The Journal of Symbolic Logic
Volume 80, Number 3, September 2015

CONSERVATIVITY FOR THEORIES OF COMPOSITIONAL TRUTH VIA CUT ELIMINATION

GRAHAM E. LEIGH

Abstract. We present a cut elimination argument that witnesses the conservativity of the compositional axioms for truth (without the extended induction axiom) over any theory interpreting a weak subsystem of arithmetic. In doing so we also fix a critical error in Halbach's original presentation. Our methods show that the admission of these axioms determines a hyper-exponential reduction in the size of derivations of truth-free statements.

§1. Overview. Let $\mathsf{I}\Delta_0 + \mathsf{exp}$ and $\mathsf{I}\Delta_0 + \mathsf{exp}_1$ be the first-order theories extending Robinson's arithmetic by Δ_0-induction and, respectively, axioms expressing the totality of the exponentiation and hyper-exponentiation function. If S is a first-order theory interpreting $\mathsf{I}\Delta_0 + \mathsf{exp}$ then by $\mathsf{CT}[\mathsf{S}]$ we denote the extension of S by a fresh unary predicate T and the *compositional axioms of truth* for T.[1]

In this paper we provide syntactic proofs for the following theorems.

Theorem 1.1. *Let* S *be an elementary axiomatised theory in a finite language interpreting* $\mathsf{I}\Delta_0 + \mathsf{exp}$. *Then* $\mathsf{CT}[\mathsf{S}]$ *conservatively extends* S. *Moreover, this fact is verifiable in* $\mathsf{I}\Delta_0 + \mathsf{exp}_1$.

Let p be a fresh unary predicate symbol not present in the language $\mathcal{L}$ of S. An $\mathcal{L}$-formula D is an S-*schema* if $\mathsf{S} \vdash \mathsf{D}\ulcorner\sigma\urcorner \to \sigma$ for every $\mathcal{L}$-formula σ and there exists a finite set of formulæ U such that $\mathsf{S} \vdash \mathsf{D}x \to \exists\psi \bigvee_{\varphi\in U}(x = \ulcorner\varphi[\psi/p]\urcorner)$.

Theorem 1.2. *Let* S *be as above. For any* S-*schema* D, *the theory* $\mathsf{CT}[\mathsf{S}] + \forall x(\mathsf{D}x \to \mathrm{T}x)$ *is a conservative extension of* S. *Moreover, this fact is verifiable in* $\mathsf{I}\Delta_0 + \mathsf{exp}_1$.

The first part of Theorem 1.1 was first established by Barwise and Schlipf in the early 70s (see Theorem IV.5.3 of [1]) and later independently proved by Kotlarski, Krajewski, and Lachlan [13] for the case of $\mathsf{S} = \mathsf{PA}$, also establishing the first part of Theorem 1.2 in this case. Both proofs are model-theoretic, showing that a countable nonstandard model of S contains a full satisfaction class if it is recursively saturated. Since every model of S is elementarily extended by a recursively saturated model of the same cardinality, conservativity is obtained. Recently, Enayat and

Received July 31, 2013.
Key words and phrases. proof theory, cut elimination, axiomatic truth, compositional theory of truth.

[1] See definition 2.1 below. Note axiom schemata of S are not expanded in $\mathsf{CT}[\mathsf{S}]$.

0022-4812/15/8003-0006
DOI:10.1017/jsl.2015.27

Visser [2] provided an alternative argument (again model-theoretic) that establishes both theorems as well as their formalisation in weak arithmetic (the special case of $\mathsf{S} = \mathsf{PA}$ is outlined in [3]).

Halbach [7] offers a proof-theoretic account of Theorem 1.1. The strategy proceeds as follows. First the theory CT[S] is reformulated as a finitary sequent calculus with a cut rule and rules of inference in place of each of the compositional axioms for truth. A typical derivation in this calculus features cuts on formulæ involving the truth predicate and for general S the system will not admit cut elimination. Instead, Halbach outlines a method of partial cut elimination whereby every cut on a formula involving the truth predicate is systematically replaced by a derivation with cuts only on truth-free formulæ. Halbach's proof, however, contains a critical error (see Section 3.7 below and also Theorem 8.5 of [9]). Nevertheless, the argument yields a method to eliminate cuts of a very particular kind, namely those on formulæ $\mathrm{T}(s)$ for which it is derivable (within S) that the logical depth of the formula coded by s is bounded by some closed term.

The present paper provides the necessary link between the CT[S] and its fragment with bounded cuts. This takes the form of the following lemma (proved in Section 5).

Bounding Lemma. *If Γ and Δ are finite sets consisting of truth-free and atomic formulæ only, and the sequent $\Gamma \Rightarrow \Delta$ is derivable in* CT[S], *then there exists a derivation of this sequent in which all cuts are either truth-free or bounded.*

Let $\mathsf{CT}^*[\mathsf{S}]$ denote the subsystem of CT[S] featuring only bounded cuts. Halbach's result shows that this calculus permits the elimination of all cuts containing the truth predicate. Thus the first part of Theorem 1.1 is a consequence of the above lemma. Moreover, the proof determines bounds on the size of the resulting derivation from which the second part of Theorem 1.1 can be readily deduced.

A particular instance of Theorem 1.2 of interest is if S is a *schematic* theory (in the sense of [4]) and D is the predicate Ax_S expressing the property of encoding an axiom of S. Ax_PA, for instance, can be seen as a PA-schema by choosing U to consist of $p(0) \wedge \forall x(p(x) \to p(x+1)) \to \forall x\, p(x)$ and the noninduction axioms of PA. In this case we notice that the reduction of CT[S] to $\mathsf{CT}^*[\mathsf{S}]$ also yields a reduction of $\mathsf{CT}[\mathsf{S}] + \forall x(\mathsf{Ax}_\mathsf{S} x \to \mathrm{T}x)$ to the extension of $\mathsf{CT}^*[\mathsf{S}]$ by the rule

$$\frac{\Gamma \Rightarrow \Delta, \mathsf{Ax}_\mathsf{S} s}{\Gamma \Rightarrow \Delta, \mathrm{T}s}\ (\mathsf{Ax}_\mathsf{S}).$$

It is not immediately clear whether this new theory admits a cut elimination process for T-cuts. Instead we show that this extension of $\mathsf{CT}^*[\mathsf{S}]$ is relatively interpretable in CT[S], whence Theorem 1.1 yields Theorem 1.2.

1.1. Outline. In Sections 2 and 3 we formally define the theory CT[S] for suitable S and its presentation as a sequent calculus, as well as the sub-theory with bounded cuts, $\mathsf{CT}^*[\mathsf{S}]$. Section 4 contains the technical lemmas necessary for the Bounding Lemma and main Theorems, the proofs of which form the content of Section 5. In Section 6 we present applications of our analysis to questions relating to interpretability and speed-up and in the final section we discuss future avenues of research.

§**2. Preliminaries.** We are interested in first-order theories that possess the mathematical resources to develop their own meta-theory. It is well-known that only a weak fragment of arithmetic is required for this task, namely $\mathsf{I}\Delta_0 + \mathsf{exp}$. For our purposes we therefore take the interpretability of $\mathsf{I}\Delta_0 + \mathsf{exp}$ as representing that a theory possess the resources to express basic properties about its own syntax.

Let $\mathcal{L}$ be a finite first-order language and S an $\mathcal{L}$-theory interpreting $\mathsf{I}\Delta_0 + \mathsf{exp}$. It will be useful to work with an extension of $\mathcal{L}$ that includes a countable list of fresh predicate symbols $\{\mathsf{p}^i_j \mid i, j < \omega\}$ where p^i_j has arity i, plus a fresh propositional constant ε; we denote this extended language by $\mathcal{L}^+$. The *logical depth* of a formula α in $\mathcal{L}^+$, denoted $d(\alpha)$, is given by: $d(\alpha) = 0$ if α is atomic; $d(\forall x\alpha) = d(\exists x\alpha) = d(\neg\alpha) = d(\alpha) + 1$; and $d(\alpha_0 \vee \alpha_1) = d(\alpha_0 \wedge \alpha_1) = \max\{d(\alpha_0), d(\alpha_1)\}$.

We fix some standard representation of $\mathcal{L}^+$ in $\mathsf{I}\Delta_0 + \mathsf{exp}$, which takes the form of a fixed simple Gödel coding of $\mathcal{L}^+$ into $\mathcal{L}$ with:

1. Predicates $\mathsf{Term}_{\mathcal{L}^{(+)}}x$, $\mathsf{Form}_{\mathcal{L}^{(+)}}x$, $\mathsf{Sent}_{\mathcal{L}^{(+)}}x$, and $\mathsf{Var}\,x$ of $\mathcal{L}$ expressing respectively the relations that x is the code of a closed term, a formula, a sentence, and a variable symbol of $\mathcal{L}^{(+)}$.
2. A Σ_1-predicate $\mathsf{val}(x, y)$ such that $\mathsf{val}(\ulcorner t \urcorner, t)$ is provable in the base theory for every term t. We view val as defining a function, writing $\mathsf{val}\,x$, and write $\mathsf{eq}(r, s)$ in place of $\forall x \forall y(\mathsf{val}(r, x) \wedge \mathsf{val}(s, y) \rightarrow x = y)$.
3. Predicates defining operations on codes; namely the binary terms $\underset{\cdot}{=}$, $\underset{\cdot}{\wedge}$, $\underset{\cdot}{\vee}$, $\underset{\cdot}{\rightarrow}$, $\underset{\cdot}{\forall}$, $\underset{\cdot}{\exists}$, $\underset{\cdot}{\mathsf{p}}$, unary terms $\underset{\cdot}{\mathsf{Q}}$ for each relation Q in $\mathcal{L}$ and $\underset{\cdot}{d}$, and a ternary term *sub* with:
   - $\underset{\cdot}{\mathsf{Q}}(\langle \ulcorner t_1 \urcorner, \dots, \ulcorner t_n \urcorner \rangle) = \ulcorner \mathsf{Q}(t_1, \dots, t_n) \urcorner$ for each $\mathsf{Q} \in \mathcal{L}$,
   - $\underset{\cdot}{\mathsf{p}}(\bar{\jmath}, \langle \ulcorner t_1 \urcorner, \dots, \ulcorner t_i \urcorner \rangle) = \ulcorner \mathsf{p}^i_j(t_1, \dots, t_i) \urcorner$,
   - $\underset{\cdot}{d}(\ulcorner \alpha \urcorner) = x$ if the logical depth of the $\mathcal{L}^+$ formula α is x, and
   - $sub(x, y, z)$ denoting the usual substitution function that replaces in the term or formula (encoded by) x each occurrence of the variable with code y by the term with code z. We abbreviate uses of this function by writing $x[z/y]$ in place of $sub(x, y, z)$.

Definition 2.1. Let S be some fixed theory in a recursive language $\mathcal{L}$ which interprets $\mathsf{I}\Delta_0 + \mathsf{exp}$. The theory $\mathsf{CT}[\mathsf{S}]$ is formulated in the language $\mathcal{L}_\mathrm{T} = \mathcal{L} \cup \{\mathrm{T}\}$ and consists of the axioms of S together with

$$\begin{aligned}
\mathsf{Term}_{\mathcal{L}}x \wedge \mathsf{Term}_{\mathcal{L}}y &\rightarrow (\mathrm{T}(x \underset{\cdot}{=} y) \leftrightarrow \mathsf{eq}(x, y)),\\
\mathsf{Sent}_{\mathcal{L}}x \wedge \mathsf{Sent}_{\mathcal{L}}y &\rightarrow (\mathrm{T}(x \underset{\cdot}{\wedge} y) \leftrightarrow \mathrm{T}x \wedge \mathrm{T}y),\\
\mathsf{Sent}_{\mathcal{L}}x \wedge \mathsf{Sent}_{\mathcal{L}}y &\rightarrow (\mathrm{T}(x \underset{\cdot}{\vee} y) \leftrightarrow \mathrm{T}x \vee \mathrm{T}y),\\
\mathsf{Sent}_{\mathcal{L}}x \wedge \mathsf{Sent}_{\mathcal{L}}y &\rightarrow (\mathrm{T}(x \underset{\cdot}{\rightarrow} y) \leftrightarrow (\mathrm{T}x \rightarrow \mathrm{T}y)),\\
\mathsf{Sent}_{\mathcal{L}}x &\rightarrow (\mathrm{T}(\underset{\cdot}{\neg} x) \leftrightarrow \neg \mathrm{T}x),\\
\mathsf{Var}\,y \wedge \mathsf{Sent}_{\mathcal{L}}(\underset{\cdot}{\forall} yx) &\rightarrow (\mathrm{T}(\underset{\cdot}{\forall} yx) \leftrightarrow \forall z(\mathsf{Term}_{\mathcal{L}}z \rightarrow \mathrm{T}(x[z/y]))),\\
\mathsf{Var}\,y \wedge \mathsf{Sent}_{\mathcal{L}}(\underset{\cdot}{\forall} yx) &\rightarrow (\mathrm{T}(\underset{\cdot}{\exists} yx) \leftrightarrow \exists z(\mathsf{Term}_{\mathcal{L}}z \wedge \mathrm{T}(x[z/y]))),\\
\mathsf{Term}_{\mathcal{L}}x_1 \wedge \cdots \wedge \mathsf{Term}_{\mathcal{L}}x_n &\rightarrow (\mathrm{T}(\underset{\cdot}{\mathsf{Q}}\langle x_1, \dots, x_n \rangle) \leftrightarrow \mathsf{Q}(\mathsf{val}x_1, \dots, \mathsf{val}x_n)).
\end{aligned}$$

for each relation Q of $\mathcal{L}$ (with arity n). We call the formulæ above the *compositional axioms for* $\mathcal{L}$ and any formula not containing the truth predicate T-*free*.

