Reinhard Kahle · Thomas Strahm
Thomas Studer
Editors

Advances in Proof Theory

Birkhäuser

Editors
Reinhard Kahle
CMA and DM, FCT
Universidade Nova de Lisboa
Caparica
Portugal

Thomas Studer
Institute of Computer Science
University of Bern
Bern
Switzerland

Thomas Strahm
Institute of Computer Science
University of Bern
Bern
Switzerland

ISSN 2297-0576 ISSN 2297-0584 (electronic)
Progress in Computer Science and Applied Logic
ISBN 978-3-319-29196-3 ISBN 978-3-319-29198-7 (eBook)
DOI 10.1007/978-3-319-29198-7

Library of Congress Control Number: 2016931426

Mathematics Subject Classification (2010): 03F03, 03F15, 03F05, 03F50, 03B20, 03B30, 03B35, 68T15

Printed on acid-free paper

This book is published under the trade name Birkhäuser
The registered company is Springer International Publishing AG Switzerland
(www.birkhauser-science.com)

Restricting Initial Sequents: The Trade-Offs Between Identity, Contraction and Cut

Peter Schroeder-Heister

Abstract In logical sequent calculi, initial sequents expressing the axiom of identity can be required to be atomic, without affecting the deductive strength of the system. When extending the logical system with right- and left-introduction rules for atomic formulas, one can analogously require that initial sequents be restricted to "uratoms", which are undefined (not even vacuously defined) atoms. Depending on the definitional clauses for atoms, the resulting system is possibly weaker than the unrestricted one. This weaker system may however be preferable to the unrestricted system, as it enjoys cut elimination and blocks unwanted derivations arising from non-wellfounded definitions, for example in the context of paradoxes.

1 Introduction

In standard sequent calculi of first-order logic, initial sequents

$$A \vdash A,$$

which express the axiom of identity, can be restricted to atomic A. For non-atomic A, the sequent $A \vdash A$ can then be derived using the right- and left-introduction rules for the logical constants occurring in A. This way of presenting sequent calculi is quite common and has certain technical advantages, such as in the area of automated theorem proving, and also in proof theory itself. For example, arguments establishing

---

This work goes back to ideas developed during a sabbatical stay as Gerhard Jäger's guest at the University of Bern in the winter semester 1993–1994. It was completed within the French-German ANR-DFG projects "Hypothetical Reasoning—Its Proof-Theoretic Analysis" (DFG Schr 275/16-2) and "Beyond Logic—Hypothetical Reasoning in Philosophy of Science, Informatics and Law" (DFG Schr 275/17-1). I am very grateful to Roy Dyckhoff, Thomas Piecha and an anonymous reviewer for their helpful comments and suggestions.

P. Schroeder-Heister (✉)
Wilhelm-Schickard-Institut für Informatik, Universität Tübingen,
Sand 13, 72076 Tübingen, Germany
e-mail: psh@uni-tuebingen.de

R. Kahle et al. (eds.), *Advances in Proof Theory*, Progress in Computer Science and Applied Logic 28, DOI 10.1007/978-3-319-29198-7_10

height-preserving admissibility (see [12]) or arguments concerning rank-preserving admissibility as in Lemma 2 below depend on it. There is even a philosophical rationale behind this procedure: If there are *specific* rules to generate a complex proposition on the right or left side of the turnstile, these specific rules should be used. The statement $A \vdash A$, which is completely *unspecific* as to the structure of A, should only be made when no specific way of introducing A is available, that is, when A is atomic.

This situation changes, when we extend our logical system with rules for atomic formulas ("atoms"). If right- and left-introduction rules for atoms are available, then these atoms are still atomic in the logical sense, that is, they do not contain logical constants, but they are no longer atoms in the semantical sense as they have a specific meaning given by these rules. This is the case in the theory of definitional reflection (see [6–8, 14]). There one extends the logical framework with right-introduction rules for atoms based on the clauses of a definition. This definition has the form of an extended logic program allowing for logically complex formulas in bodies of clauses. The rule of definitional reflection complements these rules with a left-introduction rule based on a kind of inversion principle. However, the points discussed here apply to any extension of logical systems that provides right- and left-introduction rules for atoms (see, for example, [1, 11]).

2 The Formal System of Intuitionistic Logic with Definitional Reflection

We consider intuitionistic propositional logic, which is sufficient to make our point. Let upper case Latin letters denote formulas in this language, let lower case Roman letters denote atoms (which in our simplified framework are propositional letters), and let upper case Greek letters denote finite multisets of formulas. We suppose that a *definition* $\mathbb{D}$ is given, which consists of finitely many clauses of the form

$$a \;\Leftarrow\; A$$

Such a clause is called a *(defining) clause for* a. We furthermore assume that with every $\mathbb{D}$ its domain $dom(\mathbb{D})$ is associated, which is a set of atoms containing those atoms for which there is a definitional clause in $\mathbb{D}$, but possibly further atoms. The elements in $dom(\mathbb{D})$ are called the atoms *defined* by $\mathbb{D}$. We allow for atoms defined by $\mathbb{D}$ without there being a clause for them. These elements of $dom(\mathbb{D})$ are called *vacuously defined*. Atoms that are not defined by $\mathbb{D}$ and thus do not belong to $dom(\mathbb{D})$ are called *uratoms*. If $a \in dom(\mathbb{D})$, let $\mathbb{D}(a)$ be the set of *defining conditions* of a, that is, the set $\{A_1, \ldots, A_n\}$ if the clauses for a in $\mathbb{D}$ are as follows:

$$\begin{cases} a \Leftarrow A_1 \\ \quad\vdots \\ a \Leftarrow A_n \end{cases}$$

If a is vacuously defined, then $\mathbb{D}(a)$ is empty. If a is an uratom, then $\mathbb{D}(a)$ is undefined. (In other words, if a is not defined by $\mathbb{D}$, then $\mathbb{D}(a)$ is undefined in the metalogical sense.)

