Atomic Metadeduction

Serge Autexier and Dominik Dietrich

German Research Centre for Artificial Intelligence (DFKI GmbH), Bremen, Germany

Abstract. We present an extension of the first-order logic sequent calculus SK that allows us to systematically add inference rules derived from arbitrary axioms, definitions, theorems, as well as local hypotheses – collectively called assertions. Each derived deduction rule represents a pattern of larger SK-derivations corresponding to the use of that assertion. The idea of metadeduction is to get shorter and more concise formal proofs by allowing the replacement of any assertion in the antecedent of a sequent by derived deduction rules that are available locally for proving that sequent. We prove the soundness and completeness for atomic metadeduction, which builds upon a permutability property for the underlying sequent calculus SK with liberalized δ^{++} -rule.

1 Introduction

In spite of almost four decades of research on automated theorem proving, mainly theorems considered easy by human standards can be proved fully automatically without human assistance. Many theorems still require a considerable amount of user interaction, and will require it for the foreseeable future. Hence, there is a need that proofs are presented and ideally constructed in a form that suits human users in order to provide an effective guidance.

To come close to the style of proofs as done by humans, Huang [8] introduced the so-called *assertion-level*, where individual proof steps are justified by axioms, definitions, or theorems, or even above at the so-called *proof level*, such as "by analogy". The idea of the assertion-level is, for instance, that given the facts $U \subset V$ and $V \subset W$ we can prove $U \subset W$ directly using the assertion:

 $\subset_{Trans}: \forall U. \forall V. \forall W. U \subset V \land V \subset W \Rightarrow U \subset W$

An assertion level step usually subsumes several deduction steps in standard calculi, say the classical sequent calculus [7]. To use an assertion in the classical sequent calculus it must be present in the antecedent of the sequent and be processed by means of decomposition rules, usually leading to new branches in the derivation tree. Some of these branches can be closed by means of the axiom rule which correspond to "using" that assertion on known facts or goals.

Huang followed the approach of a human oriented proof style by hiding decomposition steps once detected by abstracting them to an assertion application. Since he was mainly concerned with using the abstract representation for proof presentation in natural language [8, 6] there was no proof theoretic foundation for the assertion level. Hence, assertion level proofs could only be checked once

2 Serge Autexier and Dominik Dietrich

expanded to the underlying calculus, and the actual proof had still to be found at the calculus level and only proof parts of a specific form could be abstracted.

More recently, work was devoted to analyze the assertion-level proof theoretically: [10] defined supernatural deduction that extends the natural deduction calculus by inference rules derived from assertions and showed its soundness and completeness. This work was extended to the classical sequent calculus in [5] to obtain the superdeduction calculus. However, both approaches are restricted to closed, universally quantified equations or equivalences and the premises and conclusions of the derived inference rules are restricted to atomic formulas. In this paper we extend that work to derive and use inference rules from arbitrary formulas, including non-closed formulas such as local hypotheses, but still allow only for atomic premises and conclusions. Hence the name atomic metadeduction. Compared to [5] we use a different meta-theory based on a sequent calculus with a liberalized δ -rule (δ^{++} , [4]) which enables the necessary proof transformations to establish soundness and completeness.

The paper is organized as follows: In Sec. 2 we present a minimal sequent calculus for first-order logic with liberalized δ^{++} -rule and give two permutability results due to the use of that rule. In Sec. 3 we present the technique to compute derived inference rules from arbitrary assertions. In Sec. 4 we prove the soundness and completeness of the calculus using derived inference rules, define the metadeduction calculus and prove that the rule that allows us to move assertions from sequents to the inference level is invertible, i.e. we do not lose provability by applying it. We conclude the paper by summarizing the main results and comparing it to related work in Sec. 5.

2 Sequent Calculus with Liberalized Delta Rule

The context of this work is first-order logic. First-order terms and atomic formulas are build as usual inductively over from functions \mathcal{F} , predicates \mathcal{P} and variables \mathcal{V} . The *formulas* are then build inductively from falsity \bot , atomic formulas, the connective \Rightarrow and universal quantification \forall . For the formal parts of this paper we use the restricted set of connectives, but also the other connectives for sake of readability. Finally, *syntactic equality* on formulas is modulo renaming of bound variables (α -renaming) and denoted by =. Our notion of *substitution* is standard: A substitution is a function $\sigma: \mathcal{V} \to T(\Sigma, \mathcal{V})$ which is the identity but for finitely many $x \in \mathcal{V}$ and whose homomorphic extension to terms and formulas is idempotent. We use $t\sigma$ to denote the application of σ to t.

