A$$

2. *A presentation A is a basis for a theory θ if*

$$\Theta A = \theta$$

3. *A presentation A is deductively closed if*

$$\Theta A = A$$

4. *Presentations A and B are equivalent if they allow exactly the same theorems:*

$$A \equiv B \Leftrightarrow \Theta A = \Theta B$$

Most of our results depend on the following three standard properties of Tarskian consequence relations:

Monotonicity:

$$A \subseteq B \Rightarrow \Theta A \subseteq \Theta B$$

Reflexivity:

$$A \subseteq \Theta A$$

Closure:

$$\Theta \Theta A \subseteq \Theta A$$

We will assume that these postulates hold for all proof systems considered in this paper.

Note. The Monotonicity Postulate requires inference systems to include structural weakening rules like $\frac{a}{a}$.

We begin with some basic properties of proof systems satisfying these postulates. First, transitivity of consequences follows from Monotonicity and Closure: For all presentations A, B and C , $\Theta A \supseteq B \wedge \Theta B \supseteq C \Rightarrow \Theta A \supseteq C$. For all presentations A , by reflexivity, $A \subseteq \Theta A$, and by monotonicity, $\Theta A \subseteq \Theta \Theta A$. Finally, Closure allows one to conclude that the theories are equal, that is, a presentation A and its theory ΘA support exactly the same theorems: $\Theta \Theta A = \Theta A$, or:

$$\Theta A \equiv A$$

Thanks to Monotonicity,

$$\Theta A = \Delta \Gamma^{-1} A = \Delta \overline{\Pi} A$$

Lemma 1

$$\overline{\Pi} A \subseteq \overline{\Pi} B \Leftrightarrow A \subseteq B$$

3. Reduced Systems

Proof orderings allow for the following notion, central to the development of the theory of canonical inference:

Definition 3 (Minimal Proofs)

$$\mu P \stackrel{!}{=} \{p \in P : \neg \exists q \in P. q < p\}$$

Recall that $q < p$ would only hold for proofs p, q with the same conclusion.

Well-foundedness of the proof ordering means that minimal proofs suffice: $\Delta \overline{\Pi} A = \Delta \mu \overline{\Pi} A$.

Proof orderings can be lifted to sets of proofs (using a Smyth [33] powerdomain construction), as follows:

Definition 4 (Better Proofs) *Justification Q is better than justification P if:*

$$P \sqsupseteq Q \stackrel{!}{\Leftrightarrow} \forall p \in P. \exists q \in Q. p \geq q$$

It is much better if:

$$P \sqsubset Q \stackrel{!}{\Leftrightarrow} \forall p \in P. \exists q \in Q. p > q$$

Better is clearly a quasi-order. Also, it is clear from the definitions that $P \sqsubset Q \sqsupseteq R \Rightarrow P \sqsubset R$ and $P \sqsupseteq Q \sqsupseteq R \Rightarrow P \sqsupseteq R$.

On account of well-foundedness, minimal proofs always exist. Hence the following statements hold:

Proposition 1 For all justifications P, Q :

1. $P \sqsupseteq \mu P$
2. $P \subseteq Q \Rightarrow P \sqsupseteq Q$
3. $P \sqsupseteq Q \Rightarrow \Delta P \subseteq \Delta Q$
4. $P \sqsupseteq Q \Leftrightarrow \mu P \sqsupseteq \mu Q$
5. $P \sqsupset Q \Leftrightarrow \mu P \sqsupset \mu Q$

and for all presentations A, B :

1. $\overline{\Pi} A \sqsupseteq \overline{\Pi} B \Rightarrow \Theta A \subseteq \Theta B$
2. $A \subseteq B \Rightarrow \overline{\Pi} A \sqsupseteq \overline{\Pi} B$
3. $B \subseteq A \wedge \overline{\Pi} A \sqsupseteq \overline{\Pi} B \Rightarrow A \equiv B$

As a consequence: \sqsupseteq is a partial ordering on *minimal* proofs.

Definition 5 (Flattening) Those assumptions employed in minimal proofs are denoted

$$A^b \stackrel{!}{=} \Gamma \mu \overline{\Pi} A$$

A presentation A is reduced (or flat) if

$$A = A^b$$

It is always the case that $A^b \subseteq A$: By definition, $\mu \overline{\Pi} A \subseteq \overline{\Pi} A$. By monotonicity of Γ and the fact that $\Gamma \overline{\Pi} A \subseteq A$ (Eq. 1), we get

$$A^b = \Gamma \mu \overline{\Pi} A \subseteq \Gamma \overline{\Pi} A \subseteq A \quad (4)$$

Lemma 2 Minimal proofs use the assumptions of minimal proofs:

$$\mu \overline{\Pi} A^b = \mu \overline{\Pi} A$$

Proof: Suppose $p \in \mu \overline{\Pi}_c A$ for some c . Then $\Gamma p \subseteq A^b$ and $p \in \overline{\Pi}_c A^b$. Were there a $q \in \overline{\Pi}_c A^b \subseteq \overline{\Pi} A$ such that $q < p$, p would not be minimal in $\overline{\Pi} A$. For the other direction, suppose $p \in \mu \overline{\Pi}_c A^b \subseteq \overline{\Pi} A$, but p is not minimal in $\overline{\Pi} A$. In other words, there is some $q \in \overline{\Pi}_c A$ such that $p > q$. There must be some $r \in \mu \overline{\Pi}_c A \subseteq \mu \overline{\Pi} A^b$ such that $q \geq r$. This contradicts the minimality of p in $\overline{\Pi} A^b$. \square