Our system of intuitionistic logic with definitional reflection over the definition $\mathbb{D}$, called **LI**($\mathbb{D}$), has the following rules of inference, where the antecedent of a sequent is understood as a multiset of formulas.

$$(I)\ \frac{}{\Gamma, a \vdash a}$$

$$(\top)\ \frac{}{\Gamma \vdash \top} \qquad (\bot)\ \frac{}{\Gamma, \bot \vdash A}$$

$$(\vdash\wedge)\ \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \qquad (\wedge\vdash)\ \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C}$$

$$(\vdash\vee)\ \frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \qquad (\vee\vdash)\ \frac{\Gamma, A \vdash C \quad \Gamma, B \vdash C}{\Gamma, A \vee B \vdash C}$$

$$(\vdash\rightarrow)\ \frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \qquad (\rightarrow\vdash)\ \frac{\Gamma, A \rightarrow B \vdash A \quad \Gamma, B \vdash C}{\Gamma, A \rightarrow B \vdash C}$$

$$(\vdash\mathbb{D})\ \frac{\Gamma \vdash C}{\Gamma \vdash a}\ C \in \mathbb{D}(a) \qquad (\mathbb{D}\vdash)\ \frac{\{\Gamma, a, C \vdash A : C \in \mathbb{D}(a)\}}{\Gamma, a \vdash A} \quad a \in dom(\mathbb{D})$$

Without the definitional rules in the last line, this is a standard variant of the intuitionistic propositional sequent calculus which we call **LI**. The last line contains the rules of definitional closure ($\vdash\mathbb{D}$) and definitional reflection ($\mathbb{D}\vdash$), which for any atom a which is defined by $\mathbb{D}$, delivers right- and left-introduction rules. The right-introduction rule says that a can be inferred from each defining condition of a. The left-introduction rule says that everything that can be inferred from each defining condition of a can be inferred from a itself. If the clauses defining a are viewed as its inductive definition, the left-introduction rule for a expresses the extremal clause for this inductive definition: "Nothing else defines a". For further discussion see [8, 14]. Note that the rules ($\vdash\mathbb{D}$) and ($\mathbb{D}\vdash$) only apply to those atoms a which are defined by $\mathbb{D}$ (i.e., $a \in dom(\mathbb{D})$) and not to uratoms. In view of ($\bot$), we can disregard vacuously defined atoms, if we identify a vacuously defined atom a with an atom a defined by the clause $a \Leftarrow \bot$.

For **LI** it is well-known that the structural rules of thinning, contraction and cut

$$(Thin)\ \frac{\Gamma \vdash C}{\Gamma, A \vdash C} \qquad (Contr)\ \frac{\Gamma, A, A \vdash C}{\Gamma, A \vdash C} \qquad (Cut)\ \frac{\Gamma \vdash A \quad \Gamma, A \vdash C}{\Gamma \vdash C}$$

are admissible (see, for example, [2, 12]). As thinning is admissible, the version of cut with separated contexts

$$\frac{\Gamma \vdash A \quad \Delta, A \vdash C}{\Gamma, \Delta \vdash C}$$

is admissible, too.

The admissibility of thinning and contraction extends from **LI** to **LI**($\mathbb{D}$). For thinning this is obvious, for contraction this is due to the fact that in ($\mathbb{D}\vdash$) the atom a is repeated in the premisses, which means that we have already built an implicit contraction into ($\mathbb{D}\vdash$). In fact, this implicit contraction is not even needed. Without loss of deductive power, we can replace ($\mathbb{D}\vdash$) with its contraction-free variant $(\mathbb{D}\vdash)^{\text{cf}}$:

$$(\mathbb{D}\vdash)^{\text{cf}} \; \frac{\{\Gamma, C \vdash A : C \in \mathbb{D}(a)\}}{\Gamma, a \vdash A}$$

This can be seen as follows. Consider the system $\mathbf{LI}^{\text{cf}}(\mathbb{D})$, which results from **LI**($\mathbb{D}$) by replacing ($\mathbb{D}\vdash$) with $(\mathbb{D}\vdash)^{\text{cf}}$. For a derivation $\mathcal{D}$ in $\mathbf{LI}^{\text{cf}}(\mathbb{D})$, the $\mathbb{D}$-rank $r_{\mathbb{D}}(\mathcal{D})$ is the maximum number of applications of $(\mathbb{D}\vdash)^{\text{cf}}$, where the maximum is taken over all branches of $\mathcal{D}$. More precisely,