The sequent calculus for first-order logic is given in Fig. 1 is mostly standard. The specificities are: (i) The axiom rule AXIOM is restricted to atomic formulas; (ii) the \forall_L -rule allows us to substitute terms with free variables to postpone the choice of instances; (iii) there is a substitution rule SUBST to substitute free variables globally in the derivation tree; the idempotency of substitutions ensures the admissibility of the substitution; (iv) the \forall_R -rule uses Skolemization with an optimization regarding the used Skokem-function: Standard Skolemization requires that the Skolem-function f is new wrt. the *whole* sequent and takes

AXIOM
$$\frac{\Gamma, F, F \vdash \Delta}{\Gamma, F \vdash \Delta}$$
 $A = \operatorname{contr}_{L} \frac{\Gamma, F, F \vdash \Delta}{\Gamma, F \vdash \Delta}$
 $\Rightarrow_{L} \frac{\Gamma \vdash F, \Delta}{\Gamma, F \Rightarrow G \vdash \Delta} \Rightarrow_{R} \frac{\Gamma, F \vdash G, \Delta}{\Gamma \vdash F \Rightarrow G, \Delta}$ $\forall_{L} \frac{\Gamma, F[t/x] \vdash \Delta}{\Gamma, \forall x.F \vdash \Delta} t \text{ term that}$

$$\forall_{R} \frac{\Gamma \vdash F[f_{[\forall x.F]}(\vec{Z})/x], \Delta}{\Gamma \vdash \forall x.F, \Delta}$$
SUBST $\forall x \neq \sigma$ $\forall t \neq \sigma$ substitution

where
$$\vec{Z} = \mathcal{FV}(\forall x.F)$$
 SUBST $x \rightarrow t$ if σ substitution

Fig. 1: The Sequent Calculus SK with liberalized δ^{++} -rule and Skolemization

as arguments *all* variables that occur free in the sequent. Contrary to that we use the even more liberalized ${\delta^+}^+$ approach [4]: when eliminating the universal quantification for some succedent formula $\forall x.F$, first it allows us to take the *same* Skolem function for all formulas that are equal modulo α -renaming to $\forall x.F$; such Skolem-functions are denoted by $f_{[\forall x.F]}$ where $[\forall x.F]$ denotes the set of all formulas equal to $\forall x.F$. Secondly, the arguments to the Skolem function are only those variables that actually occur freely in $\forall x.F$. The result of using the ${\delta^+}^+$ -approach is that \forall_R for some sequent $\Gamma, \forall x.F$. The result of using the ${\delta^+}^+$ -approach is that \forall_R for some sequent $\Gamma, \forall x.F \vdash \Delta$ is invariant with respect to different Γ and Δ , and that it allows for shorter and also more natural proofs (see [11] for a survey and [12] for soundness and completeness proofs¹). The CUT-rule $\frac{\Gamma \vdash F, \Delta}{\Gamma \vdash \Delta} \stackrel{\Gamma, F \vdash \Delta}{\Gamma \vdash \Delta}$ is admissible for this sequent calculus, that is every proof using CUT can be transformed into a proof without CUT.

For every rule, any formula occurring in the conclusion but not in any of the premises is a *principal formula* of that rule,² while any formula occurring in some premise but not in the conclusion is a *side formula*. All rules have the subformula property, that is all side formulas are subformulas of the principal formulas (or instances thereof).

Permutability. The use of the liberalized δ^{++} -rule (\forall_R) allows for a specific permutability result on proof steps. The observation is that applying it to $\Gamma \vdash \forall x.F, \Delta$ the used Skolem function and variables used as arguments only depend on the formula $\forall x.F$ and variables that occur free in it. Hence, if we always introduce new free variables in \forall_L -applications and postpone the application of SUBST to the end of any proof attempt (only followed by AXIOM-rule applications), the chosen Skolem-functions and their arguments only depend on F and any previous \forall_L -applications to a formula of which $\forall x.F$ is a subformula. As a result the \forall_R -step with principal formula $\forall x.F$ can be permuted with any proof step having a principal formula that is "independent" from $\forall x.F$. More generally, any two successive rule applications with principal formulas that are "independent" of each other can be permuted (cf. Lemma 2.3 in [3])).

