# Abstract Canonical Inference

MARIA PAOLA BONACINA

Dipartimento di Informatica, Università degli Studi di Verona

and

NACHUM DERSHOWITZ

School of Computer Science, Tel Aviv University

An abstract framework of canonical inference is used to explore how different proof orderings induce different variants of saturation and completeness. Notions like completion, paramodulation, saturation, redundancy elimination, and rewrite-system reduction are connected to proof orderings. Fairness of deductive mechanisms is defined in terms of proof orderings, distinguishing between (ordinary) "fairness," which yields completeness, and "uniform fairness," which yields saturation.

*They are not capable to ground a canonicity of universal consistency.*

—Alexandra Deligiorgi (ΠΑΙΔΕΙΑ, 1998)

## 1. INTRODUCTION

For effective automated reasoning, the ability to ignore irrelevant data is just as important as the capability to derive consequences from given information. Thus, theorem provers generally incorporate various mechanisms for controlling the growth of the collection of inferred formulæ or derived goals. It is a challenge, however,

to ensure that such rules for simplification or deletion of formulæ do not impinge upon the completeness of the resulting theorem proving strategy.

One class of inference engines that make heavy use of simplification includes the Knuth-Bendix completion procedure for equational inference [Knuth and Bendix 1970] and Buchberger's Gröbner-basis algorithm for polynomial ideals [Buchberger 1985]. These forward-reasoning systems aim at generating sets of formulæ that are "complete" in the sense that completion can provide a rewriting-based decision procedure for validity in the given equational theory, and that the Gröbner basis is similarly used to decide membership in the ideal. Ballantyne (cited in [Dershowitz et al. 1988]) and Metivier [1983] took note of the fact that the fully reduced result of completion is unique for given axioms and term ordering.

Brown [1975], for the Horn case, and Lankford [1975], for the general case, showed how to combine equational completion with clausal resolution improving on the original paramodulation [Robinson and Wos 1969], a line of investigation that later produced methods based on ordered resolution and ordered paramodulation [Hsiang and Rusinowitch 1991; Bachmair and Ganzinger 1994; Nieuwenhuis and Rubio 2001]. Huet [1981] showed how Knuth's completion procedure can also play the rôle of an incomplete prover for equational validity. Hsiang and Rusinowitch [1987] and Bachmair, Dershowitz and Plaisted [Bachmair et al. 1989] designed unfailing versions of completion without compromising the powerful rôle of simplification in controlling the completion process.

In the following sections, we suggest that proof orderings, rather than formula orderings, take center stage in theorem proving with contraction (simplification and deletion of formulæ). Given a specific proof ordering, completeness of a set of formulæ—which we refer to as a *presentation*—will mean that all derivable theorems enjoy a minimal proof, while completeness of an inference system will mean that all formulæ needed as premises in such ideal proofs can be inferred. This formalism is very flexible, since it allows small proofs to use large premises, and vice-versa.

Well-founded orderings of proofs, as developed in [Bachmair and Dershowitz 1994], distinguish between economical, "direct" proofs, namely, those that are of a computational flavor (e.g. rewrite proofs), and expensive "indirect" proofs, those that are discovered after performing a search (e.g. equational proofs). These proof orderings are lifted from orderings on terms and formulæ. Given a formula ordering, one can, of course, choose to compare proofs by simply comparing (the multiset of) their premises.

Our proof-ordering based approach to deduction suggests generalizations of the current concepts of "saturation," "redundancy," and "fairness." Saturated, for us, will mean that *all* inexpensive proofs are supported, as opposed to completeness, which makes do with one minimal proof per theorem. Accordingly, we define two notions of fairness: a *fair* derivation generates a complete set in the limit, while a *uniformly fair* derivation generates a saturated limit. By considering different orderings on proofs, one gets different kinds of saturated sets. The notion of saturation in theorem proving, in which superfluous deductions are not necessary for completeness, was suggested in [Rusinowitch 1991]. In our terminology: A presentation was said to be saturated when all inferrible formulæ are syntactically

subsumed by formulæ in the presentation.[1]

We also define redundancy in terms of the proof ordering, as propounded in [Bonacina and Hsiang 1995]: A sentence is redundant if adding it to the presentation does not decrease any minimal proof. (See Bonacina 1992, Chap. 2.) The definition of redundancy in [Bachmair and Ganzinger 1994]—an inference is redundant if its conclusion can be inferred from smaller formulæ—coincides with ours when proofs are measured first by their maximal premises. In [Bachmair and Ganzinger 1994; 2001; Nieuwenhuis and Rubio 2001], saturated means that every possible inference is redundant.

The present work continues the development of an abstract theory of "canonical inference," begun in [Dershowitz and Kirchner 2006], which, in turn, grew out of the theory of rewriting (see, for example, Dershowitz and Plaisted 2001; Terese 2003) and deduction (see, for example, Bonacina 1999; Bachmair and Ganzinger 2001; Nieuwenhuis and Rubio 2001). Although we will use ground equations as an illustrative example, this framework applies equally well in the first-order setting, whether equational or clausal. Our motivations and contributions are primarily æsthetic and intellectual:

— organizing the theory of "canonical inference" in an architecture with primitive objects (such as presentations and proofs), their properties (canonical presentations, normal-form proofs), mappings between objects (inferences, derivations), their properties (good inferences, fair derivations), and theorems that state the weakest possible sufficient conditions for the desirable properties;

— keeping the treatment throughout as abstract as possible, so as to maximize generality, without losing sight of concrete instances;

— providing a terminology that is simultaneously general and precise; and

— assembling a notation that is at the same time elegant, compact, and helpful.

Since good theory produces the simplicity of concepts and clarity of priorities that are key to the building of strong systems, our hope is that this work might also nurture practical applications.

The next section sets the stage, with basic notions and notations, and introduces a running example. To keep this article self-contained, Section 3 recapitulates relevant definitions and results from [Dershowitz and Kirchner 2006].[2] Specifically, the *canonical* basis of an abstract deductive system is defined in three equivalent ways: (1) formulæ appearing in minimal proofs; (2) minimal trivial theorems; (3) nonredundant lemmata. Section 4 articulates the abstract framework, by introducing inferences and proof procedures, providing proofs with structure, and characterizing good inference sequences. Sections 5–7 carry out the study of derivation and

---

[1]In [Rusinowitch 1991], the language is clausal, and a clause $C$ subsumes a clause $D$ if there is a substitution $\sigma$ such that $C\sigma \subseteq D$ and $C$ does not have more literals than does $D$. We refer to this as "syntactic subsumption" to distinguish it from the general semantic principle, under which $C$ subsumes $D$ if $\models \forall \bar{x} C \Rightarrow \forall \bar{y} D$, where $\bar{x}$ and $\bar{y}$ are the variables of $C$ and $D$, respectively.

[2]The study in [Dershowitz and Kirchner 2006] is concerned with defining abstract properties of sets of formulæ. It is extended here with notions, such as fairness, that describe properties of *derivations*. That article is about properties of *objects* (presentations); we study properties required of *processes* (derivations) so as to generate the desired presentations.

completion processes. Finally, we close with a discussion in Section 8, including related work and connections with the praxis of theorem proving.

## 2.   ORDERED PROOF SYSTEMS

Let $\mathbb{A}$ be the set of all formulæ (ground equations and disequations, in our examples) over some fixed vocabulary. Let $\mathbb{P}$ be the set of all (ground equational) proofs. These sets of abstract objects are linked by two functions: $Pm : \mathbb{P} \to 2^{\mathbb{A}}$ gives the premises (assumptions) in a proof, and $Cl : \mathbb{P} \to \mathbb{A}$ gives its conclusion. For example, if $p \in \mathbb{P}$ is a proof of $a = b, a = c \vdash f(b, c) = f(c, b)$, then $[p]^{Pm}$ is $\{a = b, a = c\}$ and $[p]_{Cl}$ is $f(b, c) = f(c, b)$. Both functions extend to sets of proofs in the usual fashion.

The framework proposed here is predicated on two *well-founded* partial orderings over $\mathbb{P}$: a *proof ordering* $\geq$ and a *subproof relation* $\unrhd$. They are related by a monotonicity requirement given below (Eq. 7). If the best proof of a theorem $c$ requires some lemma $b$, this monotonicity condition precludes the possibility that the best proof of $b$ turn around and use $c$, since then ultimately both $b$ and $c$ would be needed to support all ideal proofs, and there would be no "localized" way of knowing when a formula is never needed and truly redundant. On the other hand, this monotonicity condition does allow $b$ to be better in some proof contexts and $c$ in others.

For convenience, we assume that the proof ordering only compares proofs with the same conclusion ($p \geq q \Rightarrow [p]_{Cl} = [q]_{Cl}$), rather than mention this condition each time we have cause to compare proofs.

We use the standard notation $A \vdash c$, for premises $A \subseteq \mathbb{A}$ and conclusion $c \in \mathbb{A}$, to mean that there exists a proof $p \in \mathbb{P}$ such that $[p]^{Pm} = A$ and $[p]_{Cl} = c$. We will use the term *presentation* to mean a set of formulæ, and *justification* to mean a set of proofs. Given a presentation $A$, the set of all proofs using all *or some* premises of $A$ is denoted by:[3]

$$Pf(A) \quad \overset{!}{=} \quad \{p \in \mathbb{P} : [p]^{Pm} \subseteq A\}$$

We reserve the term *theory* for deductively-closed presentations. Let $Th\ A$ denote the theory of presentation $A$, that is, the set of conclusions of all proofs with premises in $A$:

$$Th\ A \quad = \quad \{[p]_{Cl} : p \in \mathbb{P},\ [p]^{Pm} \subseteq A\} \quad = \quad [Pf(A)]_{Cl}$$

Presentations $A$ and $B$ are *equivalent* ($A \equiv B$) if their theories are identical ($Th\ A = Th\ B$).

We presume the following standard properties of Tarskian consequence relations:

$$A \vdash c \ \Rightarrow \ A \cup B \vdash c \tag{1}$$

$$A \ \subseteq \ Th\ A \tag{2}$$

$$Th\ Th\ A \ = \ Th\ A \tag{3}$$

for all $A$, $B$ and $c$. It follows from the definition of $Th$ that

$$Th\ A \quad \subseteq \quad Th\ (A \cup B) \tag{4}$$

---

[3]We use $\overset{!}{=}$ and $\overset{!}{\equiv}$ for definitions.

Thus, *Th* is a closure operation. On account of the (left) weakening property of Eq. (1), we need not distinguish between $A \vdash c$, meaning that there is a proof of $c$ using all the premises $A$, or using just some.

As a very simple running example, let the vocabulary consist of the constant $0$ and unary symbol $s$. Abbreviate tally terms $s^i 0$ as numeral $i$. The set $\mathbb{A}$ consists of all *unordered* equations $i = j$; so symmetry is built into the structure of proofs (we postpone dealing with disequations for the time being). An equational inference system (with this vocabulary) might consist of the following five inference rules:

$$\frac{\Box}{0 = 0}\ \mathbf{Z} \qquad \frac{\boxed{i = j}}{i = j}\ \mathbf{I}_{i=j}$$

$$\frac{i = j}{si = sj}\ \mathbf{S} \qquad \frac{a\ \ c}{c}\ \mathbf{P} \qquad \frac{i = j\ \ j = k}{i = k}\ \mathbf{T}$$

where boxes surround premises, $\mathbf{Z}$ is an axiom, $\mathbf{I}$ introduces premises, and $\mathbf{S}$ infers $i + 1 = j + 1$ from a proof of $i = j$. Proof-tree branches of the transitivity rule $\mathbf{T}$ are *unordered*. Projections $\mathbf{P}$ allow irrelevant premises to be ignored and are needed to accommodate monotonicity, that is, Eq. (1).