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

REMARK 2.2. The quantifier axioms of $\mathsf{CT}[\mathsf{S}]$ formalise the thought that a formula $\forall x\varphi$ is true iff $\varphi[s/x]$ is true for every term s. It is necessary, therefore, that the encoding of $\mathcal{L}$ provides a name for every element in the intended domain. This is already the case if S is an arithmetic theory. For set theories it can be achieved by adding a term, say $\langle\emptyset, x\rangle$, to the (encoded) lanaguage with the interpretation $\mathsf{val}\langle\emptyset, x\rangle = x$ for every x, whereby the quantifier axiom would read '$\forall x\varphi$ is true iff $\varphi[\langle\emptyset, y\rangle/x]$ is true for every y' and from the final axiom one can conclude $\forall x\forall y(\mathrm{T}(\langle\emptyset, x\rangle \dot{\in} \langle\emptyset, y\rangle) \leftrightarrow x \in y)$.

An alternative approach, used for example in [3] and [11], is to consider $\mathsf{CT}[\mathsf{S}]$ derived from the theory 'S *with a full satisfaction class*'. In place of the compositional axioms for truth one instead states axioms for a binary *satisfaction* predicate $\mathsf{S}(x, y)$ expressing 'x is a variable assignment satisfying y' in accordence with usual Tarskian semantics. Truth, in the sense of $\mathsf{CT}[\mathsf{S}]$, becomes a defined notion: $\mathsf{T}(y) \leftrightarrow \mathsf{Sent}_{\mathcal{L}}(y) \land \forall x\, \mathsf{S}(x, y)$.

For the purposes of the present paper there is no essential difference between the two formulations. We opt for the former as it permits a more concise presentation (at least from the perspective of cut elimination) and matches more closely with the formulations of Halbach.

Finally, we fix a few notational conventions for the remainder of the paper. The start of the Greek lower-case alphabet, α, β, γ, etc., will be used to represent formulæ of $\mathcal{L}_\mathrm{T} = \mathcal{L} \cup \{\mathrm{T}\}$, while the end, φ, χ, ψ, ω, as well as Roman lower-case symbols r, s, etc. denote terms in $\mathcal{L}$ (the former list will be used exclusively as meta-variables ranging over terms representing formulæ of $\mathcal{L}^+$). Upper-case Greek letters, Γ, Δ, Σ etc., are for finite sets of $\mathcal{L}_\mathrm{T}$ formulæ and boldface lower-case Greek symbols $\boldsymbol{\varphi}$, $\boldsymbol{\psi}$, etc. represent finite sequences of $\mathcal{L}$-terms. For a sequence $\boldsymbol{\varphi} = (\varphi_0, \ldots, \varphi_k)$, $\mathrm{T}\boldsymbol{\varphi}$ denotes the set $\{\mathrm{T}\varphi_i \mid i \le k\}$. As usual, Γ, α is shorthand for $\Gamma \cup \{\alpha\}$ and Γ, Δ for $\Gamma \cup \Delta$.

§3. Two sequent calculi for compositional truth.

Let S be a fixed theory extending $\mathsf{I}\Delta_0 + \mathsf{exp}$ formulated in the language $\mathcal{L}$. We present sequent calculi for $\mathsf{CT}[\mathsf{S}]$ and $\mathsf{CT}^*[\mathsf{S}]$. In the former calculus, derivations are finite and the calculus supports the elimination of all cuts on nonatomic formulæ containing the truth predicate. The latter system replaces the cut rule of $\mathsf{CT}[\mathsf{S}]$ by two restricted variants: one of these is the ordinary cut rule applicable to only formulæ not containing T; the other is a cut rule for the atomic truth predicate which is only applicable if the formula under the truth predicate subject to the cut has, provably, a fixed finite logical depth. This second variant turns out to be admissible, so any sequent derivable in $\mathsf{CT}^*[\mathsf{S}]$ has a derivation containing only T-free cuts. It follows therefore, that $\mathsf{CT}^*[\mathsf{S}]$ is a conservative extension of S. We show that any $\mathsf{CT}[\mathsf{S}]$ derivation can be transformed into a derivation in $\mathsf{CT}^*[\mathsf{S}]$ and hence obtain the conservativity of $\mathsf{CT}[\mathsf{S}]$ over S.

We now list the axioms and rules of $\mathsf{CT}[\mathsf{S}]$ and $\mathsf{CT}^*[\mathsf{S}]$.

3.1. Axioms.

1. $\Gamma \Rightarrow \Delta, \varphi$ if φ is an axiom of S,
2. $\Gamma, r = s, \mathrm{T}r \Rightarrow \Delta, \mathrm{T}s$ for all terms r and s,
3. $\Gamma, \mathrm{T}r \Rightarrow \mathsf{Sent}(r), \Delta$ for every r.

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

3.2. Basic rules.

$$\frac{\Gamma \Rightarrow \Delta, \alpha}{\Gamma \Rightarrow \Delta, \forall v_i \alpha}\ (\forall R) \qquad \frac{\Gamma, \alpha(s/v_i) \Rightarrow \Delta}{\Gamma, \forall v_i \alpha \Rightarrow \Delta}\ (\forall L)$$

$$\frac{\Gamma \Rightarrow \Delta, \alpha, \beta}{\Gamma \Rightarrow \Delta, \alpha \vee \beta}\ (\vee R) \qquad \frac{\Gamma, \alpha \Rightarrow \Delta \qquad \Gamma, \beta \Rightarrow \Delta}{\Gamma, \alpha \vee \beta \Rightarrow \Delta}\ (\vee L)$$

$$\frac{\Gamma, \alpha \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg\alpha}\ (\neg R) \qquad \frac{\Gamma \Rightarrow \Delta, \alpha}{\Gamma, \neg\alpha \Rightarrow \Delta}\ (\neg L)$$

$$(\mathrm{Cut}^*)\ \frac{\Gamma, \alpha \Rightarrow \Delta \qquad \Gamma \Rightarrow \Delta, \alpha}{\Gamma \Rightarrow \Delta}\ \text{provided } \alpha \text{ is T-free}$$

3.3. Truth rules.

$$(\vee_{\mathrm{T}}\mathrm{R})\ \frac{\Gamma \Rightarrow \Delta, \mathrm{T}\psi_0, \mathrm{T}\psi_1}{\Gamma, \mathsf{Sent}\psi, \psi = \psi_0 \underset{\cdot}{\vee} \psi_1 \Rightarrow \Delta, \mathrm{T}\psi} \qquad (\forall_{\mathrm{T}}\mathrm{R})\ \frac{\Gamma \Rightarrow \Delta, \mathrm{T}(\psi_0[v_i/s])}{\Gamma, \mathsf{Sent}\psi, \psi = \underset{\cdot}{\forall} s\psi_0 \Rightarrow \Delta, \mathrm{T}\psi}$$

$$(\vee_{\mathrm{T}}\mathrm{L})\ \frac{\Gamma, \mathrm{T}\psi_0 \Rightarrow \Delta \qquad \Gamma, \mathrm{T}\psi_1 \Rightarrow \Delta}{\Gamma, \mathsf{Sent}\psi, \psi = \psi_0 \underset{\cdot}{\vee} \psi_1, \mathrm{T}\psi \Rightarrow \Delta} \qquad (\forall_{\mathrm{T}}\mathrm{L})\ \frac{\Gamma, \mathrm{T}(\psi_0[t/s]) \Rightarrow \Delta}{\Gamma, \mathsf{Sent}\psi, \psi = \underset{\cdot}{\forall} s\psi_0, \mathrm{T}\psi \Rightarrow \Delta}$$

$$(\neg_{\mathrm{T}}\mathrm{R})\ \frac{\Gamma, \mathrm{T}\psi_0 \Rightarrow \Delta}{\Gamma, \mathsf{Sent}\psi, \psi = \underset{\cdot}{\neg}\psi_0 \Rightarrow \Delta, \mathrm{T}\psi} \qquad (\neg_{\mathrm{T}}\mathrm{L})\ \frac{\Gamma \Rightarrow \Delta, \mathrm{T}\psi_0}{\Gamma, \mathsf{Sent}\psi, \psi = \underset{\cdot}{\neg}\psi_0, \mathrm{T}\psi \Rightarrow \Delta}$$

$$(=_{\mathrm{T}}\mathrm{R})\ \frac{\Gamma \Rightarrow \Delta, \mathsf{eq}(r, s)}{\Gamma, \mathsf{Sent}\psi, \psi = (r \underset{\cdot}{=} s) \Rightarrow \Delta, \mathrm{T}\psi} \qquad (=_{\mathrm{T}}\mathrm{L})\ \frac{\Gamma, \mathsf{eq}(r, s) \Rightarrow \Delta}{\Gamma, \mathsf{Sent}\psi, \psi = (r \underset{\cdot}{=} s), \mathrm{T}\psi \Rightarrow \Delta}$$

3.4. T-cut rules. In CT[S]:

$$(\mathrm{Cut_T})\ \frac{\Gamma, \mathrm{T}\varphi \Rightarrow \Delta \qquad \Gamma \Rightarrow \Delta, \mathrm{T}\varphi}{\Gamma \Rightarrow \Delta}$$

In CT^*[S]:

$$(\mathrm{Cut}^k_{\mathrm{T}})\ \frac{\Gamma, \mathrm{T}\varphi \Rightarrow \Delta \qquad \Gamma \Rightarrow \Delta, \mathrm{T}\varphi \qquad \Gamma, \mathsf{Sent}\varphi \Rightarrow^* \underset{\cdot}{d}(\varphi) \leq \bar{k}}{\Gamma \Rightarrow \Delta}$$

Normal eigenvariable conditions apply to the four quantifier rules. In the final rule $(\mathrm{Cut}^k_{\mathrm{T}})$ the sequent arrow $\Rightarrow^*$ expresses that this sequent has a derivation using only axioms and basic rules (called a *truth-free* derivation). We refer to the two rules $(\mathrm{Cut_T})$ and $(\mathrm{Cut}^k_{\mathrm{T}})$ collectively as T-*cuts*.

3.5. Derivations. Derivations in either CT[S] or CT^*[S] are finite trees defined in the ordinary manner; the *truth depth* of a derivation is the maximum number of truth rules occurring in a path through the derivation. The *truth rank* is the least r such that for any rule $(\mathrm{Cut}^m_{\mathrm{T}})$ occurring on the derivation, $m < r$. The *rank* of a derivation is the pair (a, r), where a is the truth depth of the derivation and r the truth rank. By definition, a truth-free derivation has rank $(0, 0)$. We say (a, r) is *bounded* by (b, s) if $a \leq b$ and $r \leq s$.

3.6. Meta-theorems for compositional truth. The key fact we require from $\mathsf{I}\Delta_0 +$ exp is that the theory suffices to show that codes for $\mathcal{L}^+$ formulæ are uniquely decomposable.