3

¹ [12] who uses explicit variable conditions instead of Skolemization and substitution, but the same Eigenvariable for all syntactically equal formulas.

² For CONTR_L, F is the principal formula and its copy in the premise a side formula; for AXIOM both A are principal formulas and so is \perp in the rule \perp_L .

4 Serge Autexier and Dominik Dietrich

Given two successive sequent rule applications of respective principal formulas F and G. F and G are *independent* from each other, if G is not a side formula for F, i.e., not a subformula of F. Otherwise we say that G is a subformula of F. The notion of independent formulas serves to define rule applications that are irrelevant in a proof. A rule application R in some proof is *irrelevant*, iff none of the subformulas of R's principal formula is an active partner in an AXIOM- or \perp_L -rule application in that proof. A proof without irrelevant rule applications is called *concise* and it is folklore to show that any proof can be turned into a concise proof by removing all irrelevant proof steps. Throughout the rest of this paper we will assume concise proofs.

A consequence of being able to permute proof steps working on independent principal formulas is that we can group together in any SK^{*} derivation all rules working on a principal formula and all its subformulas, where SK^{*} denotes SK without the SUBST rule and the restriction of always introducing new free variables in \forall_L -steps. To formalize that observation, we introduce the concept of *A*-active derivations to denote those derivations where only rule applications with principal formula *A* or one of its subformulas are applied.

Definition 1 (A-Active/Passive Derivations). Let L, L' be multisets of formulas and D a derivation (possibly with open goals) for the sequent $\Gamma, L \vdash L', \Delta$. The derivation D is (L, L')-active, if it contains only calculus rules having a formula from L or L' or one of their subformulas as principal formula. Conversely, we say D is (L, L')-passive if it contains no calculus rule that has some formula from L or L' or one of their subformulas as principal formula. If $L = \{A\}$ and $L' = \emptyset$ (respectively if $L' = \{A\}$ and $L = \emptyset$) then we agree to say D is A-active if it is (L, L')-active and A-passive if it is (L, L')-passive.

It holds that every A-active rule followed by A-passive derivations can be permuted to first have an A-passive derivation followed by applications of the Aactive rule on the different open sequents (cf. Corollary 2.7 in [3]). Using that we can transform any SK* derivation into one composed of A-active derivations (cf. Lemma 2.8 in [3]). However, we can do even better and move in any A-active derivation all applications of CONTR_L on A or one of its subformulas downwards to be applied on A already. To formalize this observation, we introduce the notion of *contraction-free* derivations which are derivations without applications of the contraction rule (cf. Lemma 2.9 in [3]).

3 Derived Sequent Rules

In this section we present the technique to obtain derived inference rules for an arbitrary formula. As motivation consider the derivation from Fig. 2a illustrating the application of the assertion $(\subset_{\mathsf{Trans}}) \forall A, B, C.A \subset B \land B \subset C \Rightarrow A \subset C$. It shows given the facts $\Gamma = \{U \subset V, V \subset W\}$ we can show $\Delta = \{U \subset W\}$. The crucial steps in this derivation are to use the assertion $(\subset_{\mathsf{Trans}})$ with the instantiation [U/A, V/B, W/C] to show $U \subset W$. The other steps can be understood as unfolding or preparation steps, yielding several branches in the derivation

Fig. 2: Motivating Example

$$(C \oplus 2 \oplus) \frac{\Gamma', U \subset V, V \subset W \vdash U \subset W, \Delta'}{\Gamma', U \subset V, V \subset W \vdash U \subset W, \Delta'} (C \oplus) \frac{\Gamma', U \subset V \vdash V \subset W, \Delta'}{\Gamma', U \subset V \vdash \Delta'}$$

$$(C 2 \oplus) \frac{\Gamma', V \subset W \vdash U \subset W, U \subset V, \Delta'}{\Gamma', V \subset W \vdash U \subset W, \Delta'} (C \oplus) \frac{\Gamma', V \subset W \vdash U \subset V, \Delta'}{\Gamma', V \subset W \vdash \Delta'}$$

$$(C \oplus 3) \frac{\Gamma', U \subset V \vdash U \subset W, V \subset W, \Delta'}{\Gamma', U \subset V \vdash U \subset W, \Delta'} (C \oplus) \frac{\Gamma' \vdash U \subset V, U \subset W, \Delta'}{\Gamma' \vdash U \subset W, \Delta'}$$