For example, if $A = \{4 = 2, 4 = 0\}$, then

$$Th\, A = \{i = j : i \equiv j \pmod 2\}$$

Consider the proof schemata:

$$\frac{\Box}{0 = 0}$$
$$\frac{}{1 = 1}$$
$$\vdots$$
$$\frac{}{i = i}$$
$$\frac{\boxed{4 = 2}}{4 = 2}$$
$$\frac{}{5 = 3}$$
$$\vdots$$
$$\frac{}{i + 4 = i + 2}$$
$$\frac{\dfrac{\boxed{4 = 0}}{4 = 0}\ \dfrac{\boxed{4 = 2}}{4 = 2}\ \dfrac{p_0}{i - j - 1 = 1}}{2 = 0\qquad i - j = 2}$$
$$\frac{}{i - j = 0}$$
$$\vdots$$
$$\frac{}{i = j}$$

where $p_0$ is a proof of $i - j - 2 = 0$. Let's use proof terms for proofs, denoting the above three trees (from left to right) by $p = S^i Z$, $q = S^i I(4, 2)$ and $r = S^j T(T(I(4,0), I(4,2)), SS(p_0))$. Thus, $[p]^{Pm} = \emptyset$, $[q]^{Pm} = \{4 = 2\}$, and $[r]_{Cl}$ is the formula $i = j$.

With a (multiset) recursive path ordering [Dershowitz 1982] to order proofs, and a precedence $Z < S < T < I < P < 0 < 1 < 2 < \cdots$ on proof combinators and vocabulary symbols, the minimal proof of a theorem in *Th A* takes one of the forms

$$S^j \left( \nabla_{4k=0} \right) \qquad S^j \left( \nabla_{4k=2} \right)$$

where the subproofs $\nabla_{4k=0}$ and $\nabla_{4k=2}$ are defined recursively:

$$
\begin{array}{llll}
\nabla_{0=0} &=& Z & \quad \nabla_{0=2} &=& T(\nabla_{4=0}, \nabla_{4=2}) \\
\nabla_{4=0} &=& I(4,0) & \quad \nabla_{4(k+1)=0} &=& T(S^{4k}\nabla_{4=0}, \nabla_{4k=0}) \\
\nabla_{4=2} &=& I(4,2) & \quad \nabla_{4(k+1)=2} &=& S^2 T(\nabla_{0=2}, S^2 \nabla_{4k=0})
\end{array}
$$

We call a proof *trivial* when it proves its only premise and has no subproofs other than itself, that is, if $[p]^{Pm} = \{[p]_{Cl}\}$ and $p \trianglerighteq q \Rightarrow p = q$. We denote by $a \vdash a$ or $\hat{a}$

such a trivial proof of $a \in \mathbb{A}$, and by $\widehat{A}$ the set of trivial proofs of each $a \in A$. For example, $\widehat{4{=}0} = I(4, 0)$.

We assume that premises appear in proofs (5), that subproofs do not use non-extant premises (6), and that proof orderings are monotonic with respect to (replacement of) subproofs (7). Specifically, for all proofs $p, q, r$ and formulæ $a$:

$$a \in [p]^{Pm} \;\Rightarrow\; p \trianglerighteq \widehat{a} \tag{5}$$

$$p \trianglerighteq q \;\Rightarrow\; [p]^{Pm} \supseteq [q]^{Pm} \tag{6}$$

$$p \triangleright q > r \;\Rightarrow\; \exists v \in Pf([\{p, r\}]^{Pm}).\; p > v \triangleright r \tag{7}$$

We make no other assumptions regarding proofs or their structure.

The intuition for assumption (5), "proofs use their premises," is related to the distinction between proof and derivation. Informally, a derivation contains all formulæ generated by a deduction mechanism from a given input, while a proof of a formula generated during the derivation contains all, and only, the formulæ involved in inferring that formula within that proof (derivations will be treated formally in Section 5). The *Replacement Postulate* (7) states that $\triangleright$ and $>$ (which we have restricted to proofs with the same conclusion) commute. In other words, "replacing" a subproof $q$ of a proof $p$ with a strictly smaller proof $r$ "results" in a proof $v$ that is smaller than the original $p$, and which does not involve extraneous premises. This postulate implies the following weaker commutation property:

$$p \trianglerighteq q > r \;\Rightarrow\; \exists v \in Pf([\{p, r\}]^{Pm}).\; p > v \trianglerighteq r \tag{8}$$

Most proof orderings in the literature obey this monotonicity requirement.

Every formula $a$ admits a trivial proof $a \vdash a$ by Eqs. (2) and (5). On account of (5 and 7), proofs are also monotonic with respect to any inessential premises they refer to, should the latter admit smaller than trivial proofs.

It may be convenient to think of a proof-tree "leaf" as a subproof with only itself as a subproof; other subproofs are the "subtrees." There are two kinds of leaves: trivial proofs $a \vdash a$ (such as inferences **I**), and vacuous proofs (axioms) $\bar{a}$ with $[\bar{a}]^{Pm} = \emptyset$ and $[\bar{a}]_{Cl} = a$ (such as **Z**). By well-foundedness of $\trianglerighteq$, there are no infinite "paths" in proof trees. It follows from Replacement (7) that the transitive closure of $> \cup \triangleright$ is also well-founded.

## 3. CANONICAL PRESENTATIONS

The results in this section are extracted from [Dershowitz and Kirchner 2006], which should be consulted for proofs not given here.

Define the *minimal* proofs in a set of proofs as:

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

On account of well-foundedness, minimal proofs always exist.

Note that $Pm$, $Cl$, $Th$ and $Pf$ are all monotonic with respect to set inclusion, but $\mu Pf$ is not. Indeed, $A \subseteq B$ does not imply $\mu Pf(A) \subseteq \mu Pf(B)$, and $P \subseteq Q$ does not imply $\mu P \subseteq \mu Q$, because a proof $p$ that is minimal in $P$ need not be minimal in $Q$, since $Q$ may contain a $q < p$ such that $q \notin P$. Also, $\mu P \subseteq \mu Q$ does not imply $P \subseteq Q$, since $P$ may contain all sorts of non-minimal proofs not in $Q$.

We say that presentation $A$ is *contracted* when $A = [\mu Pf(A)]^{Pm}$, that is, $A$ contains precisely the premises used in minimal proofs based on $A$. By a "normal-form proof," we mean a minimal proof using *any* theorem as a *lemma* (that is, as a premise):

*Definition* 3.1 *(Normal-Form Proof)*. The *normal-form proofs* of a presentation $A$ are the set

$$Nf(A) \quad \overset{!}{=} \quad \mu Pf(Th\,A)$$

This leads to our main definition:

*Definition* 3.2 *(Canonical Presentation)*. The *canonical presentation* $A^\sharp$ of $A$ contains those formulæ that appear as premises of normal-form proofs:

$$A^\sharp \quad \overset{!}{=} \quad [Nf(A)]^{Pm}$$

So, we will say that $A$ is *canonical* if $A = A^\sharp$.

It follows from the definitions that

$$Nf(A) = \mu Pf(A^\sharp) \subseteq Pf(A^\sharp) \tag{9}$$

The next proposition gives a second characterization of the canonical presentation—as normal-form trivial theorems:

PROPOSITION 3.3.

$$A^\sharp = [Nf(A) \cap \widehat{Th\,A}]_{Cl}$$
$$\widehat{A^\sharp} = Nf(A) \cap \widehat{Th\,A}$$

THEOREM 3.4. *The function* $\_^\sharp$ *is "canonical" with respect to the equivalence of presentations. That is:*

$$
\begin{aligned}
A^\sharp &\equiv A & \text{(Consistency)} \\
A \equiv B &\Leftrightarrow A^\sharp = B^\sharp & \text{(Monotonicity)} \\
A^{\sharp\sharp} &= A^\sharp & \text{(Idempotence)}
\end{aligned}
$$

By lifting proof orderings to justifications and presentations, the canonical presentation can be characterized directly in terms of the ordering. First, proof orderings are lifted to sets of proofs, as follows:

*Definition* 3.5.

— Justification $Q$ is *better* than justification $P$ if:

$$P \sqsupseteq Q \quad \overset{!}{\equiv} \quad \forall p \in P.\, \exists q \in Q.\ p \geq q$$

— It is *much better* if:

$$P \sqsupset Q \quad \overset{!}{\equiv} \quad \forall p \in P.\, \exists q \in Q.\ p > q$$

— Two justifications are *similar* if:

$$P \simeq Q \quad \overset{!}{\equiv} \quad P \sqsupseteq Q \sqsupseteq P$$

Recall that only proofs with the same conclusion are compared by proof orderings.

Transitivity of these three relations follows from the definitions. They are compatible: $(\sqsupseteq \circ \sqsupset) \subseteq \sqsupset$, $(\sqsupseteq \circ \simeq) \subseteq \sqsupseteq$, etc. Since it is also reflexive, $\sqsupseteq$ is a quasi-ordering. Note that $\sqsupset$ is *not* merely the strict version of $\sqsupseteq$, since *every* proof in $P$ must have a strictly smaller one in $Q$.[4]

The next proposition states that subproofs of minimal proofs are minimal, bigger presentations may offer better proofs, and minimal proofs are the best.

PROPOSITION 3.6.

*(a) For all proofs p and q and presentations A:*

$$p \in \mu Pf(A) \ and \ p \rhd q \ \Rightarrow \ q \in \mu Pf(A)$$

*(b) For all presentations A and B:*

$$Pf(A) \ \sqsupseteq \ Pf(A \cup B)$$

*(c) For all justifications P:*

$$P \ \sqsupseteq \ \mu P$$

This "better than" quasi-ordering $\sqsupseteq$ on proofs is lifted to a "simpler than" $\succsim$ quasi-ordering on (equivalent) sets of formulæ, as follows:

*Definition* 3.7.

— Presentation $B$ is *simpler* than an equivalent presentation $A$ when $B$ provides better proofs than does $A$:

$$A \succsim B \ \overset{!}{\equiv} \ A \equiv B \ and \ Pf(A) \sqsupseteq Pf(B)$$

— Presentations are *similar* if their proofs are:

$$A \approx B \ \overset{!}{\equiv} \ Pf(A) \simeq Pf(B)$$

Similarity $\approx$ is the equivalence relation associated with $\succsim$.

These relations are also compatible.

Canonicity may be characterized in terms of this quasi-ordering:

THEOREM 3.8. *The canonical presentation is the simplest:*

$$A \ \succsim \ A^{\sharp}$$

Recalling that all subproofs of normal-form proofs are also in normal form (Proposition 3.6), we propose the following definitions:

*Definition* 3.9 *(Saturation and Completeness).*

— A presentation $A$ is *saturated* if it supports all possible normal-form proofs:

$$\mu Pf(A) \ = \ Nf(A)$$

---

[4]The strict version of $\sqsupseteq$ would say $P \sqsupseteq Q \not\sqsupseteq P$, that is, $\forall p \in P. \exists q \in Q. \ p \geq q$ and $\exists q \in Q. \forall p \in P. \ q < p$. On the other hand, $P \sqsupset Q$ says $\forall p \in P. \exists q \in Q. \ p > q$. This is why we use the term "much better" and not "strictly better."