Consequently, what is reduced cannot be further reduced, that is, $A^{b^b} = A^b$. Applying the previous statements allows us to derive $\Theta A^b = \Delta \overline{\Pi} A^b = \Delta \mu \overline{\Pi} A^b = \Delta \mu \overline{\Pi} A = \Delta \overline{\Pi} A = \Theta A$. In other words, a reduced system can prove as much as the initial one:

$$A^b \equiv A \quad (5)$$

4. Saturated Systems

Now that we know how to define good proofs, we can understand how much we can restrict a presentation and still be able to prove everything.

Definition 6 (Normal Form Proof) A proof is in normal form if it belongs to the set of minimal proofs that allow the use of all theorems as lemmata. Normal form proofs are denoted as follows:

$$\underline{\Pi} A \stackrel{!}{\Leftrightarrow} \mu \overline{\Pi} \Theta A$$

We will also have occasion to use the notation

$$\underline{\Pi}_c A \stackrel{!}{\Leftrightarrow} \mu \overline{\Pi}_c \Theta A$$

for normal-form proofs of formula c .

Considering only normal-form proofs does not restrict the theory, as we have

$$\Theta A = \Delta \underline{\Pi} A \quad (6)$$

There are two manners in which a presentation can suffice for normal form proofs:

Definition 7 (Completeness and Saturation) A presentation A is complete — denoted $\text{Compl } A$ — if every theorem has a normal form proof:

$$\text{Compl } A \stackrel{!}{\Leftrightarrow} \Theta A = \Delta (\overline{\Pi} A \cap \underline{\Pi} A)$$

and is saturated — denoted $\text{Satur } A$ — if it supports all possible normal form proofs:

$$\text{Satur } A \stackrel{!}{\Leftrightarrow} \mu \overline{\Pi} A = \underline{\Pi} A$$

A presentation is complete if it is saturated, but to prove the converse, we will need (for Proposition 2 below) an additional hypothesis:

Definition 8 Minimal proofs are unique if for all $A \subseteq \mathbb{A}$ and $c \in \mathbb{A}$ it is the case that

$$|\mu \overline{\Pi}_c A| \leq 1$$

5. Canonical Systems

Our main definition (of the *sharpening* operation) is:

Definition 9 (Canonical Presentation) The canonical presentation contains those formulæ that appear as assumptions of minimal proofs:

$$A^\# \stackrel{\dagger}{=} [\Theta A]^\flat$$

It follows from this definition that $A^\# = \Gamma \underline{\underline{A}}$.

Theorem 1 *The function $_^\#$ is canonical with respect to the equivalence of presentations. That is:*

1. $A^\# \equiv A$
2. $A \equiv B \Leftrightarrow A^\# = B^\#$
3. $A^{\#\#} = A^\#$

The canonical presentation cannot be further reduced as we can show simply that $(A^\#)^\flat = A^\#$. Furthermore, we get the usual intuition about saturated sets:

Theorem 2 *A presentation A is saturated iff it contains its own canonical presentation:*

$$\text{Satur } A \Leftrightarrow A \supseteq A^\#$$

Proof: As $A^\# \subseteq B \Leftrightarrow \underline{\underline{A}} \subseteq \overline{\overline{B}}$, and by the definition of saturated, we need to show

$$\mu \overline{\overline{A}} = \underline{\underline{A}} \Leftrightarrow \underline{\underline{A}} \subseteq \overline{\overline{A}}$$

By Reflexivity and monotonicity of $\overline{\overline{\cdot}}$: $\overline{\overline{A}} \subseteq \overline{\overline{\Theta A}}$. So, for any minimal proof $p \in \mu \overline{\overline{A}}$, $c \in \overline{\overline{\Theta A}}$ there must be a $q \in \mu \overline{\overline{\Theta A}} = \underline{\underline{A}} \subseteq \overline{\overline{A}}$ such that $p \geq q$. By minimality, $p = q \in \underline{\underline{A}}$. In other words, $\mu \overline{\overline{A}} \subseteq \underline{\underline{A}}$. So if $\mu \overline{\overline{A}} = \underline{\underline{A}}$, then, $\underline{\underline{A}} = \mu \overline{\overline{A}} \subseteq \overline{\overline{A}}$.