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

Lemma 3.1 (Unique readability lemma). *The sequent* $\Gamma \Rightarrow \Delta$ *is derivable in* $\mathsf{I\Delta_0} + \mathsf{exp}$ *whenever one of the following conditions hold.*

1. Γ *is a doubleton subset of* $\{x = y_0 \underset{.}{\vee} z_0, x = \underset{.}{\forall} y_1 z_1, x = (y_2 \underset{.}{=} z_2), x = \underset{.}{\neg} y_3\}$.
2. $\{\mathsf{Sent}_{\mathcal{L}}(y), \mathsf{Sent}_{\mathcal{L}}(z)\} \subseteq \Gamma$, $\Gamma \cap \{x = y \underset{.}{\vee} z, x = z \underset{.}{\vee} y, x = \underset{.}{\forall} zy, x = \underset{.}{\neg} y\} \neq \emptyset$ *and* $\{\mathsf{Sent}_{\mathcal{L}}(x)\} \subseteq \Delta$;
3. $\{y_0 = y_1 \wedge z_0 = z_1\} \subseteq \Delta$ *and* Γ *extends*
   (a) $\{x = y_0 \underset{.}{\vee} z_0, x = y_1 \underset{.}{\vee} z_1\}$,
   (b) $\{x = \underset{.}{\forall} y_0 z_0, x = \underset{.}{\forall} y_1 z_1\}$, *or*
   (c) $\{x = (y_0 \underset{.}{=} z_0), x = (y_1 \underset{.}{=} z_1)\}$;
4. $\{y_0 = y_1\} \subseteq \Delta$ *and* $\{x = \underset{.}{\neg} y_0, x = \underset{.}{\neg} y_1\} \subseteq \Gamma$.

Lemma 3.2. $\Gamma \Rightarrow \Delta$ *is derivable in* $\mathsf{I\Delta_0} + \mathsf{exp}$ *whenever one of the following conditions hold.*

1. $\emptyset \neq \Gamma \subseteq \{x = y \underset{.}{\vee} z, x = z \underset{.}{\vee} y, x = \underset{.}{\forall} zy, x = \underset{.}{\neg} y\}$ *and* $\{\underset{.}{d}(y) < \underset{.}{d}(x)\} \subseteq \Delta$;
2. $\emptyset \neq \Gamma \subseteq \{x = y \underset{.}{\vee} z, x = z \underset{.}{\vee} y, x = \underset{.}{\forall} zy, x = \underset{.}{\neg} y, x = (y \underset{.}{=} z), x = (z \underset{.}{=} y)\}$ *and* $\{y < x\} \subseteq \Delta$;
3. $\{\underset{.}{d}(x) \leq x\} \subseteq \Delta$.

Lemma 3.3 (Embedding lemma for CT). *Suppose* T *does not occur in* $\mathcal{L}$ *and* $\mathsf{CT[S]} \vdash \alpha$. *Then the sequent* $\emptyset \Rightarrow \alpha$ *has a derivation according to the rules of* $\mathsf{CT[S]}$.

Halbach's Theorem 3.1 of [7] (see also Theorem 8.10 of [8]) aims to provide an argument for eliminating T-cuts in $\mathsf{CT[S]}$ derivations. An important case is omitted which cannot be resolved within the context of the Theorem (this is outlined in the next section). The missing case turns out to unproblematic if the derivation in question satisfies additional assumptions that are present in $\mathsf{CT^*[S]}$ derivations.

Lemma 3.4. *If the sequents* $\Gamma \Rightarrow \Delta, \mathrm{T}\chi$ *and* $\Gamma, \mathrm{T}\chi \Rightarrow \Delta$ *have derivations in* CT^* *with ranks* (a, r) *and* (b, r) *respectively, and* $\Gamma \Rightarrow \underset{.}{d}(\chi) \leq \bar{r}$ *has a truth-free derivation, then the sequent* $\Gamma \Rightarrow \Delta$ *has a derivation with rank bounded by* $((a + b) \cdot 2, r)$.

Proof. We provide only a sketch of the argument. The missing elements can be readily constructed from the outline below and other cut-elimination arguments on theories of truth (see, for example, [8, Thm 8.10], [15], and [14]) and are left as an exercise for the reader.

The proof proceeds via induction on the sum of the heights of the two derivations. We may assume that in each of the two derivations the final rule applied introduces the (distinguished) formula $\mathrm{T}\chi$. Suppose the first derivation ends with an application of $(\forall_{\mathrm{T}}\mathrm{R})$. Then $a = a' + 1$ and there are terms s_0 and χ_0 such that the formula $\chi = \underset{.}{\forall} s_0 \chi_0$ is a member of Γ and the sequent

$$\Gamma \Rightarrow \Delta, \mathrm{T}\chi, \mathrm{T}(\chi_0[v_i/s_0])$$

is derivable with rank (a', r). As $\mathrm{T}\chi$ is assumed principal in $\Gamma, \mathrm{T}\chi \Rightarrow \Delta$, the final rule applied in this derivation is one of $(\vee_{\mathrm{T}}\mathrm{L})$, $(\forall_{\mathrm{T}}\mathrm{L})$, $(\neg_{\mathrm{T}}\mathrm{L})$, and $(=_{\mathrm{T}}\mathrm{L})$. If this is any rule other than $(\forall_{\mathrm{T}}\mathrm{L})$ there will be terms χ_0' and χ_1' such that either $\{\chi = \underset{.}{\forall} s_0 \chi_0, \chi = \chi_0' \underset{.}{\vee} \chi_1'\} \subseteq \Gamma$, $\{\chi = \underset{.}{\forall} s_0 \chi_0, \chi = (\chi_0' \underset{.}{=} \chi_1')\} \subseteq \Gamma$ or $\{\chi = \underset{.}{\forall} s_0 \chi_0, \chi = \underset{.}{\neg} \chi_0'\} \subseteq \Gamma$, whence $\Gamma \Rightarrow \Delta$ follows by the unique readability lemma. Thus we may assume $(\forall_{\mathrm{T}}\mathrm{L})$ is applied to obtain $\Gamma, \mathrm{T}\chi \Rightarrow \Delta$ and so there are terms s_1, χ_1, and t such that $\{\chi = \underset{.}{\forall} s_0 \chi_0, \chi = \underset{.}{\forall} s_1 \chi_1\} \subseteq \Gamma$ and

$$\Gamma, \mathrm{T}\chi, \mathrm{T}\chi_1[t/s_1] \Rightarrow \Delta$$

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

has a derivation with rank (b', r) for some $b' < b$. Then there is some $r' < r$ for which the sequents

$$\Gamma \Rightarrow s_0 = s_1 \wedge \chi_0 = \chi_1 \qquad\qquad \Gamma \Rightarrow \underset{\cdot}{d}(\chi_0[v_i/s_0]) \leq \bar{r}'$$

are truth-free derivable and so by term substitution we obtain a derivation of

$$\Gamma, \mathrm{T}\chi, \mathrm{T}\chi_0[t/s_0] \Rightarrow \Delta,$$

with rank (b', r). Applying the induction hypothesis yields derivations of

$$\Gamma, \mathrm{T}\chi_0[t/s_0] \Rightarrow \Delta \qquad\qquad \Gamma \Rightarrow \Delta, \mathrm{T}\chi_0[v_i/s_0]$$

with ranks bounded by $((a + b') \cdot 2, r)$ and $((a' + b) \cdot 2, r)$ respectively. Substituting t for v_i in the second derivation and applying $(\mathrm{Cut}^{r'}_{\mathrm{T}})$ yields a derivation of $\Gamma \Rightarrow \Delta$ with appropriate rank. ⊣

Iterating applications of the lemma, we obtain.

Theorem 3.5 (Cut elimination for CT^*). *Suppose* $\Gamma \Rightarrow \Delta$ *is derivable in* CT^* *with rank* $(a, r + 1)$. *Then the same sequent is derivable with rank bounded by* $(3^a, r)$.

Corollary 3.6. *If the language of* S *does not contain* T *then* $\mathsf{CT}^*[\mathsf{S}]$ *is a conservative extension of* S.

Proof. Suppose $\Gamma \Rightarrow \Delta$ is a T-free sequent with a derivation in $\mathsf{CT}^*[\mathsf{S}]$ of rank (a, r). Iterating Theorem 3.5 yields a second derivation of the same sequent with rank $(a', 0)$ for some a'. This derivation is, by necessity, free of T-cuts whence the sub-formula property implies $a' = 0$, i.e. the derivation involves only axioms and basic rules, and so is derivable in S. ⊣

3.7. Halbach's proof and remaining obstacles. Halbach's Theorem 3.1 of [7] (see also Theorem 8.10 of [8]) aims to provide an argument for the elimination of T-cuts in $\mathsf{CT}[\mathsf{S}]$ derivations. The method assigns a measure, called T-*complexity*, to each T-cut and proves that the T-complexity of the bottom-most cut in a derivation can be reduced through local operations. The T-complexity of a cut is the number of T-rules applied to ancestors of cut formula on either side of the cut.

Consider the following derivation (the presentation of which has been intentionally simplified) where φ_t denotes $\varphi[x/t]$.

$$\dfrac{\dfrac{\dfrac{\vdots}{\Gamma, \mathrm{T}\varphi_r, \mathrm{T}\varphi_s \Rightarrow \Delta}}{\Gamma, \mathrm{T}\forall x\varphi, \mathrm{T}\varphi_s \Rightarrow \Delta}(\forall_{\mathrm{T}}\mathrm{L})}{\Gamma, \mathrm{T}\forall x\varphi \Rightarrow \Delta}(\forall_{\mathrm{T}}\mathrm{L}) \quad \dfrac{\dfrac{\vdots}{\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_{\dot{x}}}}{\Pi \Rightarrow \Sigma, \mathrm{T}\forall x\varphi}(\forall_{\mathrm{T}}\mathrm{R}) \over \Gamma, \Pi \Rightarrow \Delta, \Sigma}(\mathrm{Cut_T}) \tag{1}$$

The standard reduction process transforms the derivation into the following in which both cuts have lower T-complexity than in the derivation above.

$$\dfrac{\dfrac{\dfrac{\vdots}{\Gamma, \mathrm{T}\varphi_r, \mathrm{T}\varphi_s \Rightarrow \Delta} \quad \dfrac{\vdots}{\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_r}}{\Gamma, \Pi, \mathrm{T}\varphi_s \Rightarrow \Delta, \Sigma}(\mathrm{Cut_T}) \quad \dfrac{\vdots}{\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_s}}{\Gamma, \Pi \Rightarrow \Delta, \Sigma}(\mathrm{Cut_T}) \tag{2}$$

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

A problem arises, though, when one wishes to proceed further. There is nothing to stop reductions of the top cut (on $\mathrm{T}\varphi_r$) increasing the T-complexity of the bottom cut. Suppose, for example, $\Gamma = \emptyset$, $\mathrm{T}\varphi_r \in \Delta$, $\{\mathrm{T}\varphi_s, \varphi_s = (\psi_0 \mathbin{\underset{\cdot}{\wedge}} \psi_1)\} \subseteq \Pi$, the left-most sub-derivation is an axiom and $\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_{\dot{x}}$ is obtained by an application of $(\mathbin{\underset{\cdot}{\vee}}_{\mathrm{T}}\mathrm{L})$ to $\Pi, \mathrm{T}\psi_0 \Rightarrow \Sigma, \mathrm{T}\varphi_r$. Thus derivation (2) has the form

$$\dfrac{\dfrac{\begin{array}{c}\text{Axiom 2}\\ \mathrm{T}\varphi_r, \mathrm{T}\varphi_s \Rightarrow \Delta\end{array} \qquad \dfrac{\begin{array}{c}\vdots\\ \Pi, \mathrm{T}\psi_0 \Rightarrow \Sigma, \mathrm{T}\varphi_r\end{array}}{\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_r}\,(\vee_{\mathrm{T}}\mathrm{L})}{\Pi, \mathrm{T}\varphi_s \Rightarrow \Delta, \Sigma}\,(\mathrm{Cut}_{\mathrm{T}}) \qquad \dfrac{\begin{array}{c}\vdots\\ \Pi, \mathrm{T}\psi_0 \Rightarrow \Sigma, \mathrm{T}\varphi_s\end{array}}{\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_s}\,(\vee_{\mathrm{T}}\mathrm{L})}{\Pi \Rightarrow \Delta, \Sigma}\,(\mathrm{Cut}_{\mathrm{T}}) \tag{3}$$

The upper cut has a trivial reduction which turns (3) into

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \Pi, \mathrm{T}\psi_0 \Rightarrow \Sigma, \mathrm{T}\varphi_r\end{array}}{\Pi, \mathrm{T}\varphi_s \Rightarrow \Sigma}\,(\vee_{\mathrm{T}}\mathrm{L}) \qquad \dfrac{\begin{array}{c}\vdots\\ \Pi, \mathrm{T}\psi_0 \Rightarrow \Sigma, \mathrm{T}\varphi_s\end{array}}{\Pi \Rightarrow \Sigma, \mathrm{T}\varphi_s}\,(\vee_{\mathrm{T}}\mathrm{L})}{\Gamma \Rightarrow \Sigma}\,(\mathrm{Cut}_{\mathrm{T}}) \tag{4}$$

The remaining cut in (4), however, now has a higher T-complexity than in (3) and one cannot proceed further via an inductive argument. It is such a scenario that Halbach's argument does not cover. The critical question, therefore, is how to assign a *rank* to each of the two cuts in (2) that is i) strictly smaller than the rank of the cut in (1), and ii) is preserved when further reductions are made to (arbitrary) sub-derivations.