— A presentation $A$ is *complete* if every theorem has a normal-form proof:

$$Th\ A\ =\ [Pf(A) \cap Nf(A)]_{Cl}$$

It can be shown that:

LEMMA 3.10. *A presentation $A$ is saturated if and only if*

$$Nf(A)\ \subseteq\ Pf(A)$$

A presentation is complete if it is saturated, but for the converse, we need an additional hypothesis: *minimal proofs are unique* if, for all theorems $c \in [Pf(A)]_{Cl}$, there is exactly one proof in $Nf(A)$ with conclusion $c$. In particular, this holds for proof orderings that are total (on proofs of the same theorem). Bear in mind that abstract proofs may be designed to represent whole equivalence classes of concrete proofs.

PROPOSITION 3.11.

*(a) A presentation is complete if it is saturated.*

*(b) If minimal proofs are unique, then a presentation is saturated if and only if it is complete.*

If a theorem has two distinct normal-form proofs $p$ and $q$, a presentation $A$ such that $p \in Pf(A)$, but $q \notin Pf(A)$, may be complete but not saturated. For example, suppose all rewrite (valley) proofs are minimal but incomparable. In that situation, every Church-Rosser system is complete, since every identity has a rewrite proof, but only the full deductive closure is saturated, because for every identity it offers all rewrite proofs.

The next theorem relates canonicity and saturation.

THEOREM 3.12.

*(a) A presentation $A$ is saturated if and only if it contains its own canonical presentation:*

$$A \supseteq A^{\sharp}$$

*In particular, $A^{\sharp}$ is saturated.*

*(b) Moreover, the canonical presentation $A^{\sharp}$ is the smallest saturated set:*
   *— No equivalent proper subset of $A^{\sharp}$ is saturated.*
   *— If $A$ is saturated, then every equivalent superset is also.*

Regarding completeness, we have the following:

THEOREM 3.13. *If $A$ is complete and setwise minimal (i.e. no $B \subsetneq A$, such that $B \equiv A$, is complete), then $A \subseteq A^{\sharp}$.*

PROOF. By way of contradiction, let $c \in A \backslash A^{\sharp}$. Since $A^{\sharp}$ is the set of all premises of normal-form proofs, $c$ is not a premise of any such proof. So, let $B = A \setminus \{c\}$: $B$ has the same normal-form proofs as does $A$, that is, one per theorem. It follows that $B$ is complete, contrary to the hypothesis that $A$ is setwise minimal.  □

PROPOSITION 3.14.

*(a) Presentation A is saturated if and only if Th A ≈ A.*

*(b) Similar presentations are either both saturated or neither is.*

*(c) Similar presentations are either both complete or neither is.*

The following definition sets the stage for the third characterization of canonical presentation—as non-redundant lemmata. Formulæ that can be removed from a presentation—without making proofs worse—are deemed "redundant":

*Definition* 3.15 *(Redundancy).*

— A formula $r$ is *redundant* with respect to a presentation $A$ when:

$$A \succsim A \setminus \{r\}$$

— The set of all *redundant* formulæ of a given presentation $A$ will be denoted as follows:

$$Red\, A \quad \overset{!}{=} \quad \{r \in A : A \succsim A \setminus \{r\}\}$$

— A presentation $A$ is *irredundant* if

$$Red\, A = \emptyset$$

By definition, $Red\, A \subseteq A$.

Thanks to the well-foundedness of $>$ the set of all redundant formulæ in $Red\, A$ is *globally* redundant:

PROPOSITION 3.16. *For all presentations A:*

$$A \approx A \setminus Red\, A$$

Thus, it can be shown that $A$ is *contracted* (i.e. $A = [\mu Pf(A)]^{Pm}$) if and only if it is *irredundant* ($Red\, A = \emptyset$).

Furthermore, every redundant $r \in Red\, A$ has a minimal proof $p \in \mu Pf(A)$, in which it does not appear as a premise ($r \notin [p]^{Pm}$).

The third characterization of the canonical set is central for our purposes:

THEOREM 3.17. *A presentation is canonical if and only if it is saturated and contracted.*

Informally, $A$ is contracted if it is the set of premises of its minimal proofs; it is saturated if minimal proofs in $A$ are exactly the normal-form proofs in the theory; it is canonical if it is the set of premises of normal-form proofs. Hence, saturated plus contracted is equivalent to canonical.

## 4. VARIATIONS ON CANONICITY

The idea we are promoting is that, given a set of axioms, $A$, one is interested in the (unique) set of lemmata, $A^\sharp \subseteq Th\, A$, which—when used as premises in proofs—supports *all* the *normal-form* proofs of the theorems $Th\, A$. These lemmata form *the* "canonical basis" of the theory. In this section, we observe how the canonical basis varies as the proof ordering varies.

Returning to our simple example, we take the five rules of Section 2 (reproduced here for convenience),

$$\frac{\Box}{0=0}\ \mathbf{Z} \qquad \frac{\boxed{i=j}}{i=j}\ \mathbf{I}_{i=j}$$

$$\frac{i=j}{si=sj}\ \mathbf{S} \qquad \frac{a\ \ c}{c}\ \mathbf{P} \qquad \frac{i=j\ \ j=k}{i=k}\ \mathbf{T}$$

extend $I$ and $T$ to disequalities, and add a third rule for disequalities as follows:

$$\frac{\boxed{i\neq j}}{i\neq j}\ \mathbf{I}_{i\neq j} \qquad \frac{i=j\ \ j\neq k}{i\neq k}\ \mathbf{T} \qquad \frac{i\neq i}{j=k}\ \ \mathbf{F}_{j=k}$$

With these rules, one can infer, for instance, $0\neq 0$ from $1\neq 1$ and $1\neq 0$, by applying $\mathbf{I}_{1\neq 1}$, $\mathbf{F}_{0=1}$, $\mathbf{I}_{1\neq 0}$ and $\mathbf{T}$:

$$\frac{\dfrac{\mathbf{I}_{1\neq 1}}{0=1}\ \ \ \mathbf{I}_{1\neq 0}}{0\neq 0}$$

Suppose we are using a proof ordering based on a precedence on the inference rules, or proof combinators, $Z, I, S, P, T, F$. For simplicity, we use $>$ for both proof ordering and precedence. The intended meaning will be clear from the context.

If $F$ is smaller than all other proof combinators in the precedence, and $I$ nodes are incomparable in the proof ordering, then the canonical basis of any inconsistent set is $\{i\neq j : i, j \in \mathbf{N}\}$. All positive equations are redundant, because $\mathbf{F}_{j=k}$ is a smaller proof than $\mathbf{I}_{j=k}$.

If $P > I$ in the precedence, then

$$\frac{a\ \ c}{c} > c$$

or $P(a,c) > I(c)$. By the Replacement Postulate (7), every application of $P$ can be replaced by an application of $I$ to yield a smaller proof. Hence, no minimal proof includes $\mathbf{P}$ steps.

If proofs are compared in a simplification ordering (that is, in an ordering for which subproofs are always smaller than their superproofs), then minimal proofs will never have superfluous transitivity inferences of the form

$$\frac{u=t\ \ t=t}{u=t}$$

because the trivial proof of $u=t$ (made of $u=t$ itself) is smaller.

More specifically, suppose we are using something like the recursive path ordering for proof terms and consider the above inference rules for ground equality and disequality, with the rule for successor extended to apply to all function symbols of any arity. That is, rule $\mathbf{S}$, which infers $si=sj$ from $i=j$, is generalized here to an inference rule for functional reflexivity, that infers $f(\bar{x})=f(\bar{y})$ from $\bar{x}=\bar{y}$, for any function symbol $f$, of any arity $n$, and $n$-tuples $\bar{x}$ and $\bar{y}$ of variables.

*Deductive closure.* If the proof ordering prefers introduction $\mathbf{I}$ of premises over all other inferences (including $\mathbf{Z}$), then trivial proofs are best. In this case, the whole

theory is irredundant ($Red\ Th\ A = \emptyset$); and the canonical basis includes the whole theory ($A^\sharp = Th\ A$). In other words, everything is needed, because each sentence constitutes the smallest proof of itself.

*Congruence closure.* If the precedence makes functional reflexivity $S$ smaller than $I$ (more precisely: $S < T < I$), but the only ordering on leaves is $I(u,t) \leq I(c[u], c[t])$ for any context $c$, then inferring $c[u] = c[t]$ from $u = t$ by repeated applications of $S$ yields a less expensive proof than $I(c[u], c[t])$. Ground paramodulation can deduce $c[u] = c[t]$ from $u = t$ and $c[u] = c[u]$ in one step. The canonical basis will be the congruence closure, as generated by paramodulation. Redundancies will have the form $f(u_1, \ldots, u_n) = f(t_1, \ldots, t_n)$ for all $u_1 = t_1, \ldots, u_n = t_n \in Th\ A$ and function symbol $f$ (of any arity $n$) in the vocabulary. Theory $Th\ A$ is the closure under functional reflexivity of the basis $A^\sharp$. If $A$ is as in our first example (i.e. $A = \{4 = 2, 4 = 0\}$), then $A^\sharp = \{2j = 0 \colon j > 0\}$. The other equalities in $Th\ A = \{i = j \colon i \equiv j \pmod 2\}$ are obtained from those in $A^\sharp$ by applying $S$ (e.g. $8 = 4$ is derived from $4 = 0$ by applying $S^4$ to both sides).

*Completion.* On the other hand, if the ordering on leaves compares terms in some simplification ordering $\geqslant$ (still assuming $S < T < I$), then the canonical basis will be the fully contracted set, as generated by (ground) completion. The redundancies will be the trivialities $u = u$, for all terms $u$, and equalities $u = t$, when there is a $t = v \in Th\ A$ ($v$ different than $u$), such that $t \gg v$. Operationally, $u = t$ can be contracted to $u = v$. For our first example, with $A = \{4 = 2, 4 = 0\}$, we have $A^\sharp = \{2 = 0\}$, as all equations in $\{2j = 0 \colon j > 0\}$ reduce to $2 = 0$. For another example, if $A = \{a = c, sa = b\}$ and $sa \gg sb \gg sc \gg a \gg b \gg c$, then $I(sa, b) > T(S(I(a,c)), I(sc, b))$, and $I(sc, b) < T(S(I(a,c)), I(sa, b))$, hence $A^\sharp = \{a = c, sc = b\}$.

*Refutation.* If $T < I$, the combinator $F$ is the smallest in the precedence and $I(i, j)$ nodes are measured by the values of $i$ and $j$, then the canonical basis of any inconsistent presentation is a (smallest) trivial disequation $\{t \neq t\}$. Indeed, all positive equations can be obtained by applying $F$ to $t \neq t$, and all negated equations can be obtained by two applications of $T$:

$$\frac{\dfrac{n = t \quad t \neq t}{n \neq t} \qquad t = m}{n \neq m}$$

for all numerals $m$, $n$ and $t$. Thus, the process of searching for a refutation of a given input set is the process of seeking its canonical basis, or forcing a minimal nucleus of inconsistency to emerge.