$r_{\mathbb{D}}(\mathcal{D}) = 0,$ if $\mathcal{D}$ is(I), $(\top)$ or $(\bot)$,
$r_{\mathbb{D}}(\mathcal{D}) = r_{\mathbb{D}}(\mathcal{D}_1)$, if $\mathcal{D}$ is of the form $\frac{\mathcal{D}_1}{\Gamma \vdash A}$, and the last step is different from $(\mathbb{D}\vdash)^{\text{cf}}$,
$r_{\mathbb{D}}(\mathcal{D}) = max\{r_{\mathbb{D}}(\mathcal{D}_1), r_{\mathbb{D}}(\mathcal{D}_2)\}$, if $\mathcal{D}$ is of the form $\frac{\mathcal{D}_1 \quad \mathcal{D}_2}{\Gamma \vdash A}$, and the last step is different from $(\mathbb{D}\vdash)^{\text{cf}}$,
$r_{\mathbb{D}}(\mathcal{D}) = max_{C \in \mathbb{D}(a)}\{r_{\mathbb{D}}(\mathcal{D}_C)\} + 1$, if $\mathcal{D}$ is of the form $\frac{\{\mathcal{D}_C : C \in \mathbb{D}(a)\}}{\Gamma, a \vdash A}(\mathbb{D}\vdash)^{\text{cf}}$.

Then we can show the following.

Lemma 1 (Invertibility lemma) *If $\mathcal{D}$ is a derivation of* $\Gamma, a \vdash A$ *in* $\mathbf{LI}^{\text{cf}}(\mathbb{D})$*, then for each* $C \in \mathbb{D}(a)$ *we can find a derivation* $\mathcal{D}_C$ *of* $\Gamma, C \vdash A$ *such that* $r_{\mathbb{D}}(\mathcal{D}_C) \leq r_{\mathbb{D}}(\mathcal{D})$.

Proof The only non-trivial case obtains when $\mathcal{D}$ is an initial sequent $\Gamma, a \vdash a$. Here $r_{\mathbb{D}}(\mathcal{D}) = 0$. We use that for any C we can find a derivation $\mathcal{D}'_C$ of $\Gamma, C \vdash C$ in **LI**. Since $(\mathbb{D}\vdash)^{\text{cf}}$ is not used in $\mathcal{D}'_C$, we know that $r_{\mathbb{D}}(\mathcal{D}'_C) = 0$. Then, for $C \in dom(\mathbb{D})$, let $\mathcal{D}_C$ be

$$\frac{\begin{array}{c}\mathcal{D}'_C \\ \Gamma, C \vdash C\end{array}}{\Gamma, C \vdash a}(\vdash \mathbb{D})$$

Since $r_{\mathbb{D}}(\mathcal{D}'_C) = 0$, we know that $r_{\mathbb{D}}(\mathcal{D}_C) = 0$. □

Now it is easy to show that contraction is admissible in $\mathbf{LI}^{\text{cf}}(\mathbb{D})$. More precisely, we can show the following.

Lemma 2 *If $\mathcal{D}$ is a derivation of* $\Gamma, A, A \vdash C$ *in* $\mathbf{LI}^{\text{cf}}(\mathbb{D})$*, then we can find a derivation* $\mathcal{D}'$ *of* $\Gamma, A \vdash C$ *in* $\mathbf{LI}^{\text{cf}}(\mathbb{D})$*such that* $r_{\mathbb{D}}(\mathcal{D}') \leq r_{\mathbb{D}}(\mathcal{D})$.

psh@uni-tuebingen.de

Proof by induction on the triple $\langle r_{\mathbb{D}}(\mathcal{D}), deg(A), h(\mathcal{D})\rangle$, where $deg(A)$ is the logical complexity of A and $h(\mathcal{D})$ is the height of $\mathcal{D}$ (that is, the length of its longest branch). As an example we present the case in which $(\mathbb{D}\vdash)^{\text{cf}}$ is applied in the last step and the atom a introduced by $(\mathbb{D}\vdash)^{\text{cf}}$ is the contraction formula:

$$\mathcal{D}: \quad \dfrac{\left\{\begin{matrix}\mathcal{D}_C \\ \Gamma, a, C \vdash A\end{matrix} \quad : \ C \in \mathbb{D}(a)\right\}}{\Gamma, a, a \vdash A}(\mathbb{D}\vdash)^{\text{cf}}$$

We assume that a is not vacuously defined–otherwise the case is trivial. Obviously, $r_{\mathbb{D}}(\mathcal{D}_C) < r_{\mathbb{D}}(\mathcal{D})$ for every $C \in \mathbb{D}(a)$. Applying the invertibility lemma (Lemma 1) to the premiss derivations $\mathcal{D}_C$ we obtain derivations

$$\left\{\begin{matrix}\mathcal{D}'_C \\ \Gamma, C, C \vdash A\end{matrix} \quad : \ C \in \mathbb{D}(a)\right\}$$

such that $r_{\mathbb{D}}(\mathcal{D}'_C) < r_{\mathbb{D}}(\mathcal{D})$ for every C. Therefore, by induction hypothesis, we obtain derivations

$$\left\{\begin{matrix}\mathcal{D}''_C \\ \Gamma, C \vdash A\end{matrix} \quad : \ C \in \mathbb{D}(a)\right\}$$

such that $r_{\mathbb{D}}(\mathcal{D}''_C) \leq r_{\mathbb{D}}(\mathcal{D}'_C) < r_{\mathbb{D}}(\mathcal{D})$ for every C. From those we obtain a derivation