The example demonstrates that the rank associated to the cut on $\mathrm{T}\varphi_s$ in (2) should depend on the rank assigned to the cut on $\mathrm{T}\varphi_r$ and this dependence should be more significant than it's own T-complexity (which is, of course, still relevant). Thus, if there is an appropriate way to assign ranks to occurrences of the truth predicate so the natural reduction procedure can be proven to succeed, it will require a deep analysis of the derivation as a whole.

Consider, for example, a derivation with a subderivation of the form

$$\begin{array}{c}\dfrac{\begin{array}{c}\vdots\\ \Rightarrow \mathrm{T}r\end{array} \qquad \dfrac{\mathrm{T}r \Rightarrow \mathrm{T}r}{s = (r \mathbin{\underset{\cdot}{\vee}} r), \mathrm{T}r \Rightarrow \mathrm{T}s}\,(\vee_{\mathrm{T}}\mathrm{R})}{s = (r \mathbin{\underset{\cdot}{\vee}} r) \Rightarrow \mathrm{T}s}\,(\mathrm{Cut}_{\mathrm{T}})\\ \vdots\end{array}$$

The formula $\mathrm{T}s$ in the conclusion has T-complexity 1. However, interpreting the terms in the right sub-derivation it is clear that s should be viewed as a formula with logical depth one more than that of r. After eliminating the displayed cut, which is simply matter of applying the rule $(\vee_{\mathrm{T}}\mathrm{R})$ to the left sub-derivation, this measure of complexity remains unchanged. If cuts further down the derivation are assigned ranks derived from this measure, removal of the displayed cut will not alter ranks assigned to them either.

In the following we show how the observation above can be generalised in order to define a robust rank operation that permits standard cut elimination arguments.

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

§4. Approximations. Recall the language $\mathcal{L}^+$ which extends $\mathcal{L}$ by countably many fresh predicate symbols

$$\mathcal{P} = \{\mathsf{p}^i_j \mid i, j < \omega \text{ and } \mathsf{p}^i_j \text{ is a predicate symbol of arity } i\} \cup \{\varepsilon\},$$

where ε is a fresh propositional constant. The additional predicate symbols enable us to explicitly reduce the complexity of formulæ that occur under the truth predicate in CT-derivations. This is achieved by the use of *approximations*, an idea utilised by Kotlarski et al in [13].

The definitions and lemmas of Sections 4.1–3 are taken from [13] with only minor modification; in Section 4.4 we present their formal counterparts. The reader familiar with the concepts involved in [13] may wish to skip directly to Section 4.5 which contains the technical applications of these results to CT-derivations.

An *assignment* is any function $g\colon X \to \mathcal{L}^+$ such that $X \subseteq \mathcal{P}$ is a finite set and for every i, j, if $\mathsf{p}^i_j \in X$ then $g(\mathsf{p}^i_j)$ is a formula with at most variables $x_1, \dots, x_i$ occurring free. If $\varepsilon \in X$, $g(\varepsilon)$ may be an arbitrary formula. Given an assignment g and an $\mathcal{L}^+$ formula φ, we write $\varphi[g]$ for the result of replacing each occurrence of ε by the formula $g(\varepsilon)$ and each predicate $\mathsf{p}^i_j(s_1, \dots, s_i)$ occurring in φ by $g(\mathsf{p}^i_j)(s_1, \dots, s_i)$, if $g(\mathsf{p}^i_j)$ is defined, and ε otherwise. Note that some variables free in $g(\varepsilon)$ may become bound in this substitution. We write $\varphi[\psi]$ as shorthand for $\varphi[g_\psi]$ where $g_\psi\colon \{\varepsilon\} \to \{\psi\}$.

If $\boldsymbol{\varphi} = (\varphi_0, \dots, \varphi_m)$ and $\boldsymbol{\psi} = (\psi_0, \dots, \psi_m)$ are two sequences of closed $\mathcal{L}^+$ formulæ we say $\boldsymbol{\varphi}$ *approximates* $\boldsymbol{\psi}$ if there exists an assignment g such that $\psi_i = \varphi_i[g]$ for each $i \leq m$.

For a given sequence $\boldsymbol{\varphi}$ of $\mathcal{L}^+$, a collection of approximations to $\boldsymbol{\varphi}$ are distinguished. The *n-th approximation* of $\boldsymbol{\varphi}$, defined below, is a particular approximation to $\boldsymbol{\varphi}$ that has logical depth no more than $lh(\boldsymbol{\varphi}) \cdot 2^n$, where $lh(\boldsymbol{\varphi})$ denotes the length of the sequence $\boldsymbol{\varphi}$.

4.1. Occurrences. Let $w, z, z_1, z_2, \dots$ be fresh variable symbols. Given a formula φ of $\mathcal{L}$ we first define a formula $\bar{\varphi}$ of $\mathcal{L} \cup \{w\}$ in two steps: φ^* is the result of replacing in φ every free variable by w, and $\bar{\varphi}$ is obtained from φ^* by replacing each term in which the only variable that occurs is w, by w. Thus any term occurring in $\bar{\varphi}$ is either simply the variable w or contains a bound occurrence of a variable different from w.

For each formula φ, we let $O(\varphi)$ denote the set of *occurrences in* φ, pairs (ψ, s) such that ψ is a formula of $\mathcal{L} \cup \{w, z\}$ in which the variable z occurs exactly once, s is a term of $\mathcal{L} \cup \{w\}$ which is free for z in ψ and $\varphi = \psi[z/s]$ (the result of substituting in ψ all occurrences of z by s). Notice that if $(\psi, s) \in O(\bar{\varphi})$ then $s = w$.

The construction of $\bar{\varphi}$ and $O(\varphi)$ is such that for each formula φ of $\mathcal{L}$ there is a uniquely determined function $t_\varphi\colon O(\bar{\varphi}) \to \mathsf{Term}_{\mathcal{L}}$ for which φ is the result of replacing within $\bar{\varphi}$ each occurrence of the variable w by the appropriate value of t_φ. We call two formulæ φ, ψ *equivalent*, written $\varphi \sim \psi$, if $\bar{\varphi} = \bar{\psi}$.

LEMMA 4.1. *Let Φ be a set of $\mathcal{L}$ formulæ such that for every $\varphi, \psi \in \Phi$, $\varphi \sim \psi$. Then there exists a number l and formula $\vartheta(z_1, \dots, z_l)$, called a* template *of Φ, such that for every $\varphi \in \Phi$ there are terms $s_1, \dots, s_l$ so that $\varphi = \vartheta(s_1, \dots, s_l)$.*

PROOF. Suppose Φ is a set of formulæ satisfying the hypotheses of the lemma. As $O(\bar{\varphi}) = O(\bar{\psi})$ for every $\varphi, \psi \in \Phi$, $O(\Phi)$ has a natural definition as $O(\bar{\varphi})$ for

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

some $\varphi \in \Phi$. The functions $\{t_\varphi \mid \varphi \in \Phi\}$ induce an equivalence relation E_Φ on $O(\Phi)$ by setting

$$(\chi, s)\; E_\Phi\; (\psi, t) \iff \text{for every } \varphi \in \Phi,\ t_\varphi(\chi, s) = t_\varphi(\psi, t).$$

Let l be the number of E_Φ-equivalence classes in Φ. For each $\varphi \in \Phi$, the function t_φ is constant on $O(\Phi)/E_\Phi$, whence $\vartheta(z_1, \dots, z_l)$ is easily defined. ⊣

4.2. Parts. If $\boldsymbol{\varphi} = (\varphi_0, \varphi_1, \dots, \varphi_k)$ is a nonempty sequence of $\mathcal{L}$ formulæ, then the *set of parts of* $\boldsymbol{\varphi}$, $\Pi(\boldsymbol{\varphi})$, is the collection of pairs (ψ, χ) such that ψ is a formula of $\mathcal{L} \cup \{\varepsilon\}$ in which ε occurs exactly once, χ is a formula of $\mathcal{L}$ and for some $i \leq n$, $\varphi_i = \psi[\chi]$, the result of substituting χ for ε in ψ. Notice that $|\Pi(\boldsymbol{\varphi})| < k \cdot 2^{d(\boldsymbol{\varphi})}$ where $d(\boldsymbol{\varphi})$ denotes maximal logical depth of formulæ occurring in $\boldsymbol{\varphi}$ with atomic formulæ having depth 0.

We now define an ordering $\preceq$ on $\Pi(\boldsymbol{\varphi})$. Given pairs $(\varphi, \chi), (\varphi', \chi')$ of $\mathcal{L}^+$ formulæ, $(\varphi, \chi) \preceq (\varphi', \chi')$ iff there exists $\psi \in \mathcal{L}^+$ such that $\varphi'[\psi] = \varphi$ and $\psi[\chi] = \chi'$. Notice that in this case we also have $\varphi[\chi] = \varphi'[\chi']$.[2] Let $\prec$ be the irreflexive version of $\preceq$. The *depth* of a part $(\varphi, \chi) \in \Pi(\boldsymbol{\varphi})$, denoted $d(\varphi, \chi)$, is its (reverse) order-type in $\prec$, that is the number of logical connectives and quantifiers between φ and the occurrence of ε in φ. Making use of $\prec$ and $\sim$ the following sets can be defined.

$$\begin{aligned}
\Pi^0(\boldsymbol{\varphi}, n) &= \{(\varphi, \chi) \in \Pi(\boldsymbol{\varphi}) \mid d(\varphi, \chi) \leq n\} \\
\Pi^{m+1}(\boldsymbol{\varphi}, n) &= \{(\varphi, \chi) \in \Pi(\boldsymbol{\varphi}) \mid \exists(\varphi_0, \chi_0) \in \Pi^0(\boldsymbol{\varphi}, n)\, \exists(\varphi_1, \chi_1) \in \Pi^m(\boldsymbol{\varphi}, n) \\
&\qquad\qquad \wedge \chi_0 \sim \chi_1 \wedge (\varphi, \chi) \preceq (\varphi_1, \chi_1) \\
&\qquad\qquad \wedge d(\varphi, \chi) - d(\varphi_1, \chi_1) \leq n - d(\varphi_0, \chi_0)\} \\
\Pi(\boldsymbol{\varphi}, n) &= \bigcup_{m<\omega} \Pi^m(\boldsymbol{\varphi}, n).
\end{aligned}$$

The requirement '$\exists(\varphi_0, \chi_0) \in \Pi^0(\boldsymbol{\varphi}, n)$' serves only to ensure the set $\Pi^{m+1}(\boldsymbol{\varphi}, n)$ is bounded in size. Thus $\Pi^{m+1}(\boldsymbol{\varphi}, n)$ consists of those parts of $\boldsymbol{\varphi}$ that are approximated by some (φ_1, χ_1) in $\Pi^m(\boldsymbol{\varphi}, n)$ such that

i) a template for χ_1 appears somewhere in $\boldsymbol{\varphi}$ with depth at most n, and
ii) the depth of (φ, χ) is regulated by the position of this template in $\boldsymbol{\varphi}$ and the depth of (φ_1, χ_1).

The following two lemmas are consequences of the definition.