Suppose now that $\underline{\underline{A}} \subseteq \overline{\overline{A}}$. In general, we have

$$\mu \overline{\overline{A}} \cap \underline{\underline{A}} = \overline{\overline{A}} \cap \underline{\underline{A}} \quad (7)$$

Since $\mu \overline{\overline{A}} \subseteq \overline{\overline{A}}$, we need only show that $\overline{\overline{A}} \cap \underline{\underline{A}} \subseteq \mu \overline{\overline{A}}$. Suppose $p \in \overline{\overline{A}} \setminus \mu \overline{\overline{A}}$. Then there is a $q \in \mu \overline{\overline{A}} \subseteq \overline{\overline{A}} \subseteq \overline{\overline{\Theta A}}$ (by Reflexivity) such that $p > q$. But then $p \notin \mu \overline{\overline{\Theta A}} = \underline{\underline{A}}$. By (7)

$$\begin{aligned} \underline{\underline{A}} \subseteq \overline{\overline{A}} &\Leftrightarrow \underline{\underline{A}} = \underline{\underline{A}} \cap \overline{\overline{A}} = \underline{\underline{A}} \cap \mu \overline{\overline{A}} \\ &\Leftrightarrow \underline{\underline{A}} \subseteq \mu \overline{\overline{A}} \end{aligned}$$

□

As a corollary, we get a new property of canonical presentations:

$$\text{Satur } A^\#$$

When we enforce equality instead of inclusion, that is, when we consider presentations that are their own canonical presentation, we arrive at presentations that are canonical: A presentation A is *canonical* if

$$A = A^\#$$

As a consequence, a presentation is canonical iff it is saturated and reduced.

We can now state the second characterization of canonical presentations, namely as the smallest saturated set:

Theorem 3

$$A \equiv B \Rightarrow [\text{Satur } B \Leftrightarrow B \supseteq A^\#]$$

Thus, the canonical presentation is minimal in the sense that no equivalent proper subset of $A^\#$ is saturated.

Finally, if A is saturated, then every equivalent superset also is:

$$\text{Satur } A \wedge A \equiv B \wedge A \subseteq B \Rightarrow \text{Satur } B \quad (8)$$

Consider again the resolution calculus: The canonical presentation for $A = \{a \vee \bar{b}, b \vee c, \bar{a} \vee \bar{c}\}$ includes, in addition, $\{b \vee \bar{a}, a \vee c, \bar{b} \vee \bar{c}\}$. The canonical basis of $B = A \cup \{a\}$ is just $\{a, b, \bar{c}\}$. The canonical basis of $B \cup \{c\}$ is the empty clause.

Proposition 2 *A presentation is complete if it is saturated. If minimal proofs are unique, then a presentation is saturated iff it is complete.*

Proof: If $c \in \Theta A$, then (by 6) there is a proof $q \in \underline{\underline{A}}$ of c . If $\text{Satur } A$, then (by Theorem 2)

$$\Gamma q \subseteq \Gamma \underline{\underline{A}} = A^\# \subseteq A$$

and $q \in (\overline{\overline{A}} \cap \underline{\underline{A}})$, as required for completeness.

For the other direction, by completeness and (7), for all $c \in \Theta A$, $\mu \overline{\overline{A}} \cap \underline{\underline{A}} \neq \emptyset$. By uniqueness of minimal proofs, $|\mu \overline{\overline{A}} \cap \underline{\underline{A}}| \leq 1$. Hence, $\text{Satur } A$, with $\mu \overline{\overline{A}} \cap \underline{\underline{A}} = \underline{\underline{A}}$ for all c . □

6. Redundancy

Formulæ that when removed from a presentation do not hurt proof quality will be called redundant. The notion of redundancy lies at the heart of efficient theorem proving: one seeks to perform inferences on non-redundant formulæ so as to avoid redundancy propagation, whose cost could be prohibitive.

Definition 10 (Simpler Presentation) *Presentation B is said to be simpler than presentation A when B provides better proofs than does A :*

$$A \succsim B \stackrel{\dagger}{\Leftrightarrow} A \equiv B \wedge \overline{\overline{A}} \supseteq \overline{\overline{B}}$$

Reflexivity and transitivity are inherited from \equiv and \supseteq , therefore \succsim is a quasi-ordering. Denoting by \approx the associated equivalence, we get easily that $A \approx B \Leftrightarrow \mu \overline{\overline{A}} = \mu \overline{\overline{B}}$.

The following sequence of facts of increasing interest leads to the assertion that canonical presentations are indeed simpler:

Proposition 3

1. $A \subseteq B \wedge A \equiv B \Rightarrow A \lesssim B$
2. $B \subseteq A \wedge \overline{\Pi} A \supseteq \overline{\Pi} B \Rightarrow A \approx B$
3. $C \subseteq B \wedge A \lesssim A \setminus B \Rightarrow A \approx A \setminus C$
4. $A \approx A^b$
5. $A \lesssim A^\#$

Proof: We show only $A \approx A^b$. By (5), the two theories are equal. Thus, by (4) and the first fact, $A^b \lesssim A$. Applying Lemma 2 and Proposition 1, we get that $A \lesssim A^b$:

$$\overline{\Pi} A \supseteq \mu \overline{\Pi} A = \mu \overline{\Pi} A^b \supseteq \overline{\Pi} A^b$$

□

Theorem 4 *A canonical presentation is the simplest:*

$$A \equiv B \Rightarrow B \lesssim A^\#$$

Definition 11 (Redundancy) *A set R of formulae is redundant with respect to a presentation A when:*

$$A \cup R \approx A \setminus R$$

The set of all redundant formulae of a given presentation A is denoted ρA :

$$\rho A \stackrel{\dagger}{=} \{r \in A : A \approx A \setminus \{r\}\}$$

The next result seems remarkable in its reliance on the power of proof orderings: The set of all *individually* redundant formulae is *globally* redundant.