Fig. 3: Further Derived Rules

tree, some of which can be closed using the axiom rule and the available facts Γ or goals Δ . In general, there will be several possibilities to apply the assertion (\subset_{Trans}), and the number of the new branches created by the application of the assertion will depend on the available facts Γ and goals Δ .

For instance, if $\Delta = \{\}$ and $\Gamma = \{U \subset V, V \subset W\}$, then the axiom rule is no longer applicable in 3 which gets a new open sequent, that is the derivation in Fig. 2b. Note that this remains a valid derivation, if we add arbitrary formulas Γ' to the antecedent or Δ' to the succedent. If additionally, we drop the assertion from the conclusion sequent, then we obtain the derived inference rule in Fig. 2c (the numbers in the rule name $(\subset \mathbb{O}^{2})$ indicate the actually applied axiom rules). We get a variety of these inferences depending on which application of axiom rules are enabled by filling the Γ and Δ ; these rules all represent one possible application of the assertion (\subset_{Trans}). However, if there is not at least one axiom rule application, then we do not consider this as an application of $(\subset_{\mathsf{Trans}})$ (otherwise the rule would always be applicable); moreover, this derivation is somehow superfluous if none of the subformulas of $(\subset_{\mathsf{Trans}})$ is used in the proof. With that restriction, we get 7 possibilities to apply the assertion (3 single axiom rule applications, 3 double axiom rule applications and 1 triple axiom rule application). This results in the derived inference rules (in addition to $(\subset \mathbb{O}^{2})$) shown in Fig. 3. Thereby, U, V and W are arbitrary terms as they stem from applications of \forall_L rules. Skolem functions introduced by \forall_R -rules are always the same, which results from the use of the δ^{+^+} rule, where we use the same Skolem function for the same formulas. In the case of derived rules, these are always the subformulas of the assertion which are always the same. Systematizing that results in the following rule synthesis procedure:

$$\underbrace{\begin{array}{cccc} \Gamma_{1},H_{1}\vdash H_{2},\Delta_{1} & \dots & \Gamma_{n},H_{1}\vdash H_{2},\Delta_{n} \\ \hline H_{1},F\vdash H_{2} \\ (a) \\ \end{array}}_{(b)} \underbrace{\Gamma_{i_{1}},\sigma(H_{1})\vdash\sigma(H_{2}),\Delta_{i_{1}} & \dots & \Gamma_{i_{l}},\sigma(H_{1})\vdash\sigma(H_{2}),\Delta_{i_{l}} \\ \hline \sigma(H_{1}),F\vdash\sigma(H_{2}) \\ \hline \sigma(H_{2}),\sigma(H_{2}) \\ (b) \\ \end{array}}_{(b)}$$

Fig. 4: Intermediate Stages of the Computation of Derived Rules

Definition 2 (Derived Inference Rules). Let SK0 denote the subset of SK^* without the rule CONTR_L. Given a not necessarily closed formula F, we compute the derived rules from F as follows: Take the sequent $H_1, F \vdash H_2$, where H_1 and H_2 are place-holders for lists of formulas and apply exhaustively all rules from SK0. All obtained derivation trees are of the form in Fig. 4a Then enable one or more application of the AXIOM rule by instantiating H_1 and H_2 with atoms respectively from some Δ_i or some Γ_i which results in (multiple) derivations of the form in Fig. 4b, where σ is the respective instantiation of H_1 and H_2 . For each of these trees we introduce the derived rule

BY F
$$\frac{\Gamma, \Gamma_{i_1}, \sigma(H_1) \vdash \sigma(H_2), \Delta_{i_1}, \Delta}{\Gamma, \sigma(H_1) \vdash \sigma(H_2), \Delta} \dots \frac{\Gamma, \Gamma_{i_l}, \sigma(H_1) \vdash \sigma(H_2), \Delta_{i_l}, \Delta}{\Gamma, \sigma(H_1) \vdash \sigma(H_2), \Delta}$$

