Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in Theoretical Computer Science

Electronic Notes in Theoretical Computer Science 231 (2009) 227–240

www.elsevier.com/locate/entcs

Syntactic Cut-elimination for Common Knowledge

Kai Brünnler [1] and Thomas Studer [2]

Institut für angewandte Mathematik und Informatik
Universität Bern
Bern, Switzerland

Abstract

We see a cut-free infinitary sequent system for common knowledge. Its sequents are essentially trees and the inference rules apply deeply inside of these trees. This allows to give a syntactic cut-elimination procedure which yields an upper bound of $\varphi_2 0$ on the depth of proofs, where φ is the Veblen function.

Keywords: common knowledge, cut elimination, infinitary sequent calculus, deep sequents

1 Introduction

Common knowledge is a well-studied notion in epistemic logic, where modalities express knowledge of agents. Two standard textbooks on epistemic logic and common knowledge in particular, are [6] by Fagin, Halpern, Moses, and Vardi and [12] by Meyer and van der Hoek.

The fact that a proposition A is common knowledge can be expressed by the infinite conjunction "all agents know A and all agents know that all agents know A and so on". In order to express this in a finite way we can use fixpoints: common knowledge of A is then defined to be the greatest fixpoint of $\lambda X.$*everybody knows* A *and everybody knows* X. This notion was introduced by Halpern and Moses [8] and further studied in [6].

The traditional way to formalise common knowledge is to use a Hilbert-style axiom system. Such a system has a fixpoint axiom, which states that common knowledge is a fixpoint, and an induction rule, which states that this fixpoint is the greatest fixpoint. However, this approach does not work well for designing a

[1] Web: `http://www.iam.unibe.ch/~kai/`

[2] Web: `http://www.iam.unibe.ch/~tstuder/`

1571-0661/$ – see front matter
doi:10.1016/j.entcs.2009.02.038

Gentzen-style sequent calculus. In particular, Alberucci and Jäger show in [2] that a cut-free sequent system designed in this way is not complete.

To obtain a complete cut-free system Alberucci and Jäger replace the induction rule by an infinitary ω-rule. This results in a system in which common knowledge is a greatest fixpoint. Although this system has been further studied in [11,9], no syntactic cut-elimination procedure has been found. Cut-elimination was proved only indirectly by showing completeness of the cut-free system.

In the present paper, we give a syntactic cut-elimination procedure for an infinitary system of common knowledge. Since our deductive system for common knowledge includes an ω-rule with infinitely many premises, we have proofs of transfinite depth. We will also assign transfinite ranks to formulas. We obtain our cut-elimination result by using the method of *predicative cut-elimination*, see Pohlers [14,15] and Schütte [17], which is a standard tool for the proof-theoretic analysis of systems of set theory and second order number theory.

In our system we use *deep sequents* which are essentially trees and where rules apply anywhere deep inside of these trees. The general idea of applying rules deeply has been proposed several times in different forms and for different purposes. Schütte already used it in order to obtain systems without contraction and weakening, which he considered more elegant [16]. Guglielmi used it to give a proof-theoretic system for a certain substructural logic which cannot be captured in the sequent calculus. To do so, he developed the *calculus of structures*, a formalism which is centered around deep inference and abolishes the traditional format of sequent calculus proofs [7]. The calculus of structures then has also been developed for modal logic [18]. Based on these ideas, Brünnler introduced the notion of deep sequent and gave a systematic set of sequent systems and a corresponding cut-elimination procedure for the modal logics between K and S5 [5]. Kashima had used the same notion of sequent already in [10] in order to give cut-free sequent systems for some tense logics.

Several cut-free systems for logics with common knowledge exist already. The one that is closest to our system was introduced by Tanaka in [19] for predicate common knowledge logic and is based on Kashima's ideas. It essentially also uses what we call deep sequents. In fact, if one disregards the rather different notation and some choices in the formulation of rules, then one could say that our system is the propositional part of Tanaka's system. There are also finitary systems. Abate, Goré and Widmann, for example, introduce a cut-free tableau system for common knowledge in [1]. Cut-free system have also been studied in the context of *explicit modal logic* by Artemov [4] and by Antonakos [3].

However, we do not know of syntactic cut-elimination procedures for any of the systems mentioned. Typically, cut-elimination is established only indirectly. There are cut-elimination procedures for similar logics, for example by Pliuskevicius' for an infinitary system for linear time temporal logic in [13]. For linear temporal logic he does not need deep sequents. For this logic it is enough to use indexed formulas of the form A^i which denotes A at the i-th moment in time.

The paper is organised as follows. We first present our deep sequent system for common knowledge and prove the invertibility of its rules and the admissibility of the structural rules. Then we embed the Hilbert system for common knowledge into our deep sequent system, which gives us completeness. The main part of the paper is devoted to establishing the reduction lemma, then the cut-elimination theorem follows from that in the standard way. As a result we obtain an upper bound for the depth of proofs in our system. Some discussion about future work ends this paper.

2 The Deep Sequent System

Formulas. We are considering a language with h agents for some $h > 0$. Propositions p and their negations $\bar{p}$ are *atoms*, with $\bar{\bar{p}}$ defined to be p. *Formulas* are denoted by A, B, C, D. They are given by the following grammar:

$$A ::= p \mid \bar{p} \mid (A \vee A) \mid (A \wedge A) \mid \Diamond_i A \mid \Box_i A \mid \circledast\!\!\!\Diamond A \mid \boxast A \quad ,$$

where $1 \leq i \leq h$. The formula $\Box_i A$ is read as "agent i knows A" and the formula $\boxast A$ is read as "A is common knowledge". The connectives $\Box_i$ and $\boxast$ have $\Diamond_i$ and $\circledast\!\!\!\Diamond$ as their respective De Morgan duals.