Lemma 3 *The set of redundant formulae is redundant:*

$$A \approx A \setminus \rho A$$

Proof: Let $A' = A \setminus \rho A$. We show that $\overline{\Pi} A \supseteq \overline{\Pi} A'$ and conclude using Proposition 3.1. Consider some proof $p_1 \in \overline{\Pi}_c A \setminus \overline{\Pi} A'$. Since there is a redundant $r \in \Gamma p_1 \cap \rho A$, there must be a proof $p_2 \in \overline{\Pi}_c (A \setminus \{r\}) \subseteq \overline{\Pi} A$ such that $p_1 \geq p_2$. But $\Gamma p_2 \neq \Gamma p_1$, so $p_1 > p_2$. If $p_2 \notin \overline{\Pi} A'$, then there would also be a $p_3 \in \overline{\Pi} A$, such that $p_2 > p_3$. Since the proof ordering is well-founded, this cannot go on forever, so there is, in fact, a proof $p_n \in \overline{\Pi}_c A'$ such that $p_1 \geq p_n$. □

Theorem 5 *Redundant formulae are not needed:*

$$A^b = A \setminus \rho A$$

Proof: If $a \notin A^b = \Gamma \mu \overline{\Pi} A$, then $\overline{\Pi} A \supseteq \overline{\Pi} A \setminus \{a\}$. Thus, $A \lesssim A \setminus \{a\}$ and $a \in \rho A$.

On the other hand, let $a \in A^b \subseteq A$, that is, $a \in \Gamma p$ for some $p \in \mu \overline{\Pi} A$. Suppose $a \in \rho A$, in other words, $A \lesssim A' = A \setminus \{a\}$. So, there must be a proof $q \leq p$ such that $\Gamma q \subseteq A'$. Since, then, $q \neq p$, we have $q < p$. Hence, $p \notin \mu \overline{\Pi} A$, a contradiction. Thus, $a \notin \rho A$. □

Corollary 1

$$\rho A^\# = \emptyset$$

Lemma 3 is another corollary.

7. Subproofs

In the operational quest for the best proofs, a fundamental step is to perform localized searches for bad subproofs, which could stand improvement. To that end, we now impose additional structure on proofs: a well-founded *subproof* (partial) order \triangleright . We extend this notation to sets:

$$P \triangleright Q \stackrel{\dagger}{\Leftrightarrow} \forall q \in Q. \exists p \in P. p \triangleright q$$

and use \triangleright for its reflexive closure.

A proof is deemed *trivial* when its conclusion depends only on itself, that is, if $\Gamma p = \{\Delta p\}$, and it is its own only subproof. Every formula admits a trivial proof, by Reflexivity. We denote by \hat{a} such a trivial proof of $a \in \mathbb{A}$ and by \hat{A} , the set of trivial proofs of each $a \in A$.

We will hereinafter assume three things about subproofs:

Trivialia. Assumptions are subproofs:

$$P \triangleright \widehat{\Gamma P}$$

Subproof. Subproofs use a subset of the assumptions:

$$P \triangleright Q \Rightarrow \Gamma P \supseteq \Gamma Q$$

Replacement. Most significantly, decreasing a subproof, decreases the whole proof:

$$p \triangleright q > q' \Rightarrow \exists p' \in \mathbb{P}. p > p' \triangleright q'$$

As a consequence of Replacement:

Lemma 4 *Once redundant, always redundant:*

$$A \lesssim A' \Rightarrow A' \cap \rho A \subseteq \rho A'$$

Theorem 6

$$A^b = \Delta (\mu \overline{\Pi} A \cap \hat{A})$$

Proof: Suppose $a \in A^b$. Then there is some proof $p \in \mu\overline{\Pi} A^b$ such that $p \triangleright \widehat{a}$. Were \widehat{a} not minimal, then by the Replacement Postulate, neither would p be minimal. So, $\widehat{A}^b \subseteq \mu\overline{\Pi} A$. Clearly $\widehat{A}^b \subseteq \widehat{A}$. Hence, $A^b = \Delta \widehat{A}^b \subseteq \Delta (\mu\overline{\Pi} A \cap \widehat{A})$. For the other direction, suppose $c \in \Delta (\mu\overline{\Pi} A \cap \widehat{A})$. Then $c \in \Gamma (\mu\overline{\Pi} A \cap \widehat{A}) \subseteq \Gamma \mu\overline{\Pi} A = A^b$. \square

Substituting the definition of A^\sharp :

Corollary 2 *The canonical presentation is the set of conclusions of all trivial normal-form proofs:*

$$A^\sharp = \Delta (\underline{\Pi} A \cap \widehat{\Theta A})$$

8. Inference

There are two basic applications for saturation-based inference: constructing a finite canonical presentation when such exists, and searching for proofs by forward reasoning from axioms, avoiding inferences that do not help saturate. Inference steps are defined by deduction mechanisms:

A *deduction mechanism* \rightsquigarrow is a function from presentations to presentations and we call the relation $A \rightsquigarrow B$ a *deduction step*. (We consider only functional mechanisms in this paper.)