*Superposition.* In the ground case, completion can be done by simplification only. However, with a suitable ordering, one can observe also superposition. If one distinguishes $T$ steps based on the weight of the shared term $j$, making $T > I$ when $j$ is the greatest, and $T < I$ otherwise, then the canonical basis is also closed under superposition, or paramodulation into the larger side of equations. For example, consider $k = j$ and $j = i$. If the shared term $j$ is the greatest, we have $T(I(k, j), I(j, i)) > I(k, i)$, meaning that adding $k = i$ by superposition provides a smaller proof. The transitivity proof $T(I(k, j), I(j, i))$ corresponds to the peak

$k \leftarrow j \rightarrow i$. Otherwise, we have $T(I(k,j), I(j,i)) < I(k,i)$. In particular, if the shared term $j$ is the smallest, the transitivity proof $T(I(k,j), I(j,i))$ corresponds to a valley $k \rightarrow j \leftarrow i$, and $T(I(k,j), I(j,i)) < I(k,i)$ means that valley proofs are the smallest.

## 5.  INFERENCE AND DERIVATIONS

There are two basic applications for ordering-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 the search.

Inference steps are defined by deduction mechanisms. In general, a *(one-step) deduction mechanism* $\rightsquigarrow$ is a binary relation over presentations, and we call a pair $A \rightsquigarrow B$, a *deduction step*. A deduction mechanism is *functional* if for any $A$ there is a unique $B$ (possibly $A$ itself) such that $A \rightsquigarrow B$. Practical mechanisms are functional (and usually operate deterministically); they are obtained by coupling an (nondeterministic) inference system with a *search plan* (or *strategy*), to yield a *completion procedure* or *proof procedure*. Specific procedures may impose additional structure, such as singling out one formula as the *target* theorem or "goal," in which case the deduction mechanism applies to labelled formulæ; see [Bonacina 1999] for a survey.

Here, we consider only functional mechanisms that apply to presentations, and take the notion of a deduction mechanism as a whole. Focusing attention on deduction mechanisms that apply to presentations entails no loss of generality, since the abstract set $\mathbb{P}$ may be limited on the concrete level to proofs and subproofs of a specific goal.

### 5.1  Goodness

A sequence of deductions $A_0 \rightsquigarrow A_1 \rightsquigarrow \cdots$ is called a *derivation*.[5] We write $\{A_i\}_i$ for sequences of presentations, and—in particular—for derivations. Let $A_* = \cup_i A_i$ be all formulæ appearing anywhere in $\{A_i\}_i$. The *result* $A_\infty$ of the sequence is—ever since Huet [1981]—its *persisting* formulæ:

$$A_\infty \quad \overset{!}{=} \quad \liminf_{j \to \infty} A_j \;=\; \bigcup_j \bigcap_{i \geq j} A_i$$

We say that a proof $p$ *persists* when its premises do, that is, when $[p]^{Pm} \subseteq A_\infty$. Thus, if $p$ persists, so do its subproofs, by Postulate (6). By Proposition 3.6(b), we have $Pf(A_i) \sqsupseteq Pf(A_*)$ for all $i$.

*Definition* 5.1 *(Soundness and Adequacy).*

— A deduction step $A \rightsquigarrow B$ is *sound* if $B \subseteq Th\, A$.
— It is *adequate* if $A \subseteq Th\, B$.
— It is both if $A \equiv B$.
— A derivation $\{A_i\}_i$ is *sound* if $A_\infty \subseteq Th\, A_i$, for all $i$.
— It is *adequate* if $A_i \subseteq Th\, A_\infty$.
— It is both if $A_i \equiv A_\infty$.

---

[5]We do not consider transfinite derivations in this paper.

Adequacy is essentially a monotonicity property, since it implies that $Th\,A \subseteq Th\,B$ whenever $A \rightsquigarrow B$.

We will concern ourselves only with sound and adequate derivations. In addition, we want derivations to improve gradually the presentation.

*Definition* 5.2 *(Goodness).*

— A deduction step $A \rightsquigarrow B$ is *good* if $A \succsim B$.

— A sequence $\{A_i\}_i$ is *good* if $A_i \succsim A_{i+1}$ for all $i$.

— A deduction mechanism $\rightsquigarrow$ is *good* if proofs only get better, in the sense that $A \succsim B$ whenever $A \rightsquigarrow B$.

Goodness is the cardinal principle of canonical inference. *From here on in, only good, sound, adequate derivations will be considered.*

Since the proof ordering is well-founded, we get:

LEMMA 5.3. *For each presentation $A_i$ in a good derivation $\{A_i\}_i$, we have:*

$$Pf(A_i) \sqsupseteq Pf(A_\infty)$$
$$Th\,A_i \subseteq Th\,A_\infty$$

Let $\{A \vdash c\} \stackrel{!}{=} \{p \in Pf(A) : [p]_{Cl} = c\}$ signify the proofs of formula $c$ from any subset of presentation $A$.

PROOF. Let $p_i \in \{A_i \vdash c\}$. Since the derivation is good, there are proofs $p_j \in \{A_j \vdash c\}$, for $j > i$, such that $p_i \geq p_{i+1} \geq \cdots$. By well-foundedness, from some point on these are all the same proof $q$. Thus, $[q]^{Pm} \subseteq A_\infty$, $q \in Pf(A_\infty)$ and $Pf(A_i) \sqsupseteq Pf(A_\infty)$. That $Th\,A_i \subseteq Th\,A_\infty$ follows then from the definitions.    □

NOTE 5.4. *For bad (i.e. non-good) derivations this is not the case. To wit, let*

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

*and consider $\{c\} \rightsquigarrow \{b\} \rightsquigarrow \{c\} \rightsquigarrow \{b\} \rightsquigarrow \cdots$. As the derivation oscillates perpetually between deriving $b$ from $c$ and $c$ from $b$, at the limit $A_\infty = \emptyset$ and $Th\,A_\infty = \emptyset$, whereas $Th\,A_i = \{b, c\}$ for all finite $i$.*

## 5.2 Canonicity

Canonicity of presentations leads to canonicity of derivations, in the sense that a derivation deserves to be considered canonical if it generates a canonical limit. More generally, a desirable attribute of presentations induces a corresponding characteristic of derivations that is sufficient to guarantee that the limit has the desirable attribute. The first ingredient for canonicity of derivations is the property that once something becomes redundant during a derivation, it will remain such forever, or *"once redundant, always redundant."* The following lemma implies that good derivations have this feature:

LEMMA 5.5. *For all presentations $A$ and $B$:*

$$Pf(A) \sqsupseteq Pf(B) \;\Rightarrow\; B \cap Red\,A \subseteq Red\,B$$

PROOF. Consider a proof $p \in Pf(B)$ that uses a redundant premise $a \in B \cap Red\, A \subseteq A$. Since $\widehat{a} \in Pf(A)$, by assumptions (1,2), $a$ must also have an alternative (nontrivial) proof $q \in \{A \setminus \{a\} \vdash a\}$, such that $\widehat{a} > q$. By assumption, there is an $r \in Pf(B)$ such that $q \geq r$. By the postulates of subproofs, $p \trianglerighteq \widehat{a} > r$ implies the existence of a proof $p' \in Pf(B \cup \{a\}) = Pf(B)$ such that $p > p'$. If $a \in [p']^{Pm}$, then this process continues. It cannot continue forever, so we end up with a strictly smaller proof not involving $a$, establishing $a$'s redundancy vis-à-vis $B$. □

PROPOSITION 5.6. *If a derivation $\{A_i\}_i$ is good, then its limit supports the best proofs:*

$$A_* \approx A_\infty$$

PROOF. One direction, namely $Pf(A_\infty) \sqsupseteq Pf(A_*)$, follows by Proposition 3.6(b) from the fact that $A_\infty \subseteq A_*$. To establish that $Pf(A_*) \sqsupseteq Pf(A_\infty)$, we show that $\mu Pf(A_*) \sqsupseteq Pf(A_\infty)$ and rely on Proposition 3.6(c). Suppose $p \in \mu Pf(A_*)$. It follows from Eq. (5) and Proposition 3.6(a) that $\widehat{[p]^{Pm}} \subseteq \mu Pf(A_*)$. By goodness, each $a \in [p]^{Pm}$ persists from some $A_i$ on. Hence, $[p]^{Pm} \subseteq A_\infty$ and $p \in Pf(A_\infty)$. □

*Definition* 5.7 *(Canonical Derivations).*

— A derivation $\{A_i\}_i$ is *completing* if its limit is complete.
— It is *saturating* if its limit is saturated.
— It is *contracting* if its limit is contracted.
— It is *canonical* if it is both saturating and contracting.

LEMMA 5.8.

*(a) A good derivation $\{A_i\}_i$ is completing if and only if every theorem of $A_0$ eventually admits a persistent normal-form proof:*

$$Th\, A_0 \subseteq [Pf(A_\infty) \cap Nf(A_0)]_{Cl}$$

*(b) It is saturating if and only if all normal-form proofs emerge eventually:*

$$Nf(A_0) \subseteq Pf(A_\infty)$$

*(c) It is contracting if and only if no formula remains persistently redundant:*

$$Red\, A_* \cap A_\infty = \emptyset$$

PROOF. Completeness of the limit is $Th\, A_\infty = [Pf(A_\infty) \cap Nf(A_\infty)]_{Cl}$. By Lemma 5.14, we know that $A_\infty \equiv A_0$ ($Th\, A_0 = Th\, A_\infty$) for all derivations of concern to us. Therefore, $[Pf(A_\infty) \cap Nf(A_\infty)]_{Cl} = [Pf(A_\infty) \cap Nf(A_0)]_{Cl} \subseteq [Pf(A_\infty)]_{Cl} = Th\, A_\infty = Th\, A_0$. With the preceding condition, we get $Th\, A_\infty = [Pf(A_\infty) \cap Nf(A_\infty)]_{Cl}$, as desired. The "only-if" direction is straightforward.

Similarly, by Lemma 3.10, the condition $Nf(A_0) \subseteq Pf(A_\infty)$ gives saturation.