$$\mathcal{D}': \quad \dfrac{\left\{\begin{matrix}\mathcal{D}''_C \\ \Gamma, C \vdash A\end{matrix} \quad : \ C \in \mathbb{D}(a)\right\}}{\Gamma, a \vdash A}$$

such that $r_{\mathbb{D}}(\mathcal{D}') \leq r_{\mathbb{D}}(\mathcal{D})$. □

3 The Failure of Cut in LI($\mathbb{D}$)

The system **LI**($\mathbb{D}$) does not enjoy the admissibility of cut. Consider the following definition:

$$\mathbb{D}_r \ \left\{ r \Leftarrow \neg r \right.$$

(with $\neg r$ abbreviating $r \to \bot$.) Using the right- and left-introduction rules for r and the rules for implication, this leads to derivations of both $r \vdash \bot$ and of $\vdash r$:

$$
\dfrac{\dfrac{\dfrac{}{r, r \to \bot \vdash r}\,(I) \qquad \dfrac{}{r, \bot \vdash r}\,(I)}{r \to \bot, r \vdash \bot}\,(\to\vdash)}{r \vdash \bot}\,(\mathbb{D}_r\vdash)
\qquad\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{r, r \to \bot \vdash r}\,(I) \qquad \dfrac{}{r, \bot \vdash r}\,(I)}{r \to \bot, r \vdash \bot}\,(\to\vdash)}{r \vdash \bot}\,(\mathbb{D}_r\vdash)}{\vdash \neg r}\,(\vdash\to)}{\vdash r}\,(\vdash\mathbb{D}_r)
\tag{1}
$$

If cut were admissible, $\vdash \bot$ would be derivable, which is not the case as there is no right-introduction rule for $\bot$. Note that we impose no restriction on the form of definitional clauses, in particular no well-foundedness restriction. The definition $\mathbb{D}_r$ may be considered a propositional short form of Russell's paradox, obtained from the following clause for comprehension:

$$t \in \{x : A(x)\} \Leftarrow A(t)$$

by instantiating $A(x)$ with $x \notin x$ and t with $\{x : x \notin x\}$, and then abbreviating $\{x : x \notin x\} \in \{x : x \notin x\}$ by r.

In view of the fact that $(\mathbb{D}\vdash)$ is not stronger than its contraction-free variant $(\mathbb{D}\vdash)^{\text{cf}}$, instead of (1) we may consider the following pair of derivations in $\mathbf{LI}^{\text{cf}}(\mathbb{D})$:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{r, r\to\bot\vdash r}\,(I) \qquad \dfrac{}{r, \bot\vdash\bot}\,(I)}{r\to\bot, r\vdash\bot}\,(\to\vdash)}{r\to\bot\vdash r\to\bot}\,(\vdash\to)}{r\to\bot\vdash r}\,(\vdash\mathbb{D}_r) \qquad \dfrac{}{\bot \vdash \bot}\,(I)}{r \to \bot \vdash \bot}\,(\to\vdash)}{r \vdash \bot}\,(\mathbb{D}_r\vdash)^{\text{cf}}
\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{r, r\to\bot\vdash r}\,(I) \qquad \dfrac{}{r, \bot\vdash\bot}\,(I)}{r\to\bot, r\vdash\bot}\,(\to\vdash)}{r\to\bot\vdash r\to\bot}\,(\vdash\to)}{r\to\bot\vdash r}\,(\vdash\mathbb{D}_r) \qquad \dfrac{}{\bot \vdash \bot}\,(I)}{r \to \bot \vdash \bot}\,(\to\vdash)}{r \vdash \bot}\,(\mathbb{D}_r\vdash)^{\text{cf}}}{\vdash r}\,(\vdash\mathbb{D}_r)
\tag{2}
$$

This shows that it is not a particular form of contraction of atoms that needs to be used in the derivation of the paradox, but rather contraction for implicational formulas, which is the essential form of contraction in $\mathbf{LI}(\mathbb{D})$.

4 The Trade-Off Between Contraction and Cut

We have shown by an example that in $\mathbf{LI}^{\text{cf}}(\mathbb{D})$, which by means of the rules $(\wedge\vdash)$ and $(\to\vdash)$ implicitly contains contraction, and in which therefore the explicit contraction rule is admissible (Lemma 2), the rule of cut is not admissible. If we consider a contraction-free variant, in which $(\wedge\vdash)$ and $(\to\vdash)$ are replaced with

$$\frac{\Gamma, A \vdash C}{\Gamma, A \wedge B \vdash C} \quad \frac{\Gamma, B \vdash C}{\Gamma, A \wedge B \vdash C} \quad \text{and} \quad \frac{\Gamma \vdash A \quad \Delta, B \vdash C}{\Gamma, \Delta, A \to B \vdash C}$$

respectively, then cut can be eliminated. This is shown in detail in [13]. Therefore we obtain a trade-off between contraction and cut, when we add definitional rules to the logical system.