Lemma 4.2. *For every $(\varphi, \chi) \in \Pi(\boldsymbol{\varphi}, n)$ there exists $(\varphi', \chi') \in \Pi^0(\boldsymbol{\varphi}, n)$ with $\chi \sim \chi'$.*

Lemma 4.3. *For all $(\varphi_0, \chi_0), (\varphi_1, \chi_1), (\varphi_2, \chi_2) \in \Pi(\boldsymbol{\varphi}, n)$, if $(\varphi_0, \chi_0) \preceq (\varphi_1, \chi_1)$ and $\chi_1 \sim \chi_2$ then there exists $(\varphi_3, \chi_3) \in \Pi(\boldsymbol{\varphi}, n)$ such that $(\varphi_3, \chi_3) \preceq (\varphi_2, \chi_2)$ and $\chi_3 \sim \chi_0$.*

Example 4.4. Let θ be an atomic formula and set $\varphi^{k+1} = \theta \vee \varphi^k$ with $\varphi^0 = \theta$. We calculate $\Pi(\varphi^k, n)$ for $n < k$. Set $\psi^0 = \varepsilon$ and $\psi^{i+1} = \psi^i[\theta \vee \varepsilon]$. The set $\Pi^0(\varphi^k, n)$

[2] As such the definition of $\preceq$ presented here is equivalent to the relation denoted as $\prec$ in [13, p. 286] whose (precise) definition is '$(\varphi, \chi) \prec (\varphi', \chi')$ iff $(\varphi, \chi), (\varphi', \chi') \in \Pi(\sigma)$ for some (single) formula $\sigma \in \mathcal{L}$ and there exists $\psi \in \mathcal{L} \cup \{\varepsilon\}$ such that $\varphi'[\psi] = \varphi$.'

consists of all parts of φ^k with depth no greater than n:

$$\begin{array}{cccc} (\varepsilon, \varphi^k) & (\psi^1, \varphi^{k-1}) & \dots & (\psi^n, \varphi^{k-n}) \\ (\varepsilon \vee \varphi^{k-1}, \theta) & (\psi^1[\varepsilon \vee \varphi^{k-2}], \theta) & \dots & (\psi^{n-1}[\varepsilon \vee \varphi^{k-n}], \theta). \end{array}$$

In this example, in fact $\Pi(\varphi^k, n) = \Pi^0(\varphi^k, n)$. We show, for example, that $(\psi^l, \varphi^{k-l}) \in \Pi(\varphi^k, n)$ implies $(\psi^l, \varphi^{k-l}) \in \Pi^0(\varphi^k, n)$.

Fix $l \leq k$ and suppose $(\psi^l, \varphi^{k-l}) \in \Pi^{j+1}(\varphi^k, n)$. Let $(\varphi_1, \chi_1) \in \Pi^j(\varphi^k, n)$ and $(\varphi_0, \chi_0) \in \Pi^0(\varphi^k, n)$ the witnesses to this fact as given by the definition of $\Pi^{j+1}(\varphi^k, n)$:

$$(\psi^l, \varphi^{k-l}) \preceq (\varphi_1, \chi_1)$$
$$\chi_0 \sim \chi_1$$
$$d(\psi^l, \varphi^{k-l}) - d(\varphi_1, \chi_1) \leq n - d(\varphi_0, \chi_0).$$

From the first line it follows that $\varphi_1 = \psi^m$ and $\chi_1 = \varphi^{k-m}$ for some $m \leq l$. Moreover, $\chi_0 \sim \chi_1$ implies $(\varphi_0, \chi_0) = (\varphi_1, \chi_1)$, so

$$l = d(\psi^l, \varphi^{k-l}) \leq n - d(\varphi_0, \chi_0) + d(\varphi_1, \chi_1) = n.$$

EXAMPLE 4.5. Suppose $\psi = \forall x(\theta_0 \vee \exists y \neg \theta_1)$ where $\theta_0 \sim \theta_1$ are nonatomic formulæ possibly containing variables x and y free. Denote by ψ_0 and ψ_1 the formulæ $\forall x(\varepsilon \vee \exists y \neg \theta_1)$ and $\forall x(\theta_0 \vee \exists y \neg \varepsilon)$ respectively.

Consider $(\varphi_0, \chi_0) \in \Pi(\theta_0)$ such that $0 < d(\varphi_0, \chi_0) \leq 2$. Let $(\varphi_1, \chi_1) \in \Pi(\theta_1)$ be chosen such that $\chi_1 \sim \chi_0$ and $d(\varphi_i, \chi_i) = d(\varphi, \chi)$. Such formulæ must exist as θ_0 and θ_1 have the same 'logical' form. Notice $\psi = \psi_0[\theta_0] = \psi_1[\theta_1]$ (even if x or y appears free in θ_0 or θ_1). We have $(\psi_0, \theta_0), (\psi_1, \theta_1) \in \Pi^0(\psi, 4)$ and $(\psi_0[\varphi_0], \chi_0), (\psi_1[\varphi_1], \chi_1) \in \Pi(\psi)$ but only $(\psi_0[\varphi_0], \chi_0) \in \Pi^0(\psi, 4)$. Lemma 4.3 implies $(\psi_1[\varphi_1], \chi_1) \in \Pi(\psi, 4)$, but we can see this explicitly by observing

$$(\psi_0, \theta_0), (\psi_1, \theta_1) \in \Pi^0(\psi, 4),$$
$$\theta_0 \sim \theta_1,$$
$$(\psi_1[\varphi_1], \chi_1) \preceq (\psi_1, \theta_1),$$
$$d(\psi_1[\varphi_1], \chi_1) - d(\psi_1, \theta_1) = 4 - d(\psi_0, \theta_0),$$

whence $(\psi_1[\varphi_1], \chi_1) \in \Pi^1(\psi, 4)$ follows. Lemma 4.2 and 4.3 combine to imply

$$\begin{aligned} \Pi(\psi, 4) = &\left\{(\varepsilon, \psi), (\forall x\varepsilon, \theta_0 \vee \exists y \neg \theta_1), (\psi_0, \theta_0), (\forall x(\theta_0 \vee \varepsilon), \exists y \theta_1), (\psi_1, \theta_1)\right\} \\ &\cup \left\{(\psi_i[\varphi], \chi) \mid i < 2 \wedge (\varphi, \chi) \in \Pi^0(\theta_i, 2)\right\}. \end{aligned}$$

Notice the above is entirely independent of the actual arrangement of free variables in θ_0 and θ_1. For example if $\theta_0 = \exists z(x = t(r, z))$ and $\theta_1 = \exists z(y = t(s, z))$ then provided neither r nor s contains z we have $\theta_0 \sim \theta_1$.

EXAMPLE 4.6. Let φ^k and ψ^k be as in Example 4.4 and let $\boldsymbol{\varphi}$ be the sequence $(\neg \varphi^{k+1}, \exists x \varphi^k, \varphi^k)$. Although we have

$$\Pi^0(\boldsymbol{\varphi}, n) = \Pi^0(\neg \varphi^{k+1}, n) \cup \Pi^0(\exists x \varphi^k, n) \cup \Pi^0(\varphi^k, n)$$

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

it is not the case that $\Pi(\varphi, n) = \Pi(\neg\varphi^{k+1}, n) \cup \Pi(\exists x\varphi^k, n) \cup \Pi(\varphi^k, n)$ unless $n \geq k+2$. Suppose $k \geq 3$; we determine $\Pi(\varphi, 3)$. We have, for example

$$\begin{aligned}
\Pi^0(\neg\varphi^{k+1}, 3) &= \left\{(\varepsilon, \neg\varphi^{k+1}), (\neg\varepsilon, \varphi^{k+1}), (\neg(\theta \vee \varepsilon), \varphi^k), (\neg\psi^2, \varphi^{k-1})\right\}\\
&\quad \cup \left\{(\neg(\varepsilon \vee \varphi^k), \theta), (\neg(\theta \vee (\varepsilon \vee \varphi^{k-1})), \theta)\right\}\\
\Pi^0(\exists x\varphi^k, 3) &= \left\{(\varepsilon, \exists x\varphi^k), (\exists x\varepsilon, \varphi^k), (\exists x\psi^1, \varphi^{k-1}), (\exists x\psi^2, \varphi^{k-2})\right\}\\
&\quad \cup \left\{(\exists x(\varepsilon \vee \varphi^{k-1}), \theta), (\exists x(\theta \vee (\varepsilon \vee \varphi^{k-2})), \theta)\right\}\\
\Pi^0(\varphi^k, 3) &= \left\{(\varepsilon, \varphi^k), (\psi^1, \varphi^{k-1}), (\psi^2, \varphi^{k-2}), (\psi^3, \varphi^{k-3})\right\}\\
&\quad \cup \left\{(\varepsilon \vee \varphi^{k-1}, \theta), (\theta \vee (\varepsilon \vee \varphi^{k-2}), \theta), (\psi^2[\varepsilon \vee \varphi^{k-3}], \theta)\right\}.
\end{aligned}$$

Notice that (ψ^3, φ^{k-3}) is an element of $\Pi^0(\varphi, 3)$ whereas (ψ^4, φ^{k-3}) and $(\exists x\psi^3, \varphi^{k-3})$ are not. Consider the parts (ψ^1, φ^{k-1}) and $(\exists x\psi^1, \varphi^{k-1})$ from $\Pi^0(\varphi, 3)$. We have $(\exists x\psi^3, \varphi^{k-3}) \preceq (\exists x\psi^1, \varphi^{k-1})$ and

$$d(\exists x\psi^3, \varphi^{k-3}) - d(\exists x\psi^1, \varphi^{k-1}) = 2 = 3 - d(\psi^1, \varphi^{k-1}),$$

whence $(\exists x\psi^3, \varphi^{k-3}) \in \Pi^1(\varphi, 3)$. Applying Lemma 4.3 we deduce that $(\neg\psi^4, \varphi^{k-3})$ and $(\neg\psi^3, \varphi^{k-2})$ are elements of $\Pi(\varphi, 3)$ (the former entering first at $\Pi^2(\varphi, 3)$). Thus

$$\Pi(\varphi, 3) \supseteq \Pi^0(\varphi, 3) \cup \left\{(\theta \vee \psi, \chi), (\neg(\theta \vee \psi), \chi) \mid (\psi, \chi) \in \Pi^0(\varphi^k, 3)\right\}.$$

Lemma 4.2 (and the fact that $\Pi(\varphi, 3)$ is closed upwards in $\prec$) implies the above inclusion is indeed equality.

We finish the section with an important consequence of Lemma 4.2.

Lemma 4.7. *$\Pi(\varphi, n)$ is a finite set. Indeed, $|\Pi(\varphi, n)| \leq 2^{lh(\varphi)\cdot 2^n}$.*

Proof. Lemma 4.2 implies that if

$$\Pi(\varphi, n) \ni (\varphi_0, \chi_0) \prec (\varphi_1, \chi_1) \prec \cdots \prec (\varphi_k, \chi_k)$$

is a sequence of parts of φ increasing with respect to $\prec$ then indeed $k \leq |\Pi^0(\varphi, n)| \leq lh(\varphi) \cdot 2^n$, from which we deduce that $\Pi(\varphi, n)$ is a subset of $\Pi^0(\varphi, lh(\varphi) \cdot 2^n)$ and $|\Pi(\varphi, n)| \leq 2^{lh(\varphi)\cdot 2^n}$. ⊣

4.3. Approximating formulæ. Given $\Pi(\varphi, n)$, two further sets can be defined:

$$\begin{aligned}
\Gamma(\varphi, n) &= \{\psi \in \mathcal{L} \mid \exists\varphi\, (\varphi, \psi) \in \Pi(\varphi, n)\},\\
\Gamma_I(\varphi, n) &= \{\psi \in \mathcal{L} \mid \exists\varphi\, (\varphi, \psi) \text{ is } \prec\text{-minimal in } \Pi(\varphi, n)\}.
\end{aligned}$$

We now define a function $F_{\varphi,n} \colon \Gamma(\varphi, n) \to \mathcal{L}^+$ by recursion through $\prec$ that generates the particular approximations we require.

Fix an enumeration $\Phi_0, \dots, \Phi_n$ of the $\sim$-equivalence classes of $\Gamma_I(\varphi, n)$ and for each $j \leq n$ let $\vartheta_j(z_1, \dots, z_{a_j})$ be a template for Φ_j and let $t_1^\psi, \dots, t_{a_j}^\psi$ denote the terms for which $\Phi_j \ni \psi = \vartheta_j(t_1^\psi, \dots, t_{a_j}^\psi)$. We begin by defining $F_{\varphi,n}$ on $\Gamma_I(\varphi, n)$. If $\psi \in \Gamma_I(\varphi, n)$ is atomic set $F_{\varphi,n}(\psi) = \psi$, otherwise set $F_{\varphi,n}(\psi) = \mathsf{p}_j^{a_j}(t_1^\psi, \dots, t_{a_j}^\psi)$ where $\psi \in \Phi_j$. In the remaining cases $F_{\varphi,n}(\psi)$ is defined to commute with the external connective or quantifier in ψ.