By Proposition 5.6, $A_* \approx A_\infty$ and $Pf(A_*) \simeq Pf(A_\infty)$. By applying Lemma 5.5 to $Pf(A_*) \sqsupseteq Pf(A_\infty)$, one gets $Red\, A_* \cap A_\infty \subseteq Red\, A_\infty$. If the limit is contracted, $Red\, A_\infty = \emptyset$, so that we have $Red\, A_* \cap A_\infty \subseteq Red\, A_\infty = \emptyset$. For the "if" direction, by applying Lemma 5.5 to $Pf(A_\infty) \sqsupseteq Pf(A_*)$, one gets $Red\, A_\infty \cap A_* \subseteq Red\, A_*$. Since $Red\, A_\infty \subseteq A_\infty \subseteq A_*$, we have $Red\, A_\infty = Red\, A_\infty \cap A_* \subseteq Red\, A_*$. So, if the condition $Red\, A_* \cap A_\infty = \emptyset$ holds, then $Red\, A_\infty = Red\, A_\infty \cap A_\infty \subseteq Red\, A_* \cap A_\infty = \emptyset$, and $A_\infty$ is fully contracted. □

LEMMA 5.9. *A sufficient condition for a good derivation $\{A_i\}_i$ to be completing is that each non-normal-form proof eventually becomes much better:*

$$\bigcup_i \mu Pf(A_i) \setminus Nf(A_0) \sqsupset \bigcup_i Pf(A_i)$$

PROOF. By Lemma 5.3, if $p_i \in \mu\{A_i \vdash c\}$ then $q \in \{A_\infty \vdash c\}$, for some $q$. If $q \in Nf(A_0)$, then $c \in [Pf(A_\infty) \cap Nf(A_0)]_{Cl}$ and we are done. Otherwise, the sufficient condition implies that, for some $k$, there is a proof $q_k \in Pf(A_k)$ of $c$ such that $p_i \geq q > q_k$. Completeness follows by induction on proofs. $\square$

LEMMA 5.10. *A good derivation $\{A_i\}_i$ is canonical if and only if*

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

PROOF. Assume the derivation is canonical, that is, saturating and contracting. Saturating means $Nf(A_0) \subseteq Pf(A_\infty)$, hence $[Nf(A_0)]^{Pm} = A_0^\sharp \subseteq A_\infty$. Contracting means $Red\, A_\infty = \emptyset$, from which it follows that $A_\infty \subseteq A_0^\sharp$ (by way of contradiction, if there were an $x \in A_\infty$, but $x \notin A_0^\sharp$, this $x$ would be redundant, contradicting the contracting hypothesis). Together, these conclusions give $A_0^\sharp = A_\infty$. The other direction is trivial. $\square$

In summary, the limit of a derivation is *complete, contracted, saturated*, if the derivation is *completing, contracting, saturating*, respectively, where saturated is stronger than complete, and saturated and contracted together mean *canonical*.

## 5.3  Compactness

Goodness implies that if any proof shows up during a derivation, then there is a better or equal proof in the limit (cf. Lemma 5.3). The converse property, namely that if there is a proof in the limit, then there must also have been a proof along the way, is ensured by *continuity*:

*Definition* 5.11 *(Continuity).* (Minimal) Proofs are *continuous* if

$$\liminf_{i \to \infty} \mu Pf(A_i) \;=\; \mu Pf(A_\infty) \;\left(= \mu Pf(\liminf_{i \to \infty} A_i)\right)$$

for any good sequence $A_0 \succsim A_1 \succsim \cdots$.

In other words, the operator $\mu Pf$ is continuous for any chain: the limit of the chain of the images is equal to the image of the limit of the chain.

In turn, for continuity suffices that minimal proofs use only a finite number of premises. We call this property *compactness* (of proofs), because it is used traditionally to infer compactness of a logic (namely, that a set of formulæ is unsatisfiable if and only if it has a finite unsatisfiable subset) from its completeness (viz. a presentation is unsatisfiable if and only if it is inconsistent).[6]

---

[6]Indeed, if a set $A$ is unsatisfiable, there is a proof of $F$ (falsehood) in $Pf(A)$ (unsatisfiable implies inconsistent). Take a minimal proof $p \in \mu Pf(A)$ of $F$, and let $A'$ be the finite set $[p]^{Pm}$; since $p \in Pf(A')$, $A'$ is unsatisfiable (inconsistent implies unsatisfiable), and is a finite subset of $A$.

*Definition* 5.12 *(Compactness).*  An ordered proof system is *compact* if minimal proofs use only a finite number of premises:

$$\forall p \in \mu Pf(\mathbb{A}). \quad \left| [p]^{Pm} \right| < \infty$$

For ordinary inference systems, even non-minimal proofs are finitely based.

LEMMA 5.13. *Compactness implies continuity.*

PROOF.  Continuity requires $\bigcup_j \bigcap_{i \geq j} \mu Pf(A_i) = \mu Pf(\bigcup_j \bigcap_{i \geq j} A_i)$ for good sequences.

To show $\mu Pf(\cup_j \cap_{i \geq j} A_i) \subseteq \cup_j \cap_{i \geq j} \mu Pf(A_i)$: Let $p \in \mu Pf(\cup_j \cap_{i \geq j} A_i) = \mu Pf(A_\infty)$. By compactness, there are only finitely many $a \in [p]^{Pm}$. Let $j$ be the smallest index in the derivation such that all $a \in [p]^{Pm}$ are in $A_j$. Then $p \in Pf(A_j)$. Second, $p \in \mu Pf(A_j)$, because $p \in \mu Pf(A_\infty)$, and (by the previous lemma) $A_j$ cannot provide a strictly better proof. Third, $p \in \cap_{i \geq j} \mu Pf(A_i)$, because all $a \in [p]^{Pm}$ persist, since $[p]^{Pm} \subseteq A_\infty$. It follows that $p \in \cup_j \cap_{i \geq j} \mu Pf(A_i)$.

For $\cup_j \cap_{i \geq j} \mu Pf(A_i) \subseteq \mu Pf(\cup_j \cap_{i \geq j} A_i)$: Let $p \in \cap_{i \geq j} \mu Pf(A_i)$ for some $j$. It follows that for every premise $a \in [p]^{Pm}$, $a \in \cap_{i \geq j} A_i$, whence $a \in \cup_j \cap_{i \geq j} A_i = A_\infty$. This means that $p \in Pf(A_\infty)$. As above, were $p$ not minimal, on account of compactness and goodness, it would have already turned non-minimal at some stage $k$. But $p$ is minimal at all stages $i \geq j$, so $p \in \mu Pf(A_\infty)$.  □

LEMMA 5.14. *If proofs are continuous, then any good derivation $\{A_i\}_i$ is sound and adequate. That is, for all $i$,*

$$A_i \quad \equiv \quad A_\infty$$

PROOF.  Lemma 5.3 gives adequacy, regardless of continuity:  $Th\, A_i \subseteq Th\, A_\infty$. Suppose, now, that $c \in Th\, A_\infty$, with proof $p \in \mu Pf(A_\infty)$. By continuity, $p \in \cap_{i \geq j} \mu Pf(A_i)$ for some $j$. Thus, $c \in Th\, A_i$ for all $i \geq j$. That $c \in Th\, A_i$ also for $i < j$ follows from goodness, since $A_i \succsim A_j$ implies $A_i \equiv A_j$ (see Definition 3.7).  □

NOTE 5.15. *This does not necessarily hold for infinitary systems that violate the compactness hypothesis. Let all proofs be incomparable, including (for all $i$ and $j$):*

$$\widehat{a}_i \quad \frac{a_j}{a_i} \quad \frac{}{c} \quad \frac{a_0, a_1, \ldots}{c}$$

*The derivation $\{a_j : j \leq i\}_i$ is good, but only its limit includes the infinitary proof.*

## 6.  COMPLETION PROCEDURES AND PROOF PROCEDURES

The central concept underlying completion is the existence of critical proofs. Completion alternates "expansions" that infer the conclusions of critical proofs with "contractions" that remove redundancies. More generally, theorem proving with simplification (e.g. Dershowitz 1991b; Bonacina and Hsiang 1995; Bachmair and Ganzinger 1994) entails two processes: **Expansion**, whereby any sound deductions (anything in *Th A*) may be added to the set of derived theorems; and **Contraction**, whereby any redundancies (anything in *Red A*) may be removed. This inference-rule interpretation of completion, accommodating both expansion and contraction, was elaborated on in [Bachmair and Dershowitz 1994].

*Definition* 6.1 *(Expansion and Contraction).*

— A deduction step $A \rightsquigarrow A \cup B$ is an *expansion* provided $B \subseteq Th\,A$.

— A deduction step $A \cup B \rightsquigarrow A$ is a *contraction* provided $A \cup B \succsim A$.

It is easy to see that:

PROPOSITION 6.2.

*(a) Expansions and contractions are good.*

*(b) Derivations, whose steps are expansions or contractions, are good.*

*Definition* 6.3 *(Criticality).*

— A minimal proof $p \in \mu Pf(A)$ is *critical* if it is not in normal form, but all its proper subproofs are:

$$p \in \mu Pf(A) \setminus Nf(A)$$
$$\forall q.\ p \rhd q \ \Rightarrow \ q \in Nf(A)$$

— We use $C(A)$ to denote the set of all such critical proofs in $A$.

— The *critical theorems* of a presentation $A$ are the conclusions of its critical proofs, or $[C(A)]_{Cl}$.

— A formula is *critical* for $A$ if it is a premise of a proof smaller than a critical proof in $C(A)$.

LEMMA 6.4. *The canonical presentation has neither critical formulæ nor critical theorems.*

PROOF. By the definition of critical proof, $C(A^\sharp) \subseteq \mu Pf(A^\sharp) \setminus Nf(A^\sharp)$. Since $\mu Pf(A^\sharp) \setminus Nf(A^\sharp) = \emptyset$, by the definition of $Nf$, it follows that $C(A^\sharp) = \emptyset$, and $A^\sharp$ has no critical theorems or critical formulæ. □

Since [Huet 1981], fairness has been seen as the fundamental requirement of derivations generated by completion procedures. Here, we define two fairness properties, one each for complete or saturated limits:

*Definition* 6.5 *(Fairness).*

— A good derivation $\{A_i\}_i$ is *fair* if

$$C(A_\infty) \ \sqsupset \ Pf(A_*)$$

— It is *uniformly fair* if

$$\widehat{A_\infty} \setminus \widehat{A^\sharp} \ \sqsupset \ Pf(A_*)$$

Fairness means that all critical proofs with persistent premises are "subsumed" eventually by strictly smaller proofs, whereas uniform fairness predicates the same for trivial proofs with persistent premises.

THEOREM 6.6. *Presentation $A$ is complete if $C(A) \sqsupset Pf(A)$.*

PROOF. Assume, by way of contradiction, that $A$ is incomplete, in other words, that $[Pf(A) \cap Nf(A)]_{Cl} \subsetneq Th\,A$. Then there is a $c \in Th\,A$ such that $c \notin [Pf(A) \cap Nf(A)]_{Cl}$, or there is no proof of $c$ in $Pf(A) \cap Nf(A)$. However, there are proofs of $c$ in $Pf(A)$: let's take a minimal one, that is, let $p \in \mu\{A \vdash c\}$. By the above, $p \notin Nf(A)$. If $p$ is not in normal form, it means that it has some subproof(s) that is not in normal form, that is, some $q \trianglelefteq p$ that is not in normal form. By the well-foundedness of $\trianglelefteq$, let $q$ be a minimal (with respect to $\trianglelefteq$) such proof. Minimality with respect to $\trianglelefteq$ means that all subproofs of $q$ are in normal form. Thus, we have a (possibly trivial) subproof $q$ of $p$, which is not in normal form, but such that all its subproofs are. But this is the definition of critical proof: $q \in C(A)$. The hypothesis $C(A) \sqsupset Pf(A)$ implies that there exists a proof $r \in Pf(A)$ such that $r < q$. Since we have $p \trianglerighteq q > r$, by Replacement (8), there exists a $p' \in Pf(A)$, such that $p' < p$, with $r$ in place of $q$, that is, $p > p' \trianglerighteq r$. This contradicts the fact that $p$ is minimal. $\square$

COROLLARY 6.7. *If a good derivation is fair, then its limit is complete.*

PROOF. By the definition of fairness we have $C(A_\infty) \sqsupset Pf(A_*)$. By Proposition 5.6, $Pf(A_*) \simeq Pf(A_\infty)$, so that $C(A_\infty) \sqsupset Pf(A_\infty)$. By Theorem 6.6, $A_\infty$ is complete. $\square$

This suggests completing an axiomatization $A_0$ by adding, step by step, what is needed to make for better proofs than the critical ones.