Result 1 *If the logical system contains implicit or explicit contraction, then the admission of cut makes the system inconsistent. If it contains neither implicit nor explicit contraction, then cut is admissible.*

This corresponds to the observation dating back to Fitch [5] that removing the rule of contraction may be used as a strategy to cope with the paradoxes. This result can even be refined. The form of contraction used in counterexample (1) is contraction of an atom defined by $\mathbb{D}_r$ with an atom of the same shape used in an initial sequent. In [15, 16] it was claimed that prohibiting this specific sort of contraction would be a more specific way of keeping cut admissible in the system with definitional reflection than abolishing contraction altogether. This claim is not invalidated by the fact that in $\mathbf{LI}^{\mathrm{cf}}(\mathbb{D})$ no contraction of atoms is needed, as example (2) shows. In fact, it carries over to the present situation *mutatis mutandis*. The contraction of $r \to \bot$, which is implicit in the lowermost application of $(\to\vdash)$, is a contraction of an occurrence of $r \to \bot$, in which r stems from an initial sequent, with an occurrence of $r \to \bot$, in which r is a result of $(\vdash\mathbb{D}_r)$. This means that there is still an identification of occurrences of atoms which are generated by different (structural vs. meaning-giving) rules, though in example (2) it is not definitional reflection $(\mathbb{D}_r\vdash)^{\mathrm{cf}}$ but the introduction of the atom r on the right side according to $(\vdash\mathbb{D}_r)$, which is involved. Identifying such critical forms of contraction can, for example, be achieved by attaching labels to formulas that indicate when definitional rules are applied (see [3, 4]). A further elaboration of this topic, which will result in a more finegrained specification of the rules of contraction, is beyond the scope of this paper.

5 Restricting Initial Sequents: The Admissibility of Cut

In initial sequents $\Gamma, a \vdash a$ of $\mathbf{LI}(\mathbb{D})$, a can be an arbitrary atom. According to what was said in the introduction, we now restrict the atom a in initial sequents to uratoms by replacing (I) with the following rule:

$$(I)^{\circ} \quad \frac{}{\Gamma, a \vdash a} \quad a \notin dom(\mathbb{D})$$

As the rule of definitional reflection we use the contraction-free rule $(\mathbb{D}\vdash)^{\mathrm{cf}}$. The resulting system, with $(I)^{\circ}$ instead of (I) and $(\mathbb{D}\vdash)^{\mathrm{cf}}$ instead of $(\mathbb{D}\vdash)$, is called $\mathbf{LI}^{\circ}(\mathbb{D})$. Lemmas 1 and 2 continue to hold for $\mathbf{LI}^{\circ}(\mathbb{D})$. The proof of Lemma 1 is now trivial, as initial sequents involving a defined atom a can no longer occur. Due to the restriction on initial sequents, Lemma 2 is easier to prove. Therefore in $\mathbf{LI}^{\circ}(\mathbb{D})$ contraction is admissible. The counterexample (2) to cut no longer works, as it uses initial sequents for r. These initial sequents are not available in $\mathbf{LI}^{\circ}(\mathbb{D}_r)$, because r is defined in $\mathbb{D}_r$, and is thus not an uratom.

In fact, in the system $\mathbf{LI}^{\circ}(\mathbb{D})$ we can eliminate cuts. To demonstrate this, we use as an induction measure the $\mathbb{D}$-*weight* of a formula occurrence in a derivation. Unlike the $\mathbb{D}$-rank as used in Lemmas 1 and 2, the $\mathbb{D}$-weight is not a measure of a derivation $\mathcal{D}$, but a measure of a formula occurrence at a certain place in (a sequent in) a derivation $\mathcal{D}$. In the following, upper indices distinguish occurrences of formulas. For example, C^1 and C^2 denote different occurrences of the formula C. It is always assumed that a formula occurrence below an inference line corresponds to or results from a particular formula occurrence (or from particular formula occurrences) above the line. For example, in an application of $(\vee\vdash)$ of the form

$$\frac{A_1^1, \ldots, A_n^1, A^1 \vdash C^1 \quad A_1^2, \ldots, A_n^2, B^1 \vdash C^2}{A_1^3, \ldots, A_n^3, (A \vee B)^1 \vdash C^3}$$

it is assumed that, for all i $(1 \leq i \leq n)$, the occurrence A_i^3 corresponds to the occurrences A_i^1 and A_i^2, the occurrence C^3 corresponds to the occurrences C^1 and C^2, and the occurrences of A and B as immediate subformulas of $(A \vee B)^1$ correspond to the occurences A^1 and B^1, respectively.

Then the $\mathbb{D}$-weight $w_{\mathbb{D}}(C^1)$ of a formula occurrence C^1 in a given derivation is defined by induction on the construction of the derivation.

Each formula occurrence in $(I)^{\circ}$, $(\top)$ or $(\bot)$ has $\mathbb{D}$-weight 0.

If the last step is

$$(\vdash\to)\ \frac{A_1^1, \ldots, A_n^1, A^1 \vdash B^1}{A_1^2, \ldots, A_n^2 \vdash (A \to B)^1}$$

then $w_{\mathbb{D}}(A_i^2) = w_{\mathbb{D}}(A_i^1)$ for all i $(1 \leq i \leq n)$, and $w_{\mathbb{D}}((A \to B)^1) = max\{w_{\mathbb{D}}(A^1), w_{\mathbb{D}}(B^1)\}$. For $(\wedge\vdash)$ and $(\vdash\vee)$ the $\mathbb{D}$-weight is defined in the same way.