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

The *n-th approximation* of $\varphi = (\varphi_0, \dots, \varphi_k)$ is chosen to be the sequence

$$F_{\varphi,n}(\varphi) = (F_{\varphi,n}(\varphi_0), \dots, F_{\varphi,n}(\varphi_k)).$$

EXAMPLE 4.8. Let $\varphi^k = (\neg\varphi^{k+1}, \exists x\varphi^k, \varphi^k[x/y])$ where φ^k is as defined in Example 4.4, $k > 3$ and assume x appears free in θ. In this example we calculate $F_{\varphi^k,3}(\varphi^k)$. Let ψ_y abbreviate $\psi[x/y]$. Since $\psi \sim \psi_y$ for every ψ, $\Pi(\varphi^k, 3)$ is straightforward to calculate given $\Pi(\varphi, 3)$ in Example 4.6 and we deduce

$$\begin{aligned}\Gamma(\varphi^k, 3) &= \{\neg\varphi^{k+1}, \varphi^{k+1}, \exists x\varphi^k, \varphi^k, \varphi^{k-1}, \varphi^{k-2}, \varphi^{k-3}, \theta\} \\ &\quad \cup \{\varphi^k_y, \varphi^{k-1}_y, \varphi^{k-2}_y, \varphi^{k-3}_y, \theta_y\} \\ \Gamma_I(\varphi^k, 3) &= \{\varphi^{k-3}, \varphi^{k-3}_y, \theta, \theta_y\}.\end{aligned}$$

The $\sim$-equivalence classes of $\Gamma_I(\varphi^k, 3)$ are therefore $\{\varphi^{k-3}, \varphi^{k-3}_y\}$ and $\{\theta, \theta_y\}$. The definition of $F_{\varphi^k,3}$ yields

$$\begin{aligned}F_{\varphi^k,3}(\varphi^{k-3}) &= \mathsf{p}(x) \qquad & F_{\varphi^k,3}(\theta) &= \theta \\ F_{\varphi^k,3}(\varphi^{k-3}_y) &= \mathsf{p}(y) \qquad & F_{\varphi^k,3}(\theta_y) &= \theta_y\end{aligned}$$

for $\mathsf{p}(x)$ a fresh unary predicate symbol. Notice that if x did not occur in φ^k, $\mathsf{p}(x)$ and $\mathsf{p}(y)$ above would be replaced by a propositional constant and if $k = 3$, $F_{\varphi^k,3}(\varphi^{k-3}) = \theta$. Thus

$$F_{\varphi^k,3}(\varphi^k) = \left(\neg\psi^4[\varepsilon/\mathsf{p}(x)], \exists x\psi^3[\varepsilon/\mathsf{p}(x)], \psi^3[\varepsilon/\mathsf{p}(y)]\right)$$

for every $k > 3$.

The following simple lemmas establish the main properties of the approximations.

LEMMA 4.9. *For every $\psi_0, \psi_1 \in \Gamma(\varphi, n)$, if $F_{\varphi,n}(\psi_0) \sim F_{\varphi,n}(\psi_1)$ if and only if $\psi_0 \sim \psi_1$.*

LEMMA 4.10. *The i-th approximation of φ is an approximation to φ and an approximation to the j-th approximation whenever $i \leq j$.*

LEMMA 4.11. *Every occurrence of a predicate symbol from $\mathcal{P}$ in the n-th approximation of φ has depth at least n in φ. Moreover, every formula in the n-th approximation of φ has logical depth no greater than $lh(\varphi) \cdot 2^n$.*

LEMMA 4.12. *If $\varphi' = (\varphi'_0, \dots, \varphi'_k)$ is an approximation to φ and for each $i \leq n$, φ'_i has logical depth at most n, then φ' is an approximation to the n-th approximation of φ.*

The upper bound of Lemma 4.11 holds on account of Lemma 4.7. A consequence of the previous lemmas is the following.

LEMMA 4.13. *For all sequences φ, φ', formulæ ψ_0, ψ_1, ψ'_0, ψ_1, and $m < n$,*

1. *If $(\varphi', \psi'_0 \vee \psi'_1)$ is the n-th approximation of $(\varphi, \psi_0 \vee \psi_1)$ then the m-th approximation of (φ, ψ_i) is an approximation to (φ', ψ'_i).*
2. *If $(\varphi', \neg\psi'_0)$ is the n-th approximation of $(\varphi, \neg\psi_0)$ then the m-th approximation of (φ, ψ_0) is an approximation to (φ', ψ'_0).*
3. *If $(\varphi', \forall x\psi'_0)$ is the n-th approximation of $(\varphi, \forall x\psi_0)$ then for every $a < \omega$ the m-th approximation of $(\varphi, \psi_0[x/\bar{a}])$ is an approximation to $(\varphi', \psi'_0[x/\bar{a}])$.*

4.4. Approximating sequents. We begin by noting that all the definitions and results of the previous section can be formalised and proved within $\mathsf{I}\Delta_0 + \mathsf{exp}$. Thus we fix the following formal notation.

1. $s[g] = t$ expresses that either g is not an assignment and $s = t$ or g is an assignment and t is the result of replacing within the $\mathcal{L}^+$ formula s, each occurrence of the predicate symbol p^i_j by $g(\mathsf{p}^i_j)$ if defined, otherwise by ε.
2. For $\varphi = (\varphi_0, \dots, \varphi_m)$ we set $\ulcorner \varphi \urcorner = (\ulcorner \varphi_0 \urcorner, \dots, \ulcorner \varphi_m \urcorner)$.
3. $\underset{\cdot}{F}_{r,k}(s) = t$ expresses that there exists a sequence φ and $\psi \in \Gamma(\varphi, k)$ such that $r = \ulcorner \varphi \urcorner$, $s = \ulcorner \psi \urcorner$ and $t = \ulcorner F_{\varphi,k}(\psi) \urcorner$; if there is no sequence of $\mathcal{L}$-formulæ φ such that $r = \ulcorner \varphi \urcorner$ then $s = t$.
4. For $\boldsymbol{s} = (s_0, \dots, s_m)$ and $\boldsymbol{t} = (t_0, \dots, t_m)$ (external) sequences of terms of the same length we introduce
   (a) $\boldsymbol{s} = \boldsymbol{t}$ to abbreviate $\bigwedge_{i \le m}(s_i = t_i)$;
   (b) $\boldsymbol{s}[g]$ to abbreviate the sequence of terms $(s_0[g], \dots, s_m[g])$;
   (c) $\underset{\cdot}{F}_{r,u}(\boldsymbol{s})$ to abbreviate the sequence $(\underset{\cdot}{F}_{r,u}(s_0), \dots, \underset{\cdot}{F}_{r,u}(s_m))$;
   (d) $\underset{\cdot}{d}(\boldsymbol{s}) \le u$ to abbreviate the formula $\bigwedge_{i \le m} \underset{\cdot}{d}(s_i) \le u$.
5. A Δ_0 predicate $\mathsf{Seq}(x)$ expressing that x encodes a sequence and additional terms and terms:
   (a) $lh(r)$ denoting the length of the sequence encoded by r;
   (b) $(r)_i$ denoting the i-th element of the sequence r;
   (c) $r \frown s$ denoting $((r)_0, \dots, (r)_{lh(r)-1}, (s)_0, \dots, (s)_{lh(s)-1})$ of length $lh(r) + lh(s)$.

The above notation is expanded to cover complex expressions involving sequences. For example, $\underset{\cdot}{F}_{r,u}(\boldsymbol{s})[g] = \underset{\cdot}{F}_{r',u'}(\boldsymbol{t})$ is shorthand for the formula $\bigwedge_{i \le m} \underset{\cdot}{F}_{r,u}(s_i)[g] = \underset{\cdot}{F}_{r',u'}(t_i)$.

Collecting together the results of the previous section we obtain

LEMMA 4.14. *The following sequents are* (*truth-free*) *derivable in* $\mathsf{I}\Delta_0 + \mathsf{exp}$.

$$\begin{aligned}
\emptyset &\Rightarrow (x \mathbin{\underset{\cdot}{\vee}} y)[z] = (x[z] \mathbin{\underset{\cdot}{\vee}} y[z]),\\
\emptyset &\Rightarrow (\underset{\cdot}{\neg} x)[z] = \underset{\cdot}{\neg}(x[z]),\\
\emptyset &\Rightarrow (\underset{\cdot}{\forall} xy)[z] = \underset{\cdot}{\forall} x(y[z]),\\
\emptyset &\Rightarrow (y(x/w))[z] = (y[z])(x/w),\\
\mathsf{Seq}(x), (x)_i = y \mathbin{\underset{\cdot}{\vee}} z &\Rightarrow \underset{\cdot}{F}_{x,w+1}(y \mathbin{\underset{\cdot}{\vee}} z) = \underset{\cdot}{F}_{x,w+1}(y) \mathbin{\underset{\cdot}{\vee}} \underset{\cdot}{F}_{x,w+1}(z),\\
\mathsf{Seq}(x), (x)_i = \underset{\cdot}{\neg} y &\Rightarrow \underset{\cdot}{F}_{x,w+1}(\underset{\cdot}{\neg} y) = \underset{\cdot}{\neg} \underset{\cdot}{F}_{x,w+1}(y),\\
\mathsf{Seq}(x), (x)_i = \underset{\cdot}{\forall} yz &\Rightarrow \underset{\cdot}{F}_{x,w+1}(\underset{\cdot}{\forall} yz) = \underset{\cdot}{\forall} y(\underset{\cdot}{F}_{x,w+1}(z)),\\
\mathsf{Seq}(y_0 \frown y_1 \frown y_2) &\Rightarrow \underset{\cdot}{F}_{x,w}(y_0 \frown y_1 \frown y_2) = \underset{\cdot}{F}_{x,w}(y_0 \frown y_1 \frown y_2),\\
\emptyset &\Rightarrow \underset{\cdot}{d}(\underset{\cdot}{F}_{x,z}(\boldsymbol{s})) \le \underset{\cdot}{lh}(x) \cdot 2^z.
\end{aligned}$$

LEMMA 4.15. *There is a term* $g(w, x, y, z)$ *such that the following sequents are truth-free derivable in* $\mathsf{I}\Delta_0 + \mathsf{exp}$.

$$\begin{aligned}
\emptyset &\Rightarrow \underset{\cdot}{d}(g) \le \underset{\cdot}{lh}(x) \cdot 2^z,\\
y < z, w = x &\Rightarrow \underset{\cdot}{F}_{w,y}(u)[g] = \underset{\cdot}{F}_{x,z}(u),\\
y < z, x = x' \frown (x_0 \mathbin{\underset{\cdot}{\vee}} x_1), w = x' \frown x_i &\Rightarrow \underset{\cdot}{F}_{w,y}(w)[g] = \underset{\cdot}{F}_{x,z}(w),\\
y < z, x = x' \frown (\underset{\cdot}{\neg} x_0), w = x' \frown x_0 &\Rightarrow \underset{\cdot}{F}_{w,y}(w)[g] = \underset{\cdot}{F}_{x,z}(w),
\end{aligned}$$

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

$$y < z, x = x'^\frown(\dot\forall x_0 x_1), w = x'^\frown subn(x_1, x_2, u) \Rightarrow \underset{\cdot}{F}_{w,y}(w)[g] = \underset{\cdot}{F}_{x,z}(x')^\frown subn(\underset{\cdot}{F}_{x,z}(x_2), x_1, u),$$

$$w = x^\frown w', \forall u(\underset{\cdot}{d}(\underset{\cdot}{F}_{w,y}(u)) \leq z) \Rightarrow \underset{\cdot}{F}_{w,y}(x)[g] = \underset{\cdot}{F}_{x,z}(x).$$

The first sequent of Lemma 4.15 formalises Lemma 4.11, the second Lemma 4.10, the third, fourth, and fifth Lemma 4.13 (the last of these expressing that the y-th approximation of $(\varphi, \varphi[a/x_2])$ can be viewed as an approximation to the z-th approximation of $(\varphi, \forall x\varphi)$ whenever $y < z$), and the final line combines Lemmas 4.12 and 4.11.