For example, suppose a proof ordering makes $\widehat{c} > \frac{b}{c}$ and $\frac{c}{b} > \widehat{b}$. Start with $A_0 = \{c\}$ and consider $\widehat{c}$. Were $\widehat{c}$ to persist, then by fairness a better proof would evolve, the better proof being $\frac{b}{c}$. If $\widehat{b}$ is in normal form, then $b \in A_\infty$ and both minimal proofs $\frac{b}{c}$ and $\widehat{b}$ persist.

Another example: $\mu\mathbb{P} = \{\widehat{b}, \widehat{c}, \frac{b}{c}\}$ and $A = \{b\}$, then $A \rightsquigarrow A \rightsquigarrow \cdots$ is fair, since $A_\infty = A$ and $C(A_\infty) = \emptyset$. The result is complete but unsaturated ($c$ is missing).

Clearly, a fair derivation is also completing. On the other hand, completing does not imply fair, because the limit could feature a normal-form proof of some $c \in Th\,A_0$, without having reduced all persistent critical proofs of $c$. The two notions serve different purposes: completing is the more abstract and represents the condition for attaining a complete limit. Fair is stronger and more concrete, as it specifies a way to achieve completeness by reducing all persistent critical proofs.

A saturated limit is not necessarily contracted, unless the derivation is contracting, in which case it is canonical:

THEOREM 6.8 (FAIR COMPLETION). *Contracting, fair derivations are canonical, provided minimal proofs are unique.*

PROOF. This follows from Lemma 5.8(c) (contracting derivation implies contracted limit), Corollary 6.7 (fair derivation implies complete limit), Proposition 3.11 (saturated and complete are equivalent if minimal proofs are unique), and Theorem 3.17 (saturated and contracted imply canonical). $\square$

By Proposition 3.3, this also means that each $a \in A_\infty$ ($= A^\sharp$) is its own ultimate proof $\widehat{a} \in Nf(A)$, so is not susceptible to contraction.

We are left with the task of identifying sufficient conditions for saturation, in case minimal proofs are not unique:

THEOREM 6.9. *Presentation $A$ is saturated if and only if $\widehat{A} \setminus \widehat{A^\sharp} \sqsupset Pf(A)$.*

PROOF. Recall that $A$ saturated means $\mu Pf(A) = Nf(A)$.

First, we show that $\widehat{A} \setminus \widehat{A^\sharp} \sqsupset Pf(A)$ implies saturation, assuming, by way of contradiction, that $\mu Pf(A) \neq Nf(A)$. Then, there is a theorem $c \in Th\,A$ for which a normal-form proof $p^*$ is absent from $\mu Pf(A)$. Instead, there is a minimal non-normalized proof $p \in \mu Pf(A) \setminus Nf(A)$. So, there is some $x \in [p]^{Pm} \setminus A^\sharp$, since $p$ would be in normal form were $[p]^{Pm} \subseteq A^\sharp$. By hypothesis, $\widehat{x} > r$ for some $r \in Pf(A)$. By Replacement (8), there exists a $v \in Pf(A)$, such that $p > v \unrhd r$, contradicting the minimality of $p$.

For the other direction, suppose $\mu Pf(A) = Nf(A)$. Employing Proposition 3.6(c), we have $\widehat{A} \setminus \widehat{A^\sharp} \sqsupseteq Pf(A) \sqsupseteq \mu Pf(A) = Nf(A)$. But if $x \in A \setminus A^\sharp$, then $\widehat{x} \notin Pf(A^\sharp) \supseteq Nf(A)$ (the inclusion is from (9)), so there must be some other, strictly smaller proof than $\widehat{x}$ in $Nf(A)$. So, in fact, $\widehat{A} \setminus \widehat{A^\sharp} \sqsupset Nf(A) = \mu Pf(A) \sqsupseteq Pf(A)$, as desired. $\square$

By the previous theorem, if $A$ is saturated, $A \setminus A^\sharp$ is redundant (i.e. $A \setminus A^\sharp = Red\,A$).

COROLLARY 6.10. *A good derivation is uniformly fair if and only if its limit is saturated.*

PROOF. Uniform fairness says that $\widehat{A_\infty} \setminus \widehat{A^\sharp} \sqsupset Pf(A_*)$. Since $Pf(A_*) \simeq Pf(A_\infty)$ by Proposition 5.6, this is equivalent to $\widehat{A_\infty} \setminus \widehat{A^\sharp} \sqsupset Pf(A_\infty)$, which is equivalent to $A_\infty$ being saturated by Theorem 6.9. $\square$

## 7.  INSTANCES OF THE FRAMEWORK

A class of completion procedures can be described as deduction mechanisms, wherein each step $A_i \rightsquigarrow A_{i+1}$ is the composition of an expansion that adds some formulæ, followed by a contraction that removes all redundant formulæ (cf. Dershowitz 1985, Sect. 3.1). In other words, we are looking at deductions of the form $A \rightsquigarrow (A \cup D)^\flat$, where $D$ is the expansion and $B^\flat \stackrel{!}{=} B \setminus Red\,B$ is $B = A \cup D$ after contraction.

One possibility for such a mechanism is to expand with all critical theorems:

*Definition* 7.1 *(Critical Completion).  Critical completion* is a sequence of steps:

$$\textbf{Critical}:\quad A \underset{c}{\rightsquigarrow} \left(A \cup [C(A)]_{Cl}\right)^\flat$$

An alternative is to add only something better:

*Definition* 7.2 *(Bulk Completion).  Bulk completion* is a sequence of steps:

$$\textbf{Bulk}:\quad A \underset{b}{\rightsquigarrow} \left(A \cup [B(A)]^{Pm}\right)^\flat$$

where $B(A)$ is a minimal subset of $Pf(A)$ (minimal, with respect to $\subseteq$) that is much better than critical proofs: $C(A) \sqsupset B(A)$.

Another variation on this theme is "mass completion," where the expansion component of each step $A_i \leadsto_m A_{i+1}$ adds normal-form trivial theorems, *en masse*, followed by contraction:

*Definition* 7.3 *(Mass Completion).* *Mass completion* is a sequence of steps:

$$\boxed{\textbf{Mass}: \quad A \underset{m}{\leadsto} (A \cup [M(A)]_{Cl})^{\flat}}$$

where

$$M(A) \overset{!}{=} \{p \in \mu Pf(A) : \widehat{p} < p \wedge \forall q \lhd p. \widehat{q} \not< q\}$$

and $\widehat{p}$ is short for $\widehat{[p]_{Cl}}$, the *trivial* proof of the conclusion of $p$.

By Proposition 6.2:

THEOREM 7.4. *Critical completion, bulk completion and mass completion are all good.*

A presentation $A$ is *stable* under a deduction mechanism $\leadsto$ if $B = A$ whenever $A \leadsto B$.

THEOREM 7.5. *The canonical presentation is stable under critical, bulk and mass completion.*

PROOF. By the proof of Lemma 6.4, $\mu Pf(A^{\sharp}) \setminus Nf(A^{\sharp}) = \emptyset$ and $C(A^{\sharp}) = \emptyset$. It follows that $[C(A^{\sharp})]_{Cl} = \emptyset$. Second, the condition $C(A^{\sharp}) \sqsupset B(A^{\sharp})$ is satisfied vacuously and the minimal subset of $Pf(A^{\sharp})$ is $\emptyset$, so that $B(A^{\sharp}) = \emptyset$ and $[B(A)]^{Pm} = \emptyset$. Third, since there are no better proofs than those provided by $A^{\sharp}$ (Theorem 3.8), $M(A^{\sharp}) = \emptyset$ and $[M(A)]_{Cl} = \emptyset$. Hence, expansions by critical, bulk and mass completion do not apply. Because $A^{\sharp}$ is contracted (by Theorem 3.17), we have $Red\, A^{\sharp} = \emptyset$, and contraction does not apply either. So, for all three mechanisms, $A^{\sharp} \leadsto A^{\sharp}$ only. □

Let $A_{\infty}^{\text{Bulk}}$ and $A_{\infty}^{\text{Mass}}$ denote the limits of derivations by bulk and mass completion from $A$, respectively. Similarly, let $A_{*}^{\text{Bulk}}$ and $A_{*}^{\text{Mass}}$ denote the sets of *all* derived formulæ in those derivations.

THEOREM 7.6. *Bulk completion is canonical, provided proofs are continuous and minimal proofs are unique, in which case*

$$A_{\infty}^{\text{Bulk}} = A^{\sharp}$$

PROOF. Let $\{A_i\}_i$ be a derivation by bulk completion starting from $A = A_0$. By Theorem 6.8, canonicity of the limit requires that derivations by bulk completion be fair and contracting. Fairness says that

$$\forall p \in C(A_{\infty}^{\text{Bulk}}). \, \exists q \in Pf(A_{*}^{\text{Bulk}}). \, p > q$$

Let $p$ be a proof in $C(A_{\infty}^{\text{Bulk}})$ and let $i$ be the smallest index such that $p \in C(A_i)$. There must be such an $i$ by continuity (Definition 5.11), given goodness—per Theorem 7.4. By the definition of bulk completion and the nature of expansion and redundancy removal (Propositions 3.6(b) and 3.16), $C(A_i) \sqsupset B(A_i) \sqsupseteq$

$Pf((A_i \cup [B(A_i)]^{Pm})^{\flat}) = Pf(A_{i+1})$. It follows that there is some $q \in Pf(A_{i+1}) \subseteq Pf(A_*^{\mathrm{Bulk}})$, such that $q < p$, establishing fairness. As bulk completion removes redundancies immediately, its derivations are also contracting; see Lemma 5.8(c).    □

THEOREM 7.7. *Mass completion is canonical, provided proofs are continuous and minimal proofs are unique, in which case*

$$A_{\infty}^{\mathrm{Mass}} = A^{\sharp}$$

PROOF. For mass completion, it is convenient to show that the limit is saturated in terms of the characterization of $A^{\sharp}$ as all trivial normal-form theorems (Proposition 3.3). Suppose $c \in A^{\sharp}$ and $\widehat{c}$ is in normal form, and let $p \in \mu Pf(A_{\infty}^{\mathrm{Mass}})$ be a minimal proof of $c$ in the limit, which exists by virtue of Theorem 7.4 and Lemma 5.3. Since minimal proofs are unique, $\widehat{c}$ and $p$ are comparable. Suppose that $\widehat{c} < p$. Let $q$ be the smallest subproof of $p$ such that $q > \widehat{q}$, and let $i$ be the smallest index (as in the previous proof) such that $q \in \mu Pf(A_i)$. Thus, $q \in M(A_i)$, and, by the definition of mass completion, $q$ and $p$ (by (7)) have better proofs in $A_{i+1}$, and hence (by goodness and Lemma 5.3) in $A_{\infty}^{\mathrm{Mass}}$, contradicting the minimality of $p$. So $\widehat{c} = p$, and $c \in A_{\infty}^{\mathrm{Mass}}$, as desired. Hence, $A_{\infty}^{\mathrm{Mass}}$ is saturated. But $A_{\infty}^{\mathrm{Mass}}$ is also contracted, so, by Theorem 6.8, mass completion is canonical.    □

In the equational case, persistent critical pairs are at one and the same time both critical formulæ and critical theorems, since the proof ordering is designed so that the trivial proof using a critical pair is always smaller than the peak from which the critical pair is derived. So, expansions by $C(A)$, $B(A)$ and $M(A)$ are essentially the same, and bulk, mass and critical completion lead to the same result. In general, the different methods of expansion differ, as the following example demonstrates:

Suppose formula $a$ has three proofs: $\widehat{a}$, $p = \frac{b}{a}$, and $q = \frac{c}{a}$, and assume a proof ordering that orders proofs of $a$ by $\widehat{a} > p > q$, proofs of $c$ by $\frac{a}{c} > \frac{p}{c} > \frac{q}{c} > \widehat{c}$, while $\widehat{b}$ is the only proof of $b$. The only critical proof using $A = \{b\}$ is $\frac{b}{a}$: it is minimal in $Pf(A)$, it is not in normal form, and its only subproof $\widehat{b}$ is in normal form. Note that $\frac{p}{c}$ is not critical, although it is minimal and not in normal form, because its subproof $p$ is not in normal form. Critical completion generates the critical theorem $a$ and then deletes it right away, because $a$ is redundant, since $\widehat{a} > p$. Thus, derivation by critical completion is unfair, because a proof smaller than $p$ never arises. The limit of the derivation by critical completion is $\{b\}$ itself, which is not canonical, since it provides no normal form proofs for either $a$ or $c$.