Using that we can derive for any, not necessarily closed formula F a set of derived so-called F-rules. For $(\subset_{\mathsf{Trans}})$ we get the rules from Figs. 2c and 3. Note that the derived rules are strongly interrelated; they can be divided in two classes: forward rules working on antecedent formulas $((\subset ①), (\subset ②), (\subset \odot ②))$, and backward rules working on the succedent and possibly on the antecedent $((\subset \odot @)), (\subset \odot @))$. While these rules are in the spirit of Huang's assertion level and reflect all possibilities to apply an assertion – hence well suited for an interactive setting – they introduce redundancy in the search space if used without care in an automated setting. For example, the backward inference $(\subset ③)$ is more general than the other backward inferences and sufficient for the search.

4 The Metadeduction Calculus

We now formally define the metadeduction calculus: First, we define *theory sequent calculi* and prove their soundness and weak completeness. We then define the metadeduction calculus by adding a lifting rule that enables to replace assertions from the antecedent of some open goal sequent by corresponding inference rules in the meta-level theory of the sequent and prove, besides its soundness, its *inversion property*, that is we do not lose provability by this operation.

Definition 3 (Theory Sequent Calculus). Let Th be a set of not necessarily closed formulas: Then we denote by $\Gamma \vdash_{Th} \Delta$ a theory sequent wrt. Th and we allow to write $\Gamma \vdash_{Th,F} \Delta$ to denote the sequent wrt. the theory Th augmented by the formula F. The theory sequent calculus consists of the sequent calculus rules of SK and for each theory sequent $\Gamma \vdash_{Th} \Delta$ of all rules derived from all formulas in Th. The SUBST-rule now affects the antecedent and succedent of sequents and the formulas in the attached theory Th (resp. the derived rules).

7

We prove soundness of the theory sequent calculus by constructing from any proof for some sequent $\Gamma \vdash_{Th,F} \Delta$ using *F*-rules a proof for $\Gamma, F \vdash_{Th} \Delta$ not using *F*-rules. This allows us to eliminate step by step all theory formulas and end up in the classical sequent calculus (that is, $Th = \emptyset$).

Theorem 1 (Soundness). For all sets of formulas Th, all formulas F and all proofs of $\Gamma \vdash_{Th,F} \Delta$ there exists a proof for $\Gamma, F \vdash_{Th} \Delta$ without F-rules.

Conversely, we prove completeness by constructing from any proof of some sequent $\Gamma, Th \vdash \Delta$ a proof for $\Gamma \vdash_{Th} \Delta$. The completeness proof relies on an *self-derivability property* for derived rules, that is, if we lift an assertion F from the antecedent of our current open goal sequent to the calculus level then F is still derivable using F-rules. This allows us to reuse the proof of $\Gamma, Th \vdash \Delta$ using CUT (cf. [3], Sec. 4 for more details). Because of the use of CUT, we call the obtained result weak completeness. Future work is devoted to transform a given proof to an assertion proof without CUT using our permutability results.

Lemma 1 (Selfderivability). For every formula F there is a proof of $\vdash_F F$.

Theorem 2 (Weak Completeness). For all sets of formulas Th and proofs Π for $\Gamma, Th \vdash \Delta$ there exists a proof Π_{Th} for $\Gamma \vdash_{Th} \Delta$ possibly using CUT.

We further extend the theory sequent calculus by a rule to lift arbitrary assertions F from the antecedent of sequents to the calculus level at any stage of the proof and to apply the henceforth derived F-rules at any time: $\operatorname{LIFT} \frac{\Gamma \vdash_{Th,F} \Delta}{\Gamma, F \vdash_{Th} \Delta}$. Due to Lemma 1 the LIFT-rule has the inversion property. We call the resulting calculus the *metadeduction*-calculus which soundness and completeness directly follows from the Theorems 1 and 2.

5 Conclusion and Related Work

In this paper we have presented atomic metadeduction as an extension to a firstorder sequent calculus to systematically synthesize new derived inference rules from assertions at any time during the proof search. Using metadeduction, proofs can directly be constructed, checked, and presented at the assertion level. We have shown soundness, completeness of the theory sequent calculi and proved the inversion property of the LIFT-rule.