In our ground resolution calculus, we have two deduction rules: $\frac{\ell \vee f}{f \vee f'}$ and $\frac{f}{\ell \vee f}$, plus weakening.

Definition 12 (Soundness and Adequacy) A *deduction mechanism* \rightsquigarrow is *sound and adequate* if:

$$A \rightsquigarrow A' \Rightarrow A \equiv A'$$

We only consider sound, adequate mechanisms:

Definition 13 (Derivation) A *derivation* is a chain of *sound and adequate deductions*:

$$A_0 \rightsquigarrow A_1 \rightsquigarrow \dots \rightsquigarrow A_j \rightsquigarrow \dots$$

(We consider only ω -chains in this paper.)

Formulae that appear at some stage of a derivation and persist are deemed “persistent”:

Definition 14 (Persistent Formulae) The limit A_∞ of a derivation $\{A_i\}_i$ is its *persistent formulae*:

$$A_\infty = \limsup_{i \rightarrow \infty} A_i = \bigcup_{i \rightarrow \infty} \bigcap_{j > i} A_j$$

We are interested in the ability to derive minimal proofs:

Definition 15 (Saturating and Completing Derivations)

A *derivation* $\{A_i\}_i$ is *sound and adequate* if $A_0 \equiv A_\infty$. A *sound and adequate derivation* $\{A_i\}_i$ is *saturating* if $\text{Satur } A_\infty$. It is *completing* if $\text{Compl } A_\infty$.

Completing means that, at the limit, there is at least one minimal proof per theorem; saturating means that all minimal proofs are supported at the limit.

Definition 16 (Simplifying) A *deduction mechanism* \rightsquigarrow is *simplifying* if *proofs only get better*:

$$\rightsquigarrow \subseteq \succsim$$

That is, $\overline{\Pi} A \supseteq \overline{\Pi} A'$ whenever $A \rightsquigarrow A'$.

Since the proof ordering is well-founded:

Lemma 5 *If a deduction mechanism is simplifying then*

$$\overline{\Pi} A_i \supseteq \overline{\Pi} A_\infty \text{ and therefore } \Theta A_i \subseteq \Theta A_\infty$$

for all i in a derivation $\{A_i\}_i$.

Proof: Let $p_i \in \mu\overline{\Pi}_c A_i$. Since the derivation is simplifying, there are proofs $p_j \in \overline{\Pi}_c A_j$, $j > i$, such that $p_i \geq p_{i+1} \geq \dots$. By well-foundedness, from some point on these are all the same proof q . Thus, $\Gamma q \subseteq A_\infty$ and $q \in \overline{\Pi} A_\infty$. \square

Note. For non-simplifying derivations this is not the case. To wit, let

$$\mathbb{P} = \left\{ \frac{a}{b}, \frac{b}{a} \right\}$$

and consider $a \rightsquigarrow b \rightsquigarrow a$.

Lemma 6 *A sufficient condition for a simplifying derivation $\{A_i\}_i$ to be completing is that each non-normal-form proof eventually becomes much better:*

$$\bigcup_i \mu\overline{\Pi} A_i \setminus \underline{\Pi} A_0 \sqsubset \bigcup_i \overline{\Pi} A_i$$

Proof: By Lemma 5, if $p_i \in \mu\overline{\Pi}_c A_i$ then $q \in \overline{\Pi}_c A_\infty$, for some q . If $q \in \underline{\Pi} A_0$ then $c \in \Delta (\overline{\Pi} A_\infty \cap \underline{\Pi} A_0)$ and we are done. Otherwise, the sufficient condition implies that for some k , there is a proof $q_k \in \overline{\Pi} A_k$ of c such that $p_i \geq q > q_k$. Completeness follows by induction on proofs. \square

Though weaker conditions may suffice, we assume that *ordered proof systems have finitely-based proofs, in the sense that they use only a finite number of assumptions:*

$$\forall p \in \mathbb{P}. |\Gamma p| < \infty$$

Lemma 7 *Any simplifying derivation $\{A_i\}_i$ is sound and adequate:*

$$A_i \equiv A_\infty$$

Proof: Lemma 5 gives one direction, namely $\Theta \widehat{A}_i \subseteq \Theta A_\infty$. Suppose $p \in \mu \overline{\Pi} A_\infty$. Consider $a \in \Gamma p$. By the properties of subproofs, \widehat{a} must be in normal form. Since the derivation is simplifying, a persists from the moment it appears in an A_i . We are assuming that there are only finitely many such a . So from some k on, $p \in \overline{\Pi} A_k$. But $A_k \equiv A_i$. \square

Note. This does not necessarily hold for infinitary systems. Let all proofs be incomparable, including: \widehat{a}_i (for all i), $\frac{\widehat{a}_j}{a_i}$ (for all i, j), $\frac{\widehat{a}_0, \widehat{a}_1, \dots}{c}$, and $\frac{}{c}$. The derivation $\{a_j : j \leq i\}_i$ is simplifying, but only its limit includes the infinitary proof.