On the other hand, bulk completion generates the critical formula $c$, premise of $\frac{c}{a} < \frac{b}{a}$. Similarly, mass completion generates $c$, because $M(A) = \{\frac{p}{c}\}$, since $\frac{p}{c}$ is the minimal proof of $c$ in $A$, $\widehat{c} < \frac{p}{c}$, and its only subproof $p$ does not share this property, as $\widehat{a} > p$. By adding $c$, the critical proof $p$ is replaced by $q$. The critical formula $c$ is not redundant and persists. Thus, the derivation is fair, and its limit $\{b, c\}$ is canonical, with normal form proofs $\widehat{b}$, $\widehat{c}$ and $\frac{c}{a}$. The behavior of critical completion, on one hand, and bulk or mass completion, on the other, would be the same, under a non-total proof ordering defined as the one above, except with proofs of $c$ ordered by $\frac{p}{c} > \frac{a}{c} > \widehat{c}$, $\frac{q}{c} > \frac{a}{c} > \widehat{c}$, where $\frac{p}{c}$ and $\frac{q}{c}$ are incomparable.

A subtle point is that bulk completion does not add all critical formulæ, but only sufficiently many to provide a smaller proof for each critical proof. (This is the gist

of the $C(A) \sqsupset B(A)$ condition in the definition of bulk completion.) To appreciate
the difference, consider a proof ordering such that $\widehat{c} > \frac{a_i}{c} > \frac{b}{c}$, for $i \geq 0$, with all
the $\frac{a_i}{c}$ incomparable. If the definition of bulk completion required it to add all the
$a_i$'s, it could not be considered a "mechanical" process. On the other hand, the
definition of bulk completion makes it sufficient to add just one of the $a_i$'s.

Lastly, the hypothesis that minimal proofs are unique is actually needed. Indeed,
consider proofs $\widehat{a}$, $\frac{a}{b}$ and $\widehat{b}$ with an empty ordering and let $A = \{a\}$. The minimal
proofs in $A$ are $\widehat{a}$ and $\frac{a}{b}$. Since $\widehat{b} < \frac{a}{b}$ does not hold, $M(A)$ is empty and mass
completion does not generate $b$. Similarly, $C(A)$ is empty and bulk completion
cannot generate $b$ either.

Returning to the ground equational case, with inference rules $P$, $I$, $T$, $S$, $Z$,
where $S$ is the inference rule for functional reflexivity given in Section 4, let $\geqq$ be a
total simplification-ordering of terms, let $P > I > T > S > Z$ in the precedence, let
proofs be greater than terms, and compare proof trees in the corresponding total re-
cursive path simplification-ordering. *Ground completion* is an inference mechanism
consisting of the following inference rules:

---

> **Deduce:**  $E \cup \{w = t[u]\} \rightsquigarrow E \cup \{w = t[v]\}$   if $u = v \in E$
> $\hspace{10.5cm}$ and $u \gg v$
>
> **Delete:**  $E \cup \{t = t\} \rightsquigarrow E$

---

Operationally, completion implements these inferences "fairly": No persistently
enabled inference rule is ignored forever.

THEOREM 7.8 (COMPLETENESS OF COMPLETION).         *Ground completion
results—at the limit—in the canonical, Church-Rosser basis.*

PROOF. Ground completion is good, since **Deduce** and **Delete** do not increase
proofs ($\rightsquigarrow \,\subseteq\, \succsim$). In particular,

$$I(w, t[u]) > T(I(w, t[v]), S^n(I(u, v)))$$

if $u \gg v$, where $n$ is the number of applications of $S$ needed to build the context
$t$, since $t[u] \gg t[v]$ and $t[u] \geqq u \gg v$. Ground completion is fair and contracting.
For example, the critical obligation

$$\frac{w = t \quad t = v}{w = v} \ \ \mathbf{T}$$

when $t \gg w, v$, is resolved by **Deduce**. Also, since $T > S$, non-critical cases resolve
naturally:

$$\frac{\dfrac{w = t}{fw = ft} \quad \dfrac{t = v}{ft = fv}}{fw = fv} \quad > \quad \frac{\dfrac{w = t \quad t = v}{w = v}}{fw = fv}$$

or $T(S(I(w, t)), S(I(t, v))) > S(T(I(w, t), I(t, v)))$. Since the proof ordering is total,
minimal proofs are unique, and Theorem 6.8 applies.   □

## 8. DISCUSSION

Completion procedures have been studied intensively since their discovery and application to automated theorem proving by Knuth and Bendix [1970] and Buchberger [1985]. The fundamental rôle of proof orderings in automated deduction, and the interpretation of completion as nondeterministic application of inference rules, received systematic treatment in [Bachmair and Dershowitz 1994]. The completion principle can be applied in numerous situations [Dershowitz 1989; Bonacina and Hsiang 1995], including the following:

— equational rewriting [Peterson and Stickel 1981; Jouannaud and Kirchner 1986; Bachmair and Dershowitz 1989];

— Horn theories [Kounalis and Rusinowitch 1991; Dershowitz 1991a; 1991b];

— induction [Kapur and Musser 1987; Fribourg 1989; Bachmair and Dershowitz 1994];

— unification [Doggaz and Kirchner 1991]; and

— rewrite programs [Bonacina and Hsiang 1992; Dershowitz and Reddy 1993].

Our abstract framework can be applied to re-understand completion mechanisms in a fully uniform setting. Because we have been generic in our approach, the results here apply to any completion-based framework, including standard ones, like ground completion and congruence closure,[7] as illustrated herein, equational completion (see [Burel and Kirchner 2006]), or completion for unification, and also to derive new completion algorithms, such as for constraint solving.

In [Bachmair and Dershowitz 1994], a completion sequence is deemed fair if all persistent critical inferences are generated, and criteria are employed to eliminate redundant inferences from consideration. In [Nieuwenhuis and Rubio 2001, fn. 8], an inference sequence is held to be fair if all persistent inferences are either generated or become redundant. The approach of [Bonacina and Hsiang 1995] distinguishes between fairness requirements for proof search and for saturation. The notion of fairness was formulated in terms of proof reduction with respect to a proof ordering, and made relative to the target theorem, suggesting for the first time that fairness should earn one a property weaker than saturation. Specifically, a derivation was considered fair if whenever a minimal proof of the target theorem is reducible by inferences, it is reduced eventually; see [Bonacina 1992, Chap. 2]. The treatment of fairness propounded here combines all these ideas. Fairness—for us—means that all persistent critical proofs are reduced, but it only attains completeness, not saturation. As we have seen, a stronger version of fairness, namely *uniform fairness*, is needed for saturation when the proof ordering is partial.[8]

Furthermore, by putting the accent on proof search and proof reduction, the approach of [Bonacina and Hsiang 1995] leads to an appreciation of the rôle of

[7]That ground completion can be used to compute congruence closure has been known since [Lankford 1975]; using congruence closure to generate canonical rewrite systems from sets of ground equations has been investigated further in [Gallier et al. 1993; Plaisted and Sattler-Klein 1996], among others; a recent survey comparing different ground completion and congruence closure algorithms can be found in [Bachmair et al. 2003].
[8]The term "uniform fairness" was introduced in [Bonacina 1992] for that property which guarantees saturation.

contraction as productive inference, as opposed to pure deletion. This is reflected here in the emphasis on canonicity, rather than saturation alone.

Bulk completion, as investigated here, is an abstract notion. Concrete procedures are obtained by coupling the inference system with a search plan that determines the order in which expansion and contraction steps take place. From a practical point of view, *fair* and *contracting* are two requirements for the search plan: it should schedule enough expansion steps to be fair, hence complete, and enough contraction steps to be contracting. Specific search plans may settle for some approximation of these properties. The two are intertwined, as a basic control issue is how best to avoid performing expansion inferences from premises that can be contracted, because such expansions are not necessary for fairness, and would generate redundancies. This principle has led many to design search plans called by various authors *simplification-first*, *contraction-first*, or *eager contraction* plans. Our definition of critical obligations also allows one to incorporate "critical pair criteria," as, for example, in [Bachmair and Dershowitz 1988].

On the other hand, making sure that contraction takes priority over expansion is not cost-free, because it involves keeping a potentially very large database of formulæ *inter-reduced*. In turn, this involves *forward contraction*, that is, contracting newly generated formulæ with respect to already existing ones, and *backward contraction*, that is, contracting formulæ already in the database with respect to new formulæ that survived forward contraction. Conceptually, forward contraction is considered to be part of the generation of a formula, while backward contraction is considered to be a bookkeeping task for the database of formulæ. In practice, an observation that helped streamline implementations of completion, and of theorem-proving strategies based on completion, was that backward contraction can be implemented by forward contraction. That is, it suffices to detect that a formula in the database is reducible, and then subject it to forward contraction, as if it were newly generated. This way, formulæ generated by backward contraction are treated like formulæ generated by expansion. This observation appeared in implementations since the late eighties, most notably in Otter [McCune 1994].

In our framework, the endeavor to implement contraction efficiently is the endeavor to make *contracting* derivations efficient. A sufficient condition for being contracting is $Red\,A_* \cap A_\infty = \emptyset$. One may approach the problem by aiming at ensuring that $Red\,A_i = \emptyset$, for all stages $i$ of a derivation. The practical meaning and feasibility of such a requirement depends on how one defines the map between the prover's operations and the steps $A_i \rightsquigarrow A_{i+1}$ of a derivation. If every single expansion or contraction inference done by the prover is a step $A_i \rightsquigarrow A_{i+1}$, it is trivially impossible to have $Red\,A_i = \emptyset$. Thus, either $A_i \rightsquigarrow A_{i+1}$ corresponds to many inference steps (as is the case for bulk completion), or one aims at implementing $Red\,A_* \cap A_\infty = \emptyset$ by ensuring that $Red\,A_i = \emptyset$ holds periodically.