Tying in approximations with derivations we have:

Lemma 4.16. *Let Γ, Δ be sets consisting of* T*-free formulæ, and $\boldsymbol{\varphi}$, $\boldsymbol{\psi}$ be sequences of terms. If $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$ is derivable with truth depth a then for every term g,*

$$\Gamma, \mathrm{T}\boldsymbol{\varphi}[g] \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}[g]$$

is derivable with truth depth $\leq a$. Moreover, if the first derivation contains no T*-cuts, neither does the second.*

The lemma is not difficult to prove. We require, however, a more general version from which we may infer bounds on the truth rank of the resulting derivation. The next lemma achieves this.

Lemma 4.17. *Let Γ, Δ, $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ be as in the statement of the previous lemma. If the sequents $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$, and $\Gamma \Rightarrow \underset{\cdot}{d}(g) < \bar{k}$ are derivable with rank (a, r) and $(0, 0)$ respectively, the sequent*

$$\Gamma, \mathrm{T}\boldsymbol{\varphi}[g] \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}[g]$$

is derivable with rank bounded by $(a, r + k)$.

Proof. The only nontrivial case is if the last rule is $(\mathrm{Cut}^l_\mathrm{T})$ for some $l < r$. So suppose $a = a' + 1$ and we have the following derivation

$$\frac{\Gamma, \mathrm{T}\boldsymbol{\varphi}, \mathrm{T}\chi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi} \qquad \Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\chi, \mathrm{T}\boldsymbol{\psi} \qquad \Gamma \Rightarrow \underset{\cdot}{d}(\chi) \leq \bar{l}}{\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}} \,(\mathrm{Cut}^l_\mathrm{T})$$

with the two left-most premises derivable with rank (a', r) and the right-most with rank $(0, 0)$. By the induction hypothesis, the sequents

$$\Gamma, \mathrm{T}\boldsymbol{\varphi}[g], \mathrm{T}\chi[g] \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}[g] \qquad \Gamma, \mathrm{T}\boldsymbol{\varphi}[g] \Rightarrow \Delta, \mathrm{T}\chi[g], \mathrm{T}\boldsymbol{\psi}[g]$$

are both derivable with rank bounded by $(a', r+k)$. Since the sequent $\Gamma \Rightarrow \underset{\cdot}{d}(g) \leq \bar{k}$ is truth-free derivable, so is

$$\Gamma \Rightarrow \underset{\cdot}{d}(\chi[g]) \leq \bar{l} + \bar{k},$$

whence the rule $(\mathrm{Cut}^{l+k}_\mathrm{T})$ yields the desired sequent. ⊣

4.5. Approximating derivations. Given a sequent $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$, we define its *$k$-th approximation* to be the sequent

$$\Gamma, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\varphi}^\frown\boldsymbol{\psi}, \bar{k}}\boldsymbol{\varphi}) \Rightarrow \Delta, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\varphi}^\frown\boldsymbol{\psi}, \bar{k}}\boldsymbol{\psi}).$$

Let H be the function given by

$$H(k, n) = n \cdot 2^k.$$

By Lemma 4.11 each member of the k-th approximation of φ has depth at most $H(k, lh(\varphi))$.

The following lemmas hold for arbitrary derivations in $\mathsf{CT}^*[\mathsf{S}]$.

LEMMA 4.18. *Suppose $a, r, m, n, k < \omega$, Γ, and Δ are finite sets of $\mathcal{L}$-formulæ, $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ are sequences of terms, ψ, s, and t are terms, and $lh(\boldsymbol{\varphi}) + lh(\boldsymbol{\psi}) = n$. If the k-th approximation to the sequent $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}(\psi[s/t])$ is derivable with rank (a, r) then there is a derivation with rank bounded by $(a+1, r+H(k+1, n+1))$ of the $(k+1)$-th approximation to $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}(\underset{\cdot}{\forall} s\psi)$.*

PROOF. Let $\boldsymbol{\chi} = \boldsymbol{\varphi}^\frown\boldsymbol{\psi}^\frown(\psi[s/t])$. Assume the sequent

$$\Gamma, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi},\bar{k}}\boldsymbol{\varphi}) \Rightarrow \Delta, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi},\bar{k}}\boldsymbol{\psi}), \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi},\bar{k}}(\psi[s/t]))$$

is derivable with rank (a, r). Let $g(x, y, z)$ be the term given by Lemma 4.15 and set $g' = g(\boldsymbol{\chi}, \bar{k}, \bar{k}+1)$. Lemma 4.17 implies there is a derivation with rank bounded $(a, r+H(k+1, n+1))$ of the sequent

$$\Gamma, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}\boldsymbol{\varphi}) \Rightarrow \Delta, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}\boldsymbol{\psi}), \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi},\bar{k}}(\psi[s/t])[g']),$$

where $\boldsymbol{\chi}' = \boldsymbol{\varphi}^\frown\boldsymbol{\psi}^\frown(\underset{\cdot}{\forall} s\psi)$. Combining this derivation with those of Lemma 4.14 and the penultimate sequent in 4.15 and using only T-free cuts, yields a derivation of the sequent

$$\Gamma, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}\boldsymbol{\varphi}) \Rightarrow \Delta, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}\boldsymbol{\psi}), \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}(\psi)[s/t])$$

with rank $(a, r+H(k+1, n+1))$, whence $(\forall_{\mathrm{T}}\mathrm{R})$ and Lemma 4.14 yield a derivation of

$$\emptyset \Rightarrow \Delta, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}\boldsymbol{\psi}), \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\chi}',\bar{k}+1}(\forall s\psi))$$

with rank bounded by $(a+1, r+H(k+1, n+1))$. ⊣

LEMMA 4.19. *If the k-th approximation to $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}\psi_i$ is derivable with rank (a, r) then the $(k+1)$-th approximation to $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}(\psi_0 \underset{\cdot}{\vee} \psi_1)$ is derivable with rank $(a+1, r+H(k+1, n))$, where $n = lh(\boldsymbol{\varphi}) + lh(\boldsymbol{\psi}) + 1$.*

LEMMA 4.20. *If the k-th approximation to $\Gamma, \mathrm{T}\boldsymbol{\varphi}, \mathrm{T}\psi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$ is derivable with rank (a, r) then the $(k+1)$-th approximation to the sequent $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}(\underset{\cdot}{\neg}\psi)$ is derivable with rank $(a+1, r+H(k+1, n))$, where $n = lh(\boldsymbol{\varphi}) + lh(\boldsymbol{\psi}) + 1$.*

Additionally, the variations of the above relating to the rules $(\forall_{\mathrm{T}}\mathrm{L})$, $(\vee_{\mathrm{T}}\mathrm{L})$, and $(\neg_{\mathrm{T}}\mathrm{L})$ also hold, though we omit them here.

LEMMA 4.21. *Let $n = lh(\boldsymbol{\varphi}) + lh(\boldsymbol{\psi})$ and suppose $r \leq H(k, n+1)$. If the k-th approximation to the sequents $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}\chi$, and $\Gamma, \mathrm{T}\boldsymbol{\varphi}, \mathrm{T}\chi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$ are derivable with rank (a, r) then the $H(k, n+1)$-th approximation to $\Gamma, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$ is derivable with rank bounded by*

$$(a+1, H(k, n+1) + H(H(k, n+1), n))).$$

PROOF. Let $N = H(k, n+1) \geq r$, $\boldsymbol{\omega} = \boldsymbol{\varphi}^\frown\boldsymbol{\psi}$, $\boldsymbol{\omega}' = \boldsymbol{\omega}^\frown\chi$. By Lemma 4.14 there is a truth-free derivation of $\emptyset \Rightarrow \underset{\cdot}{d}(\underset{\cdot}{F}_{\boldsymbol{\omega}',\bar{k}}(x)) \leq \bar{N}$, so by an application of $(\mathrm{Cut}^N_{\mathrm{T}})$ to the two derivations in the statement of the lemma yields a derivation of

$$\Gamma, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\omega}',\bar{k}}\boldsymbol{\varphi}) \Rightarrow \Delta, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\omega}',\bar{k}}\boldsymbol{\psi}) \tag{5}$$

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

rank $(a+1, r, N)$. Let g be the term given by Lemma 4.15 and set $g' = g(\boldsymbol{\omega}', \boldsymbol{\omega}, \bar{k}, \bar{N})$. The second line of Lemma 4.15 induces a truth-free derivation of

$$\emptyset \Rightarrow (F_{\boldsymbol{\omega}', \bar{k}} u)[g'] = F_{\boldsymbol{\omega}, \bar{N}} u,$$

whence we apply Lemma 4.17 to (5) to obtain a derivation of

$$\Gamma, \mathrm{T}(F_{\boldsymbol{\omega}, \bar{N}} \varphi) \Rightarrow \Delta, \mathrm{T}(F_{\boldsymbol{\omega}, \bar{N}} \boldsymbol{\psi})$$

with rank bounded by $(a+1, N + H(N, n))$. ⊣

§5. Proofs of the main theorems. We now have all the ingredients to establish the Bounding Lemma described in Section 1 that provides an interpretation of CT[S] in CT*[S]. The next lemma is a generalisation incorporating all the relevant bounds.

Lemma 5.1 (Bounding lemma). *There are elementary functions G_1, G_2 such that for every $a, n < \omega$, if $lh(\varphi) + lh(\boldsymbol{\psi}) \leq n$ and the sequent $\Gamma, \mathrm{T}\varphi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$ is derivable in* CT[S] *with truth depth a, then its $G_1(a, n)$-th approximation is derivable in* CT*[S] *with rank bounded by $(a, G_2(a, n))$.*

Proof. The idea is to copy the CT[S] derivation into CT*[S] replacing the rule $(\mathrm{Cut_T})$ by $(\mathrm{Cut}^k_{\mathrm{T}})$ for k determined inductively. The functions G_1 and G_2 are defined according to the bounds obtained in the previous section:

$$\begin{aligned} G_1(0, r) &= 0, \\ G_1(a+1, r) &= H(G_1(a, r+1), r+1), \\ G_2(a, r) &= G_1(a+1, a+r). \end{aligned}$$

We argue by induction on a. Suppose the last rule applied to obtain $\Gamma, \mathrm{T}\varphi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$ is a T-cut on $\mathrm{T}\chi$ and that this derivation has height $a+1$. Let $\boldsymbol{\omega} = \varphi{}^\frown \boldsymbol{\psi}$ and $\boldsymbol{\omega}' = \boldsymbol{\omega}{}^\frown \chi$. The induction hypothesis implies that the $G_1(a, n+1)$-th approximations to

$$\Gamma, \mathrm{T}\varphi, \mathrm{T}\chi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi} \qquad\qquad \Gamma, \mathrm{T}\varphi \Rightarrow \Delta, \mathrm{T}\chi, \mathrm{T}\boldsymbol{\psi}$$

are each derivable in CT*[S] with ranks bounded by $(a, G_2(a, n+1))$. By Lemma 4.21 there is a derivation with height $a+1$ of the $G_1(a+1, n)$-th approximation to $\Gamma, \mathrm{T}\varphi \Rightarrow \Delta, \mathrm{T}\boldsymbol{\psi}$. This derivation has truth rank bounded by $G_2(a+1, n)$ so we are done. The other cases are similar and follow from applications of Lemmas 4.18 and 4.19. ⊣

A combination of Lemmas 4.16 and 5.1 implies that CT[S] permits the elimination of all T-cuts.

Corollary 5.2. *If $\Gamma \Rightarrow \Delta$ is derivable in* CT[S] *then it is derivable without* T*-cuts.*

Theorem 5.3. *Let* S *be an elementary axiomatised theory in a finite language interpreting* $\mathsf{I\Delta_0} + \mathsf{exp}$. CT[S] *conservatively extends* S. *Moreover, this fact is verifiable in* $\mathsf{I\Delta_0} + \mathsf{exp}_1$.

Proof. Let φ be an T-free theorem of CT[S]. By the Embedding Lemma, the sequent $\emptyset \Rightarrow \varphi$ has a derivation within CT[S]. Lemma 5.1 implies that the same sequent is derivable in CT*[S] and the cut elimination theorem for CT*[S] shows $\emptyset \Rightarrow \varphi$ is derivable without truth cuts. But this derivation is also a derivation within S. Notice that this final derivation has height bounded by $2^a_{2 \cdot G_1(a+1, a+1)}$, where a bounds the height of the original derivation of $\emptyset \Rightarrow \varphi$ in $\mathsf{CT[S]}$, G_1 is as defined in the proof of the Bounding Lemma, and 2^n_m represents the function of hyper-exponentiation: $2^n_0 = 2^n$ and $2^n_{m+1} = 2^{2^n_m}$. Thus this reduction can be formalised within $\mathsf{I\Delta_0 + exp_1}$. ⊣

Theorem 5.4. *Let* S *be an elementary* $\mathcal{L}$*-theory in a finite language interpreting* $\mathsf{I\Delta_0 + exp}$. *For any* S*-schema* D, *the theory* $\mathsf{CT[S]} + \forall x(\mathsf{D}x \to \mathrm{T}x)$ *is a conservative extension of* S. *Moreover, this fact is verifiable in* $\mathsf{I\Delta_0 + exp_1}$.