The idea of extending the natural deduction calculus or the sequent calculus by new deduction rules is not new [9, 10, 5] and was discussed in Sec. 1. Compared to that, we allow for derived rules from arbitrary, even non-closed assertions, which allows us to use intermediate facts as derived rules, for instance the induction hypothesis of an inductive proof or case conditions in a case analysis. However, in contrast to these works we have no cut admissibility result so far, which is a topic for future work. We expect to obtain similar results at least for the restricted fragment of closed equivalences of the form $P \Leftrightarrow Q$ where P is atomic – if not even for a larger fragment of formulas – by following ideas of [?].

8 Serge Autexier and Dominik Dietrich

Closely related are also focusing derivations [1] to eliminate inessential nondeterminism by alternating phases of asynchronous (invertible) and synchronous (non-invertible) steps. Focusing derivations decompose a chosen formula and, if the formula was an antecedent formula, it corresponds roughly to the synthesis of derived rules from it; the difference is that we apply both synchronous and asynchronous rules and at least one AXIOM-rule, which excludes applications of derived rules of definitely irrelevant formulas. Moreover, derived F-rules remain available, while focusing consumes F. Finally, the use of derived rules allows to study in future work how to adapt proof search techniques known from other calculi, such as, for instance, the use of term indexing techniques on the level of derived inference rules, or proof strategies based on term orderings as in superposition calculi.

Further future work will also be concerned with relaxing the atomicity restriction (requiring more complex proof transformations) as well as investigating how to adapt metadeduction to deep inference³ (following ideas from [2]) to eventually support the application of derived rules on subformulas.

References

- J.-M. Andreoli. Logic programming with focusing proofs in linear logic. J. Log. Comput., 2(3):297–347, 1992.
- S. Autexier. The CORE calculus. In R. Nieuwenhuis, editor, Proceedings of the 20th International Conference on Automated Deduction (CADE-20), volume 3632 of LNAI, Tallinn, Estonia, july 2005. Springer.
- 3. S. Autexier and D. Dietrich. Soundness and weak completeness of proof calculi augmented with derived inference rules restricted to atomic formulas. Seki-Report SR-09-03, DFKI Bremen, 2009.
- 4. B. Beckert, R. Hähnle, and P. H. Schmitt. The even more liberalized δ-rule in free variable semantic tableaux. In G. Gottlob, A. Leitsch, and D. Mundici, editors, *Proceedings, 3rd Kurt Gödel Colloquium (KGC), Brno, Czech Republic*, LNCS 713, pages 108–119. Springer, August 1993.
- P. Brauner, C. Houtmann, and C. Kirchner. Principle of superdeduction. In Luke Ong, editor, *Proceedings of LICS*, pages 41–50, jul 2007.
- Guillaume Burel and Claude Kirchner. Cut elimination in deduction modulo by abstract completion. In Sergei N. Artemov and Anil Nerode, editors, *LFCS*, volume 4514 of *Lecture Notes in Computer Science*, pages 115–131. Springer, 2007.
- A. Fiedler. *P.rex*: An interactive proof explainer. In R. Goré, A. Leitsch, and T. Nipkow, editors, *Automated Reasoning: Proceedings of IJCAR'01*, number 2083 in LNAI, pages 416–420, Siena, Italy, 2001. Springer.
- G. Gentzen. The Collected Papers of Gerhard Gentzen (1934-1938). Edited by Szabo, M. E., North Holland, Amsterdam, 1969.
- X. Huang. Human Oriented Proof Presentation: A Reconstructive Approach. Number 112 in DISKI. Infix, Sankt Augustin, Germany, 1996.
- D. Prawitz. Natural deduction; a proof-theoretical study, volume 3 of Stockholm Studies in Philosophy. Almqvist and Wiksells, 1965.

8

³ alessio.guglielmi.name/res/cos/

11. B. Wack. *Typage et déduction dans le calcul de réécriture*. Thèse de doctorat, Université Henri Poincaré (Nancy 1), octobre 2005.

12. C.-P. Wirth. Descente infinie + Deduction. Logic J. of the IGPL, 12(1):1–96, 2004.

13. C.-P. Wirth. Hilbert's epsilon as an operator of indefinite committed choice. J. of Applied Logic, 2007.