A derivation is completing if every theorem of A_0 eventually admits a persistent normal-form proof:

$$\Theta A_0 \subseteq \Delta (\overline{\Pi} A_\infty \cap \underline{\Pi} A_0) \quad (9)$$

Definition 17 (Canonical Derivations) A derivation $\{A_i\}_i$ is reducing if its persistent equations are all reduced:

$$\rho A_\infty = \emptyset$$

It is canonical if it is both saturating and reducing.

In other words, the limit of a reducing derivation is flat: $A_\infty = A_\infty^b$.

Lemma 8 A sufficient condition for a simplifying derivation $\{A_i\}_i$ to be reducing is that no formula remain persistently redundant:

$$\rho A_i \cap A_\infty = \emptyset$$

Proof: A formula r is redundant at the limit only if it has a better proof than the trivial one. That (finitely-based) proof must have come into existence at some stage k , at which point r would have already become redundant. This condition then precludes the persistence of r . \square

Lemma 9 A simplifying derivation $\{A_i\}_i$ is canonical iff

$$A_\infty = A_0^\sharp$$

Proof: We already know (Lemma 7) that $A_0 \equiv A_\infty$ and, therefore (Theorem 1.2), that $A_0^\sharp = A_\infty^\sharp$.

So, if $A_\infty = A_0^\sharp$, then the result of the derivation is saturated and reduced. On the other hands, if $A_\infty = A_\infty^\sharp$, then $A_0^\sharp = A_\infty$. \square

Now, we need to ensure that a deduction mechanism makes continual progress with respect to proof quality and that no useful inferences are forever ignored:

Definition 18 (Fairness) A sound and adequate functional deduction mechanism δ is progressive if it makes every non-normal-form proof better:

$$\mu \overline{\Pi} A \setminus \underline{\Pi} A \sqsubset \overline{\Pi} (A \cup \delta A)$$

A derivation $\{A_i\}_i$ is fair for a progressive mechanism δ if all persistently progressive formulæ are derived:

$$\delta A_\infty \subseteq \bigcup_i A_i$$

Theorem 7 Fair simplifying derivations are complete.

Proof: Suppose $c \in \Theta A_0$. By Lemma 7, it has a proof $p \in \overline{\Pi} A_\infty$. If $p \in \underline{\Pi} A_\infty$, we are done. So assume $p \in \overline{\Pi} A_\infty \setminus \underline{\Pi} A_\infty$. The progressive mechanism δ guarantees the existence of a smaller proof $q \in \overline{\Pi} (A_\infty \cup \delta A_\infty)$. Since proofs are finite, each formula in Γq appears in $\bigcup_{i \leq n} A_i$ for some n . Since the derivation is simplifying, if $a \in A_i$, then for all $j \geq i$, $\widehat{a} \geq q_j$ for some proof $q_j \in \overline{\Pi} A_j$. By Replacement, there is a proof $r \in \overline{\Pi} A_n$ such that $p > q \geq r$. By induction, we eventually get a minimal proof of c . \square

9. Completion

The central concept underlying completion [26] is the existence of critical proofs. Completion alternates “expansions” that infer the conclusions of critical proofs with “contractions” that remove redundancies.

Definition 19 (Expansion and Contraction) A deduction step $A \rightsquigarrow A \cup B$ is an expansion provided

$$B \subseteq \Theta A$$

A deduction step $A \cup B \rightsquigarrow A$ is a contraction provided

$$A \cup B \succsim A$$

With the proof ordering we gave for the resolution calculus example, subsumption elimination and tautology elimination serve as contractions.

Proposition 4 Expansions and contractions are sound and adequate.

Definition 20 (Critical Proof) A minimal proof $p \in \mu \overline{\Pi} A$ is critical if it is not in normal form, but all its subproofs are:

$$\begin{aligned} p &\in \mu \overline{\Pi} A \setminus \underline{\Pi} A \\ p \triangleright q &\Rightarrow q \in \underline{\Pi} A \end{aligned}$$

Let

$$\nabla A \stackrel{\dagger}{=} \left\{ \Gamma p' : \begin{array}{l} p \text{ is critical for } A \text{ and} \\ p' \text{ is any proof such that} \\ p' < p \end{array} \right\}$$

Lemma 10 If $\nabla A \subseteq \delta A \subseteq \Theta A$, then δ is progressive.

Proof: Every $p \in \mu\overline{\Pi} A \setminus \underline{\Pi} A$ is either bettered by the addition of ∇A or, by induction on the subproof relation, has a subproof that is bettered. In the latter case, by Replacement, p is also improved. \square

In particular, ∇ is progressive.

We are now able to describe an inference system, traditionally called “completion”, wherein each step $A \rightsquigarrow A'$ is the composition of an expansion, $A \rightsquigarrow A \cup \nabla A = B$ followed by a contraction, $B \rightsquigarrow B^b = A'$.