If the last step is

$$(\vee\vdash)\ \frac{A_1^1, \ldots, A_n^1, A^1 \vdash C^1 \quad A_1^2, \ldots, A_n^2, B^1 \vdash C^2}{A_1^3, \ldots, A_n^3, (A \vee B)^1 \vdash C^3}$$

then $w_{\mathbb{D}}(A_i^3) = max\{w_{\mathbb{D}}(A_i^1), w_{\mathbb{D}}(A_i^2)\}$ for all i $(1 \leq i \leq n)$, $w_{\mathbb{D}}(C^3) = max\{w_{\mathbb{D}}(C^1), w_{\mathbb{D}}(C^2)\}$ and $w_{\mathbb{D}}((A \vee B)^1) = max\{w_{\mathbb{D}}(A^1), w_{\mathbb{D}}(B^1)\}$, and analogously for $(\vdash\wedge)$.

If the last step is

$$(\to\vdash)\ \frac{A_1^1, \ldots, A_n^1, (A \to B)^1 \vdash A^1 \quad A_1^2, \ldots, A_n^2, B^1 \vdash C^1}{A_1^3, \ldots, A_n^3, (A \to B)^2 \vdash C^2}$$

then $w_{\mathbb{D}}(A_i^3) = max\{w_{\mathbb{D}}(A_i^1), w_{\mathbb{D}}(A_i^2)\}$ for all i $(1 \le i \le n)$, $w_{\mathbb{D}}(C^2) = w_{\mathbb{D}}(C^1)$, and $w_{\mathbb{D}}((A \to B)^2) = max\{w_{\mathbb{D}}((A \to B)^1), w_{\mathbb{D}}(A^1), w_{\mathbb{D}}(B^1)\}$.

If we consider derivations with cut and if the last step is

$$(Cut)\ \frac{A_1^1, \ldots, A_n^1 \vdash A^1 \quad A_1^2, \ldots, A_n^2, A^2 \vdash C^1}{A_1^3, \ldots, A_n^3 \vdash C^2}$$

then $w_{\mathbb{D}}(A_i^3) = max\{w_{\mathbb{D}}(A_i^1), w_{\mathbb{D}}(A_i^2)\}$ for all i $(1 \le i \le n)$, and $w_{\mathbb{D}}(C^2) = w_{\mathbb{D}}(C^1)$.

If the last step is a $\mathbb{D}$-rule, then $w_{\mathbb{D}}(a)$ is increased by 1. More precisely, if this step is

$$(\vdash\mathbb{D})\ \frac{A_1^1, \ldots, A_n^1 \vdash C^1}{A_1^2, \ldots, A_n^2 \vdash a^1}\ C \in \mathbb{D}(a)$$

then $w_{\mathbb{D}}(A_i^2) = w_{\mathbb{D}}(A_i^1)$ for all i $(1 \le i \le n)$, and $w_{\mathbb{D}}(a^1) = w_{\mathbb{D}}(C^1) + 1$. If it is

$$(\mathbb{D}\vdash)^{\mathrm{cf}}\ \frac{\{A_1^i, \ldots, A_n^i, C_i^1 \vdash A^i : 1 \le i \le k\}}{A_1^{k+1}, \ldots, A_n^{k+1}, a^1 \vdash A^{k+1}}\ \text{where } \mathbb{D}(a) = \{C_1, \ldots, C_k\},$$

then $w_{\mathbb{D}}(A_j^{k+1}) = max_{1 \le i \le k}\{w_{\mathbb{D}}(A_j^i)\}$ for all j $(1 \le j \le n)$, and $w_{\mathbb{D}}(a^1) = max_{1 \le i \le k}$ $\{w_{\mathbb{D}}(C_i^1)\} + 1$. If $\mathbb{D}(a) = \emptyset$, and thus $k = 0$, then $w_{\mathbb{D}}(A_j^1) = w_{\mathbb{D}}(A^1) = 0$ for all j $(1 \le j \le n)$ and $w_{\mathbb{D}}(a^1) = 1$.

It is important to notice that the weight increases not only at the application of $(\mathbb{D}\vdash)^{\mathrm{cf}}$, but also at the application of $(\vdash\mathbb{D})$. However, the crucial point is that only the formula a introduced by $(\vdash\mathbb{D})$ or $(\mathbb{D}\vdash)^{\mathrm{cf}}$ is affected by the increase of weight, not the parametric context formulas. Putting it another way, we may consider the $\mathbb{D}$-rank to be a measure of a sequent within a derivation, namely the $\mathbb{D}$-rank of the subderivation with this sequent as its end-sequent. In contradistinction to that, the $\mathbb{D}$-weight is a measure of a formula occurrence within a sequent, namely the $\mathbb{D}$-weight of this formula occurrence within the subderivation that has this sequent as its end-sequent.

Now cut can be eliminated even though contraction is admissible. More precisely, we can show *weight-preserving cut elimination* in the following sense.

Theorem 1 *Consider* $\mathbf{LI}^\circ(\mathbb{D})$ *extended with the cut rule* (Cut)*. Suppose a derivation* $\mathcal{D}$ *in this system is given that ends with an application*

$$\frac{A_1^1, \ldots, A_n^1 \vdash A^1 \quad A_1^2, \ldots, A_n^2, A^2 \vdash C^1}{A_1^3, \ldots, A_n^3 \vdash C^2}$$

of cut such that the derivations of its premisses are cut-free. Then we can construct a cut-free derivation of $A_1^4, \ldots, A_n^4 \vdash C^3$ *such that* $w_{\mathbb{D}}(A_i^4) \le w_{\mathbb{D}}(A_i^3)$ *for all* i $(1 \le i \le n)$ *and* $w_{\mathbb{D}}(C^3) \le w_{\mathbb{D}}(C^2)$.