For instance, take Otter's well-known *given-clause loop*. The prover maintains a list of formulæ *already selected* as expansion parents and a list of formulæ *to be selected*. At every iteration, it selects a *given clause*, performs all expansions between the *given clause* and the *already selected* clauses, and moves the *given clause* to the *already selected* list. Every new formula is forward-contracted after its generation, and those that survive forward contraction are added to the list

*to be selected*, and applied to backward-contract elements of both lists until no further backward contraction applies. Thus, if $A$ is the union of the two lists *already selected* and *to be selected*, Otter's given clause loop aims at something like $Red\,A_i = \emptyset$, for all $i$'s that correspond to a stage after an iteration of the loop.

A more conservative approach is to implement $Red\,A_* \cap A_\infty = \emptyset$ by ensuring that $Red\,B_i = \emptyset$ holds periodically and only for a subset $B_i \subset A_i$. This is the approach of the so-called DISCOUNT version of the given-clause loop, where only the subset of formulæ eligible to be expansion parents (the *already selected* list augmented with the given clause) is kept inter-reduced. However, when a formula in $B_i$ is backward-contracted, its direct descendants in $A_i \setminus B_i$ can be deleted as "orphans" [Schulz 2002]. Most of Otter's successors, such as Gandalf [Tammet 1997], SPASS [Weidenbach et al. 1999], Vampire [Riazanov and Voronkov 2002] and WALDMEISTER [Hillenbrand 2003], implement both versions of the given-clause loop, while the E prover [Schulz 2002] features only the DISCOUNT version.

Since contraction is, at the same time, an essential ingredient for efficiency and an expensive task, the appropriate balance of contraction and efficiency is still a subject of current research in the implementation of theorem provers.

## REFERENCES

BACHMAIR, L. AND DERSHOWITZ, N. 1988. Critical pair criteria for completion. *J. Symb. Comput. 6,* 1, 1–18.

BACHMAIR, L. AND DERSHOWITZ, N. 1989. Completion for rewriting modulo a congruence. *Theor. Comput. Sci. 67,* 2-3 (October), 173–202.

BACHMAIR, L. AND DERSHOWITZ, N. 1994. Equational inference, canonical proofs, and proof orderings. *J. ACM 41,* 2, 236–276.

BACHMAIR, L., DERSHOWITZ, N., AND PLAISTED, D. A. 1989. Completion without failure. In *Resolution of Equations in Algebraic Structures*, H. Aït-Kaci and M. Nivat, Eds. Vol. 2: Rewriting Techniques. Academic Press, London, 1–30.

BACHMAIR, L. AND GANZINGER, H. 1994. Rewrite-based equational theorem proving with selection and simplification. *J. Logic and Comput. 4,* 3, 217–247.

BACHMAIR, L. AND GANZINGER, H. 2001. Resolution theorem proving. In *Handbook of Automated Reasoning*, A. Robinson and A. Voronkov, Eds. Vol. I. Elsevier Science, Amsterdam, Chapter 2, 19–99.

BACHMAIR, L., TIWARI, A., AND VIGNERON, L. 2003. Abstract congruence closure. *J. Automat. Reason. 31,* 2, 129–168.

BONACINA, M. P. 1992. Distributed automated deduction. Ph.D. thesis, Department of Computer Science, State University of New York at Stony Brook, Stony Brook, NY.

BONACINA, M. P. 1999. A taxonomy of theorem-proving strategies. In *Artificial Intelligence Today – Recent Trends and Developments*, M. J. Wooldridge and M. Veloso, Eds. Lecture Notes in Artificial Intelligence, vol. 1600. Springer-Verlag, Berlin, 43–84.

BONACINA, M. P. AND HSIANG, J. 1992. On rewrite programs: Semantics and relationship with Prolog. *J. Logic Program. 14,* 1 & 2 (October), 155–180.

BONACINA, M. P. AND HSIANG, J. 1995. Towards a foundation of completion procedures as semidecision procedures. *Theor. Comput. Sci. 146*, 199–242.

BROWN, JR., T. C. 1975. A structured design-method for specialized proof procedures. Ph.D. thesis, California Institute of Technology, Pasadena, CA.

BUCHBERGER, B. 1985. *Multidimensional Systems Theory*. Reidel, Bose, N.K. Ed., Chapter Gröbner Bases: An Algorithmic Method in Polynomial Ideal Theory, 184–232.

BUREL, G. AND KIRCHNER, C. 2006. Completion is an instance of abstract canonical system inference. In *Algebra, Meaning and Computation – Essays in Honor of Joseph Goguen*, K. Futatsugi, J.-P. Jouannaud, and J. Meseguer, Eds. Lecture Notes in Computer Science, vol. 4060. Springer-Verlag, 497–520.

DERSHOWITZ, N. 1982. Orderings for term-rewriting systems. *Theor. Comput. Sci. 17,* 3 (Mar.), 279–301.

DERSHOWITZ, N. 1985. Computing with rewrite systems. *Inf. Control 64,* 2/3 (May/June), 122–157.

DERSHOWITZ, N. 1989. Completion and its applications. In *Resolution of Equations in Algebraic Structures, Volume 2: Rewriting Techniques*, H. Aït-Kaci and M. Nivat, Eds. Academic Press Inc., London, 31–86.

DERSHOWITZ, N. 1991a. Canonical sets of Horn clauses. In *Proceedings of the 18th International Colloquium on Automata, Languages and Programming (Madrid, Spain)*, J. L. Albert, B. Monien, and M. R. Artalejo, Eds. Lecture Notes in Computer Science, vol. 510. Springer-Verlag, Berlin, 267–278.

DERSHOWITZ, N. 1991b. Ordering-based strategies for Horn clauses. In *Proceedings of the 12th International Joint Conf. on Artificial Intelligence*. Morgan Kaufmann, Sydney, Australia, 118–124.

DERSHOWITZ, N. AND KIRCHNER, C. 2006. Abstract canonical presentations. *Theor. Comput. Sci. 357*, 53–69.

DERSHOWITZ, N., MARCUS, L., AND TARLECKI, A. 1988. Existence, uniqueness and construction of rewrite systems. *SIAM J. Comput. 17,* 4 (Aug.), 629–639.

DERSHOWITZ, N. AND PLAISTED, D. A. 2001. Rewriting. In *Handbook of Automated Reasoning*, A. Robinson and A. Voronkov, Eds. Vol. I. Elsevier Science, Amsterdam, Chapter 9, 535–610.

DERSHOWITZ, N. AND REDDY, U. 1993. Deductive and inductive synthesis of equational programs. *J. Symb. Comput. 15*, 467–494.

DOGGAZ, N. AND KIRCHNER, C. 1991. Completion for unification. *Theor. Comput. Sci. 85,* 1, 231–251.

FRIBOURG, L. 1989. A strong restriction to the inductive completion procedure. *J. Symb. Comput. 8,* 3, 253–276.

GALLIER, J., NARENDRAN, P., PLAISTED, D. A., RAATZ, S., AND SNYDER, W. 1993. Finding canonical rewriting systems equivalent to a finite set of ground equations in polynomial time. *J. ACM 40,* 1, 1–16.

HILLENBRAND, T. 2003. Citius, altius, fortius: Lessons learned from the theorem prover WALDMEISTER. In *Proceedings of the 4th International Workshop On First-Order Theorem Proving (FTP03)*, I. Dahn and L. Vigneron, Eds. Electronic Notes in Theoretical Computer Science, vol. 86. Elsevier, Valencia, Spain. See `http://www.sciencedirect.com`.

HSIANG, J. AND RUSINOWITCH, M. 1987. On word problems in equational theories. In *Proceedings of the 14th International Colloquium on Automata Languages and Programming*, T. Ottman, Ed. Lecture Notes in Computer Science, vol. 267. Springer-Verlag, Berlin, 54–71.

HSIANG, J. AND RUSINOWITCH, M. 1991. Proving refutational completeness of theorem proving strategies: The transfinite semantic tree method. *J. ACM 38,* 3 (July), 559–587.

HUET, G. 1981. A complete proof of correctness of the Knuth–Bendix completion algorithm. *J. Comput. Syst. Sci. 23,* 1 (August), 11–21.

JOUANNAUD, J.-P. AND KIRCHNER, H. 1986. Completion of a set of rules modulo a set of equations. *SIAM J. Comput. 15,* 4, 1155–1194.

KAPUR, D. AND MUSSER, D. R. 1987. Proof by consistency. *Artificial Intelligence 13,* 2, 125–157.

KNUTH, D. E. AND BENDIX, P. B. 1970. Simple word problems in universal algebras. In *Computational Problems in Abstract Algebra*, J. Leech, Ed. Pergamon Press, Oxford, 263–297.

KOUNALIS, E. AND RUSINOWITCH, M. 1991. On word problems in Horn theories. *J. Symb. Comput. 11,* 1 & 2 (January/February), 113–128.

LANKFORD, D. S. 1975. Canonical inference. Memo ATP-32, Automatic Theorem Proving Project, University of Texas, Austin, TX. Dec.

McCUNE, W. W. 1994. Otter 3.0 reference manual and guide. Tech. Rep. 94/6, MCS Division, Argonne National Laboratory. See `http://www-unix.mcs.anl.gov/AR/otter/`.

METIVIER, Y. 1983. About the rewriting systems produced by the Knuth-Bendix completion algorithm. *Inf. Process. Lett. 16,* 1 (Jan.), 31–34.

NIEUWENHUIS, R. AND RUBIO, A. 2001. Paramodulation-based theorem proving. In *Handbook of Automated Reasoning*, A. Robinson and A. Voronkov, Eds. Vol. I. Elsevier Science, Amsterdam, Chapter 7, 371–443.

PETERSON, G. AND STICKEL, M. E. 1981. Complete sets of reductions for some equational theories. *J. ACM 28*, 233–264.

PLAISTED, D. A. AND SATTLER-KLEIN, A. 1996. Proof lengths for equational completion. *Inf. Comput. 125,* 2, 154–170.

RIAZANOV, A. AND VORONKOV, A. 2002. The design and implementation of VAMPIRE. *J. AI Commun. 15,* 2/3, 91–110.

ROBINSON, G. AND WOS, L. 1969. Paramodulation and theorem-proving in first-order theories with equality. In *Machine Intelligence*, D. Michie and R. Meltzer, Eds. Vol. IV. Edinburgh Univ. Press, Edinburgh, 135–150.

RUSINOWITCH, M. 1991. Theorem-proving with resolution and superposition. *J. Symb. Comput. 11*, 21–50.

SCHULZ, S. 2002. E – A brainiac theorem prover. *J. AI Commun. 15,* 2–3, 111–126.

TAMMET, T. 1997. Gandalf. *J. Automat. Reason. 18,* 2, 199–204.

TERESE. 2003. *Term Rewriting Systems*. Cambridge University Press, Cambridge, U.K. M. Bezem, J. W. Klop and R. de Vrijer, eds.

WEIDENBACH, C., AFSHORDEL, B., BRAHM, U., COHRS, C., ENGEL, T., KEEN, E., THEOBALT, C., AND TOPIĆ, D. 1999. System description: SPASS version 1.0.0. In *Proceedings 16th Conf. on Automated Deduction, Trento (Italy)*, H. Ganzinger, Ed. Lecture Notes in Artificial Intelligence, vol. 1632. Springer-Verlag, Berlin, 378–382.