Bulk completion (cf. [3, 28–29]) is a sequence of steps:

$$A \rightsquigarrow [A \cup \nabla A]^b$$

Lemma 11 *Bulk completion is simplifying.*

Proof: $A \rightsquigarrow A' \Rightarrow A \lesssim A \cup \nabla A \lesssim [A \cup \nabla A]^b = A'$. \square

Lemma 12 *The canonical presentation has no critical formulae: $\nabla(A^\sharp) = \emptyset$.*

Proof: $\mu\overline{\Pi} A^\sharp \setminus \underline{\Pi} A^\sharp = \emptyset$. \square

Corollary 3 *The canonical presentation is stable under bulk completion:*

$$A^\sharp \rightsquigarrow A' \Rightarrow A' = A^\sharp$$

Theorem 8 *Bulk completion is canonical, if minimal proofs are unique:*

$$A_0^\sharp = A_\infty^{\text{bulk}}$$

Proof: Bulk completion derives all progressive formulae immediately, so is fair. Completeness follows from Theorem 7. By Proposition 2, the limit is also saturated. Bulk completion also removes redundancies immediately, so is reducing by Lemma 8. \square

Fair completion is a derivation that is fair for ∇ , and which eventually deletes every redundancy.

Theorem 9 *Fair completion is canonical, if minimal proofs are unique:*

$$A_0^\sharp = A_\infty^{\text{fair}}$$

10. Conclusion

The focus of this paper was our definition of the canonical presentation for any deductive theory supplied with a proof ordering. It is exactly what is needed for all theorems to enjoy normal-form proofs. The structure of normal-form

(or “direct”) proofs is fixed by the ordering under which they are minimal. We have given alternate characterizations of the canonical presentation, derived many of its properties, and shown how it can be generated.

The notion of “saturation” in theorem proving, in which redundant deductions are not necessary for completeness, was suggested by Rusinowitch [32], and pursued most recently in [8, 29]. Our definitions generalize the redundancy-based notion of saturation to arbitrary proof orderings. See [14].

Completion processes have been studied intensively since their independent discovery and application to automated theorem proving by Knuth [26] and Buchberger [9]. The fundamental role of proof orderings in automated deduction was conceived in [6]. The completion principle can be applied to numerous situations [13], including equational rewriting [4, 22, 30], induction [25], and unification [18]. Our abstract framework can be applied to re-understand completion mechanisms in a fully uniform setting. Because we have been generic in approach, the results here apply in any completion-based framework, including standard completion mechanisms like ground completion [34] (see [14]), equational completion, or completion for unification, and also to derive new completion algorithms, such as for constraint solving. (For constraint solving, one may—contrariwise—seek out the “interesting” redundant formulae in order to favor local checks.)

Finally, it bears mentioning that, thanks to the Curry-deBruijn-Howard morphism, one can consider a proof p as a term whose type is precisely its conclusion Δp . Considering proof orderings would then be related to the definition of an ordering on higher-order terms, as studied, for example, in [23], or, for dependently typed terms, in [10, 37].

Acknowledgments

We thank Arnon Avron for asking what corresponds to rewrite-system completion in other deductive settings and anonymous referees for corrections and suggestions.

References

- [1] M. Aiguier, D. Bahrani, and C. Dubois. Axioms for rewriting theory. In N. Dershowitz and C. Kirchner, editors, *RULE 2000: Proc. 1st International Workshop on Rule-Based Programming*, Montreal, Canada, Sept. 2000.
- [2] F. Baader and T. Nipkow. *Term Rewriting and All That*. Cambridge University Press, 1998.
- [3] L. Bachmair. *Canonical Equational Proofs*. Birkhäuser, Boston, 1991.
- [4] L. Bachmair and N. Dershowitz. Completion for rewriting modulo a congruence. *Theoretical Comput. Sci.*, 67(2–3):173–202, Oct. 1989.