Proof by induction on the triple $\langle w_{\mathbb{D}}(A), deg(A), h(\mathcal{D})\rangle$, where $w_{\mathbb{D}}(A) = max(w_{\mathbb{D}}(A^1), w_{\mathbb{D}}(A^2))$, and, as before, $deg(A)$ is the logical complexity of A and $h(\mathcal{D})$ is the height of $\mathcal{D}$. The value $w_{\mathbb{D}}(A)$ is also called the *weight of the cut formula*. The reduction steps are as usual in cut elimination. We just indicate where the $\mathbb{D}$-rules are involved, and where the definition of *weight* comes into play. The main reduction of the $\mathbb{D}$-rules reduces

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash B\end{matrix}}{\Gamma \vdash a}(\vdash \mathbb{D}) \qquad \dfrac{\left\{\begin{matrix}\mathcal{D}_C\\ \Gamma, C \vdash A\end{matrix} \quad : \ C \in \mathbb{D}(a)\right\}}{\Gamma, a \vdash A}(\mathbb{D}\vdash)^{\mathrm{cf}}}{\Gamma \vdash A} \quad B \in \mathbb{D}(a)$$

to

$$\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash B\end{matrix} \qquad \begin{matrix}\mathcal{D}_B\\ \Gamma, B \vdash A\end{matrix}}{\Gamma \vdash A}$$

by reducing a cut with cut-formula a to a cut with cut-formula B of lower weight.

An example of a permutation of cut with an application of a $\mathbb{D}$-rule is the reduction of

$$\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash A\end{matrix} \qquad \dfrac{\begin{matrix}\mathcal{D}_2\\ \Gamma, A \vdash C\end{matrix}}{\Gamma, A \vdash a}(\vdash \mathbb{D})}{\Gamma \vdash a} \quad C \in \mathbb{D}(a)$$

to

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash A\end{matrix} \qquad \begin{matrix}\mathcal{D}_2\\ \Gamma, A \vdash C\end{matrix}}{\Gamma \vdash C}}{\Gamma \vdash a}(\vdash \mathbb{D})$$

Here it is crucial that even though the weight of the cut formula A might not decrease, the weight of a is not increased, as it is solely dependent on the weight of C, which is untouched by the transformation. A measure such as the rank $r_{\mathbb{D}}$ that applies to sequents in a proof rather than formula occurrences in a sequent could not deliver this behaviour.

The restriction of initial sequents to uratoms becomes significant, when we consider the situation

$$\frac{\Gamma \vdash a^1 \quad \Gamma, a^2 \vdash a^3}{\Gamma \vdash a^4}$$

which is reduced to

$$\Gamma \vdash a^1$$

As the right premiss of the cut is an initial sequent, $w_{\mathbb{D}}(a^2) = w_{\mathbb{D}}(a^3) = w_{\mathbb{D}}(a^4) = 0$. According to our restriction on initial sequents, a is an uratom. It is easy to see that the weight of any occurrence of an uratom is 0, which means that $w_{\mathbb{D}}(a^1) = 0$. Thus $w_{\mathbb{D}}(a^1) = w_{\mathbb{D}}(a^4) = 0$. The latter equation would not necessarily hold if we admitted initial sequents with defined atoms. If a is not an uratom, then it is possible that $w_{\mathbb{D}}(a^1) > 0$ and thus $w_{\mathbb{D}}(a^1) > w_{\mathbb{D}}(a^4)$, contrary to what is claimed in the theorem. □

The admissibility of contraction and cut in $\mathbf{LI}^{\circ}(\mathbb{D})$ implies that the unrestricted rule (I) is not derivable[1] in $\mathbf{LI}^{\circ}(\mathbb{D})$. For if it were derivable, then using the definition $\mathbb{D}_r$ and applying cut to the derivations (2), we could derive $\vdash \bot$.

6 The Trade-Off Between Identity and Contraction/Cut

We have shown that by restricting initial sequents to the case where a is an uratom, we obtain a system in which both contraction and cut are admissible. If we do not restrict initial sequents, contraction is still admissible, but cut ceases to be admissible. This means that, in a sense, we have traded identity against contraction and cut.[2]

Result 2 *If identity as expressed by initial sequents is restricted to uratoms, then the rules of contraction and cut are admissible. If identity is admitted for any atom, whether defined or not, then the admission of cut makes the system inconsistent.*

It should be noted that, in the presence of unrestricted identity, a restriction on cut is an option which should not be excluded. A way of restricting cut to cases where it continues to be admissible, consists, for example, in using certain term assignments and corresponding provisos for the application of cut, as sketched in [17]. Here we do not want to enter a philosophical discussion on which one of unrestricted identity, unrestricted contraction or unrestricted cut is the preferential rule, but just point to the trade-offs between these principles in the presence of definitional reflection.

As the counterexample given in (1) or (2) is from the domain of paradoxes, this shows that restrictions on any of the principles of identity, contraction and cut can block paradoxes (see [16]).

[1] As the rule (I) is an axiom, derivability and admissibility mean the same.

[2] In the system with unseparated contexts chosen here, unrestricted identity plus cut implies contraction, so that cut and contraction cannot be separated. This was pointed out to the author by Roy Dyckhoff.