Proof. Let S and D be as given in the statement of the theorem and let U be the finite set of $\mathcal{L} \cup \{p\}$ formulæ associated with the S-schema D. First we note that the additional axiom can formulated as the sequent rule

$$\frac{\Pi \Rightarrow \Sigma, \mathsf{D}s}{\Pi \Rightarrow \Sigma, \mathrm{T}s}\ (\mathsf{D}).$$

Suppose d is a derivation with truth depth a of the truth-free sequent $\Gamma \Rightarrow \Delta$ in the expansion of $\mathsf{CT[S]}$ by the rule (D). By redefining the functions G_1 and G_2 so that $G_1(0, n)$ bounds the logical depth of the (finitely many) formulæ in U for each n, the proof of the Bounding Lemma can be carried through to obtain a derivation with rank $(a, G_2(a, 0))$ of the same sequent in the system $\mathsf{CT^*[S]}$ expanded by a bounded version of (D):

$$\frac{\Pi, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Sigma, \mathrm{T}\boldsymbol{\psi}, \mathsf{D}\sigma}{\Pi, \mathrm{T}\boldsymbol{\varphi} \Rightarrow \Sigma, \mathrm{T}\boldsymbol{\psi}, \mathrm{T}(\underset{\cdot}{F}_{\boldsymbol{\omega}, \bar{k}}\sigma)}\ (\mathsf{D}_{\boldsymbol{\omega}})$$

where Π and Σ are truth-free, $k = G_1(a, 0)$ and $\boldsymbol{\omega} = \boldsymbol{\varphi}^\frown\boldsymbol{\psi}^\frown\sigma$.

Let d^* denote this derivation. Fix n such that for each instance of $(\mathsf{D}_{\boldsymbol{\omega}})$ occurring in d^*, $lh(\boldsymbol{\omega}) < n$, and set U^+ to be the finite set of instantiations of formulæ from U by $\mathcal{L}$-formulæ that have logical depth at most $G_2(a, n)$. It follows that the sequent $\mathsf{D}x, \underset{\cdot}{d}(x) < \overline{G_2(a, n)} \Rightarrow \{x = \ulcorner\varphi\urcorner \mid \varphi \in U^+\}$ is derivable in S. Because the sequent $\sigma \Rightarrow \mathrm{T}\ulcorner\sigma\urcorner$ is derivable in $\mathsf{CT[S]}$ for each $\mathcal{L}$-sentence σ we may deduce

$$\mathsf{D}x \Rightarrow \mathrm{T}(\underset{\cdot}{F}_{\ulcorner\boldsymbol{\omega}\urcorner, \bar{k}}x)$$

is derivable in $\mathsf{CT[S]}$ whenever $lh(\boldsymbol{\omega}) < n$. Thus d^* can be interpreted in $\mathsf{CT[S]}$ and an application of Theorem 1.1 completes the proof. ⊣

Instead of providing an interpretation of $\mathsf{CT[S]}$ into $\mathsf{CT^*[S]}$, Lemma 5.1 can be read as assigning ranks to T-cuts in $\mathsf{CT[S]}$ derivations. A corollary of this observation is that the cut-elimination argument for $\mathsf{CT^*[S]}$ can be transferred directly to $\mathsf{CT[S]}$ and these rank assignments. Thus we have

Corollary 5.5. CT *supports cut-elimination for* T*-cuts.*

§6. Conservativity, interpretability and speed-up.

The following instance of Theorem 1.2 is particularly revealing.

Corollary 6.1. *Let* $\mathsf{Ind}_{\mathcal{L}}$ *be the formula expressing that* x *is the code of the universal closure of an instance of* $\mathcal{L}$*-induction. Then* $\mathsf{CT[PA]} + \forall x(\mathsf{Ind}_{\mathcal{L}}x \to \mathrm{T}x)$ *conservatively extends* PA.

Corollary 6.1 effectively shows the limit of what principles can be conservatively added to $\mathsf{CT[PA]}$. It is well known that extending $\mathsf{CT[PA]}$ by induction for formulæ involving the truth predicate (even only for bounded formulæ [12]) allows the

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

deduction of the global reflection principle, $\forall x(\mathsf{Bew}_{\mathsf{PA}} x \to \mathrm{T}x)$, and hence also the local reflection schema $\{\mathsf{Bew}_{\mathsf{PA}}\ulcorner\varphi\urcorner \to \varphi \mid \varphi \in \mathcal{L}\}$, the latter of which is a statement not provable in PA.

An analogous result holds also for other first-order systems such as set theories. For example, if PA is replaced by Zermelo–Fraenkel set theory and Ind is replaced by a formula recognising all instances of the separation and replacement axioms. Expanding the axiom schemata of CT[ZF] to apply also to formulæ involving the truth predicate, however, yields a nonconservative extension in the same way.[3]

We conclude this section with some corollaries that are specific to the proof-theoretic treatment of CT[S].

Corollary 6.2. *Let* D *be an* S*-schema*. $\mathsf{CT[S]} + \forall x(\mathsf{D}x \to \mathrm{T}x)$, *and hence also* CT[S], *attains no better than hyper-exponential speed-up over* S.

To restate Corollary 6.2, every $\mathcal{L}$-theorem of CT[S] is derivable in S with at most hyper-exponential increase in the length of the derivation. The upper-bound results from the fact the conservativeness of CT[S] over S can be established within $\mathsf{I}\Delta_0 + \mathsf{exp}_1$.

Regarding lower bounds to the speed-up phenomenon, we observe that within $\mathsf{CT[PA]} + \forall x(\mathsf{Ind}_{\mathcal{L}} x \to \mathrm{T}x)$ it is simple to prove the consistency of S *on a cut*. Thus we may conclude

Corollary 6.3. $\mathsf{CT[PA]} + \forall x(\mathsf{Ind}_{\mathcal{L}} x \to \mathrm{T}x)$ *provides between exponential and hyper-exponential speed-up over* PA.

It remains open, however, whether this is also the case for the truth axioms alone.

Fischer, in [5], discusses a further consequence of a formalised conservativeness proof for CT.

Lemma 6.4 (Fischer). *If* $\mathsf{PA} \vdash \forall x(\mathsf{Sent}_{\mathcal{L}} x \wedge \mathsf{Bew}_{\mathsf{CT[S_0]}} x \to \mathsf{Bew}_{\mathsf{S_0}} x)$ *for every* $\mathsf{I}\Sigma_1 \subseteq \mathsf{S}_0 \subseteq \mathsf{PA}$ *then* CT[PA] *is relatively interpretable in* PA.[4]

Combining this with Theorem 1.1 therefore yields

Corollary 6.5. CT[PA] *is relatively interpretable in* PA.

§7. Future work. There remain two natural open problems. The first is whether the conservativity results extend also to base theories *weaker* than $\mathsf{I}\Delta_0 + \mathsf{exp}$, for instance sequential theories, theories of bounded arithmetic or syntax theories. The main hurdle in this direction is likely the formalisation of the approximation functions $F_{\varphi,n}$ and their properties within a weaker background theory. Lemmas 4.7 and 4.11 yield at best exponential bounds on the function $F_{\varphi,n}$. However, as the function returns codes of approximations to its arguments it should be formalisable within weaker systems too. The second open problem concerns a characterisation of the speed-up afforded by the compositional axioms. A refinement of the cut elimination argument that is formalisable in $\mathsf{I}\Delta_0 + \mathsf{exp}$ say, would yield optimal bounds on the speed-up of $\mathsf{CT[PA]} + \forall x(\mathsf{Ind}_{\mathcal{L}} x \to \mathrm{T}x)$ and be a significant improvement on the results of this paper.

[3] Assuming ZF is consistent.

[4] We refer the reader to, e.g., [5] for a definition of *relatively interpretable*.

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press

A potential application of the work is in the ordinal analysis of theories of truth, in particular self-referential systems. Currently, the only type-free truth theories for which cut elimination arguments exist are for a selection of the systems introduced by Friedman and Sheard [6] (see [15]). In particular there is no (infinitary) cut-elimination argument for the Friedman-Sheard theory FS, Kripke-Feferman theory KF[5] or the intuitionistic truth theories studied in [16] and [14]. In all of these cases, a cut elimination argument may become possible by generalising the theory of approximations to formulæ contianing an untyped truth predicate and applying the techniques of this paper to obtain explicit ranks on T-cuts in infinitary calculi.

§**8. Acknowledgments.** I would like to thank Ali Enayat, Kentaro Fujimoto, Volker Halbach, Michael Rathjen, Albert Visser, and the anonymous referee for their suggestions and comments on earlier versions of this paper. This work was supported by the Arts and Humanities Research Council UK grant no. AH/H039791/1.

REFERENCES

[1] Barwise, J., ***Admissible Sets and Structures: An Approach to Definability Theory***, Springer–Verlag, Berlin, 1975.

[2] Enayat, A. and A. Visser, *Full satisfaction classes in a general setting (Part 1)*, unpublished manuscript, 2012.

[3] ———, *New constructions of satisfaction classes,* ***Unifying the philosophy of truth*** (Achourioti, T., Galinon, H., Fujimoto, K. and Martínez-Fernández, J., editors), in press.

[4] Feferman, S., *Reflecting on incompleteness*, this Journal, vol. 56 (1991), pp. 1–49.

[5] Fischer, M., *Minimal truth and interpretability*. ***Review of Symbolic Logic***, vol. 2 (2009), no. 4, pp. 799–815.

[6] Friedman, H. and M. Sheard, *An Axiomatic Approach to Self-Referential Truth*. ***Annals of Pure and Applied Logic***, vol. 33 (1987), pp. 1–21.

[7] Halbach, V., *Conservative theories of classical truth*. ***Studia Logica***, vol. 62 (1999), pp. 353–370.

[8] ———, ***Axiomatic theories of truth***, Cambridge University Press, Cambridge, 2011.

[9] ———, ***Axiomatic theories of truth***, second edition. Cambridge University Press, Cambridge, 2014.

[10] Halbach, V. and G. E. Leigh, ***Axiomatic theories of truth*** (Zalta, E. N., editors), The Stanford Encyclopedia of Philosophy (Summer 2014 Edition), url http://plato.stanford.edu/archives/sum2014/entries/truth-axiomatic/.

[11] Krajewski, S., *Nonstandard satisfaction classes*, ***Set Theory and Hierarchy Theory: A Memorial Tribute to Andrzej Mostowski*** (W. Marek et al, editors), Lecture Notes in Mathematics, vol. 537, Springer-Verlag, Berlin, 1976, pp. 121–144.

[12] Kotlarski, H., *Bounded Induction and Satisfaction Classes*. ***Zeitschrift für Mathematische Logik***, vol. 32 (1986), pp. 531–44.

[13] Kotlarski, H., S. Krajewski, and A. H. Lachlan, *Construction of satisfaction classes for nonstandard models*. ***Canadian Mathematical Bulletin***, vol. 24 (1981), pp. 283–93.

[14] Leigh, G. E., *A proof-theoretic account of classical principles of truth*. ***Annals of Pure and Applied Logic***, vol. 164 (2013), pp. 1009–1024.

[15] Leigh, G. E. and M. Rathjen, *An ordinal analysis for theories of self-referential truth*. ***Archive for Mathematical Logic***, vol. 49 (2010), no. 2, pp. 213–247.

[16] ———, *The Friedman-Sheard programme in intuitionistic logic*, this Journal, vol. 77 (2012), no. 3, pp. 777–806.

[5] See, for example, [10] for definitions of FS and KF.

INSTITUTE OF DISCRETE MATHEMATICS AND GEOMETRY
VIENNA UNIVERSITY OF TECHNOLOGY
WIEDNER HAUPTSTRAßE 8–10
1040 VIENNA, AUSTRIA
E-mail: graham.leigh@tuwien.ac.at

https://doi.org/10.1017/jsl.2015.27 Published online by Cambridge University Press