- [5] L. Bachmair and N. Dershowitz. Equational inference, canonical proofs, and proof orderings. *J. ACM*, 41(2):236–276, 1994.
- [6] L. Bachmair, N. Dershowitz, and J. Hsiang. Orderings for equational proofs. In *Proc. 1st Symposium on Logic in Computer Science (Cambridge, MA)*, pages 346–357. IEEE, June 1986.
- [7] L. Bachmair, N. Dershowitz, and D. Plaisted. Completion without failure. In H. Ait-Kaci and M. Nivat, editors, *Resolution of Equations in Algebraic Structures 2: Rewriting Techniques*, pages 1–30, New York, 1989. Academic Press.
- [8] L. Bachmair and H. Ganzinger. Resolution theorem proving. In A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, volume I, chapter 2, pages 19–99. Elsevier Science, 2001.
- [9] B. Buchberger. Gröbner-bases: An algorithmic method in polynomial ideal theory. In N. K. Bose, editor, *Multidimensional Systems Theory*, chapter 6, pages 184–232. Reidel, Dordrecht, 1985.
- [10] H. Cirstea, C. Kirchner, and L. Liquori. The Rho cube. In F. Honsell, editor, *Proc. Foundations of Software Science and Computation Structures*, Lecture Notes in Computer Science, pages 166–180, Genova, Italy, Apr. 2001.
- [11] H. Comon and C. Kirchner. Constraint solving on terms. In H. Comon, C. Marché, and R. Treinen, editors, *Proc. Constraints in Computational Logics. Theory and Applications*, volume 2002, chapter 2. Lecture Notes in Computer Science, 2001.
- [12] N. Dershowitz. Orderings for term-rewriting systems. *Theoretical Comput. Sci.*, 17(3):279–301, Mar. 1982.
- [13] N. Dershowitz. Completion and its applications. In H. Ait-Kaci and M. Nivat, editors, *Resolution of Equations in Algebraic Structures, Volume 2: Rewriting Techniques*, pages 31–86. Academic Press, 1989.
- [14] N. Dershowitz. Ground canonicity. Unpublished, 2003. <http://arXiv.org/abs/cs.LO/0304017>.
- [15] N. Dershowitz, L. Marcus, and A. Tarlecki. Existence, uniqueness and construction of rewrite systems. *SIAM Journal of Computing*, 17(4):629–639, Aug. 1988.
- [16] N. Dershowitz and M. Okada. Proof-theoretic techniques and the theory of rewriting. In *Proc. 3rd Symposium on Logic in Computer Science (Edinburgh, UK)*, pages 104–111. IEEE, 1988.
- [17] N. Dershowitz and D. A. Plaisted. Rewriting. In A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, volume I, chapter 9, pages 535–610. Elsevier Science, 2001.
- [18] N. Doggaz and C. Kirchner. Completion for unification. *Theoretical Comput. Sci.*, 85(1):231–251, 1991.
- [19] I. Gnaedig, C. Kirchner, and H. Kirchner. Equational completion in order-sorted algebras. In M. Dauchet and M. Nivat, editors, *Proc. 13th Colloquium on Trees in Algebra and Programming*, volume 299 of *Lecture Notes in Computer Science*, pages 165–184, Nancy, France, 1988. Springer-Verlag.
- [20] J. Hsiang and M. Rusinowitch. Proving refutational completeness of theorem proving strategies. The transfinite semantic tree method. *J. ACM*, 38(3):559–587, July 1991.
- [21] G. Huet. A complete proof of correctness of the Knuth–Bendix completion algorithm. *J. Comput. Syst. Sci.*, 23(1):11–21, Aug. 1981.
- [22] J.-P. Jouannaud and H. Kirchner. Completion of a set of rules modulo a set of equations. *SIAM Journal of Computing*, 15(4):1155–1194, 1986.
- [23] J.-P. Jouannaud and A. Rubio. The higher-order recursive path ordering. In G. Longo, editor, *Proc. 14th Annual Symposium on Logic in Computer Science*, pages 402–411, Trento, Italy, 1999. IEEE.
- [24] A. Kandri-Rody, D. Kapur, and F. Winkler. Knuth-Bendix procedures and Buchberger algorithm—A synthesis. In *Proc. 20th Intl. Symp. on Symbolic and Algebraic Computation (Portland, Oregon)*, pages 55–67, 1989.
- [25] D. Kapur and D. R. Musser. Proof by consistency. *J. of Artificial Intelligence*, 13(2):125–157, 1987.
- [26] D. E. Knuth and P. B. Bendix. Simple word problems in universal algebras. In J. Leech, editor, *Computational Problems in Abstract Algebra*, pages 263–297. Pergamon Press, Oxford, 1970.
- [27] D. S. Lankford. Canonical inference. Memo ATP-32, Automatic Theorem Proving Project, University of Texas, Austin, TX, Dec. 1975.
- [28] Y. Métivier. About the rewriting systems produced by the Knuth-Bendix completion algorithm. *Inf. Process. Lett.*, 16(1):31–34, Jan. 1983.
- [29] R. Nieuwenhuis and A. Rubio. Paramodulation-based theorem proving. In A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, volume I, chapter 7, pages 371–443. Elsevier Science, 2001.
- [30] G. Peterson and M. Stickel. Complete sets of reductions for some equational theories. *J. ACM*, 28(2):233–264, 1981.
- [31] J.-C. Raoult. On graph rewritings. *Theoretical Comput. Sci.*, 32(1,2):1–24, July 1984.
- [32] M. Rusinowitch. *Démonstration Automatique: Techniques de Réécriture*. Science Informatique. InterEditions, Paris, 1989.
- [33] M. B. Smyth. Powerdomains. *J. Comput. Syst. Sci.*, 16:23–36, 1977.
- [34] W. Snyder. Efficient ground completion: An $O(n \log n)$ algorithm for generating reduced sets of ground rewrite rules equivalent to a set of ground equations E . In N. Dershowitz, editor, *Proc. 3rd Conference on Rewriting Techniques and Applications (Chapel Hill, NC)*, volume 355 of *Lecture Notes in Computer Science*, pages 419–433. Springer-Verlag, Apr. 1989.
- [35] G. Struth. *Canonical Transformations in Algebra, Universal Algebra and Logic*. Dissertation, Institut für Informatik, Universität des Saarlandes, Saarbrücken, Germany, June 1998.
- [36] “Terese” (M. Bezem, J. W. Klop and R. de Vrijer, eds.). *Term Rewriting Systems*. Cambridge University Press, 2002.
- [37] R. Virga. *Higher-Order Rewriting with Dependent Types*. PhD thesis, Department of Mathematical Sciences, Carnegie Mellon University, Sept. 1999. Available as Technical Report CMU-CS-99-167.