7 Restricted Initial Sequents in Logic Programming

Within logic, and in particular within the discussion of the paradoxes, the restriction of initial sequents, and thus of identity, has never been properly considered, in contrast, for example, to the issue of contraction, which is a strong topic in this debate. However, in the realm of logic programming, this issue has always been present. In fact, the idea of definitional reflection has been developed in close parallel with related issues in logic programming (see [8]). The restriction of initial sequents considered here was first proposed by Kreuger [10]. He motivated it by considerations concerning the operational interpretation of definitional reflection in the programming language GCLA, which implemented definitional reflection. Instead of formally specifying a domain $dom(\mathbb{D})$, he adds the clause $a \Leftarrow a$ to $\mathbb{D}$ to trivialize the application of $(\mathbb{D}\vdash)$ for every a which in our terminology is an uratom.

In their proof-theoretic framework for logic programming, Jäger and Stärk [9] use a classical one-sided Schütte-Tait-style sequent calculus, which contains rules for the evaluation of atoms that correspond to our rules of definitional closure and reflection. They develop a three-valued semantics for this system and explicitly consider identity-free derivations to be the proof-theoretic approach most faithful to this semantics. For identity-free derivations they prove a cut elimination theorem by translating proofs into a system with ramified $\mathbb{D}$-rules, for which the cut elimination proof is completely standard, and then retranslate cut-free proofs. This translation and retranslation crucially depends on the fact that identity in the unrestricted form (I) is lacking. This method can easily be carried over to the situation considered here. Thus Jäger and Stärk implicitly point to the trade-off between unrestricted identity and the availability of cut, which was the main topic of this paper.

References

1. J. Brotherston, A. Simpson, Complete sequent calculi for induction and infinite descent, in *Proceedings of the 22nd Annual IEEE Symposium on Logic in Computer Science (LICS)* (IEEE Press, Los Alamitos, 2007), pp. 51–62
2. A.G. Dragalin, *Mathematical Intuitionism: Introduction to Proof Theory* (American Mathematical Society, Providence, Rhode Island, 1987). Russian original 1979
3. J. Ekman, *Normal proofs in set theory*, Ph.D. thesis, Unversity of Göteborg, 1994
4. J. Ekman, Self-contradictory reasoning, in *Advances in Proof-Theoretic Semantics*, ed. by T. Piecha, P. Schroeder-Heister (Springer, Cham, 2016), pp. 211–229
5. F.B. Fitch, A system of formal logic without an analogue to the Curry W operator. J. Symb. Log. **1**, 92–100 (1936)
6. L. Hallnäs, Partial inductive definitions. Theoret. Comput. Sci. **87**, 115–142 (1991)
7. L. Hallnäs, On the proof-theoretic foundation of general definition theory. Synthese **148**, 589–602 (2006)
8. L. Hallnäs, P. Schroeder-Heister, A proof-theoretic approach to logic programming: I. Clauses as rules. II. Programs as definitions. J. Logic Comput. **1**, 261–283, 635–660 (1990/91)
9. G. Jäger, R.F. Stärk, A proof-theoretic framework for logic programming, in *Handbook of Proof Theory*, ed. by S.R. Buss (Elsevier, Amsterdam, 1998), pp. 639–682

psh@uni-tuebingen.de

10. P. Kreuger, Axioms in definitional calculi, in *Extensions of Logic Programming. 4th International Workshop, ELP'93, St Andrews, U.K., Mar/Apr 1993. Proceedings* ed. by R. Dyckhoff. (*Lecture Notes in Computer Science*, vol. 798), (Springer, Berlin, 1994), pp. 196–205
11. P. Martin-Löf, Hauptsatz for the intuitionistic theory of iterated inductive definitions, in *Proceedings of the Second Scandinavian Logic Symposium*, ed. by J.E. Fenstad (North-Holland, Amsterdam, 1971), pp. 179–216
12. S. Negri, J. von Plato, *Structural Proof Theory* (Cambridge University Press, Cambridge, 2001)
13. P. Schroeder-Heister, Cut elimination in logics with definitional reflection, in *Nonclassical Logics and Information Processing: International Workshop (Berlin, November 1990), Proceedings*, ed. by D. Pearce, H. Wansing, (*Lecture Notes in Computer Science*, vol. 619) (Springer, Berlin, 1992), pp. 146–171
14. P. Schroeder-Heister, Rules of definitional reflection, in *Proceedings of the 8th Annual IEEE Symposium on Logic in Computer Science (Montreal 1993)* (IEEE Press, Los Alamitos, 1993), pp. 222–232
15. P. Schroeder-Heister, On the notion of *assumption* in logical systems, in *Selected Papers Contributed to the Sections of GAP5, Fifth International Congress of the Society for Analytical Philosophy*, Bielefeld, 22–26 Sept 2003, ed. by R. Bluhm, C. Nimtz (Mentis, Paderborn, 2004), pp. 27–48. Online publication: http://www.gap5.de/proceedings
16. P. Schroeder-Heister, Paradoxes and structural rules, in *Insolubles and Consequences: Essays in Honour of Stephen Read*, ed. by C. Dutilh Novaes, O. Hjortland (College Publications, 2012), pp. 203–211
17. P. Schroeder-Heister, Proof-theoretic semantics, self-contradiction, and the format of deductive reasoning, Topoi **31**, 77–85 (2012)