Given a formula A, its *negation* $\bar{A}$ is defined as usual using the De Morgan laws, $A \supset B$ is defined as $\bar{A} \vee B$ and $\bot$ is defined as $p \wedge \bar{p}$ for some proposition p. The formula $\Box A$ is an abbreviation for "everybody knows A":

$$\Box A = \Box_1 A \wedge \ldots \wedge \Box_h A \quad \text{and} \quad \Diamond A = \Diamond_1 A \vee \ldots \vee \Diamond_h A.$$

A sequence of $n \geq 0$ modal connectives can be abbreviated, for example

$$\Box^n A = \underbrace{\Box \ldots \Box}_{n-\text{times}} A$$

Formula rank. For a formula A we define its *rank* $rk(A)$ as follows:

$$rk(p) = rk(\bar{p}) = 0$$

$$rk(A \wedge B) = rk(A \vee B) = max(rk(A), rk(B)) + 1$$

$$rk(\Box_i A) = rk(\Diamond_i A) = rk(A) + 1$$

$$rk(\boxast A) = rk(\circledast\!\!\!\Diamond A) = \omega + rk(A)$$

Lemma 2.1 (Some properties of the rank) *For all formulas A we have that*
(i) $rk(A) = rk(\bar{A})$,
(ii) there are $m, n < \omega$ such that $rk(A) = \omega \cdot m + n$,
(iii) for all $k < \omega$ we have $rk(\Box^k A) < rk(\boxast A)$.

Proof. Statements (i) and (ii) are immediate. For (iii), an induction on k yields that $rk(\Box^k A) = rk(A) + k \cdot h$. By (ii) it is then enough to check that for all k we have $\omega \cdot m + n + k \cdot h < \omega + \omega \cdot m + n$. $\Box$

Deep sequents. A *(deep) sequent* is a finite multiset of formulas and boxed sequents. A *boxed sequent* is an expression $[\Gamma]_i$ where Γ is a sequent and $1 \leq i \leq h$. Sequents are denoted by $\Gamma, \Delta, \Lambda, \Pi, \Sigma$. A sequent is always of the form

$$A_1, \dots, A_m, [\Delta_1]_{i_1}, \dots, [\Delta_n]_{i_n} \quad ,$$

where the i_j denote agents and thus range from 1 to h. As usual, the comma denotes multiset union and there is no distinction between a singleton multiset and its element. The *corresponding formula* of the above sequent is $\bot$ if $m = n = 0$ and otherwise

$$A_1 \vee \dots \vee A_m \vee \Box_{i_1} D_1 \vee \dots \vee \Box_{i_n} D_n \quad ,$$

where $D_1 \dots D_n$ are the corresponding formulas of the sequents $\Delta_1 \dots \Delta_n$. Often we do not distinguish between a sequent and its corresponding formula, e.g. a model of a sequent is a model of its corresponding formula. A sequent has a *corresponding tree* whose nodes are marked with multisets of formulas and whose edges are marked with agents. The corresponding tree of the above sequent is

$$\begin{tikzcd}[column sep=small] & & \{A_1, \dots, A_m\} \arrow[dll, "i_1"'] \arrow[dl, "i_2"] \arrow[dr, "i_{n-1}"'] \arrow[drr, "i_n"] & & \\ \mathit{tree}(\Delta_1) & \mathit{tree}(\Delta_2) & \cdots & \mathit{tree}(\Delta_{n-1}) & \mathit{tree}(\Delta_n) \end{tikzcd} \quad ,$$

where $\mathit{tree}(\Delta_1) \dots \mathit{tree}(\Delta_n)$ are the corresponding trees of $\Delta_1 \dots \Delta_n$. Often we do not distinguish between a sequent and its corresponding tree, e.g. the *root* of a sequent is the root of its corresponding tree.

Sequent contexts. A *context* is a sequent with exactly one occurrence of the symbol $\{\ \}$, the *hole*, which does not occur inside formulas. Such contexts are denoted by $\Gamma\{\ \}$, $\Delta\{\ \}$, and so on. The hole is also called the *empty context*. The sequent $\Gamma\{\Delta\}$ is obtained by replacing $\{\ \}$ inside $\Gamma\{\ \}$ by Δ. For example, if $\Gamma\{\ \} = A, [[B], \{\ \}]$ and $\Delta = C, [D]$ then

$$\Gamma\{\Delta\} = A, [[B], C, [D]] \quad .$$

Inference rules. In an instance of the inference rule ρ

$$\rho \, \frac{\Gamma_1 \quad \Gamma_2 \quad \dots}{\Delta}$$

we call $\Gamma_1, \Gamma_2 \dots$ its *premises* and Δ its *conclusion*. An *axiom* is a rule without premises. We will not distinguish between an axiom and its conclusion. A *system*, denoted by $\mathcal{S}$, is a set of rules. Figure 1 shows system $\mathsf{D_C}$, our infinitary deep sequent calculus for the logic of common knowledge.

Derivations and proofs. A tree is *well-founded* if it does not have an infinite path. A *derivation* in a system $\mathcal{S}$ is a well-founded tree whose nodes are labelled with sequents and which is built according to the inference rules from $\mathcal{S}$. Derivations

$$\Gamma\{a,\bar{a}\} \qquad \wedge\,\frac{\Gamma\{A\} \quad \Gamma\{B\}}{\Gamma\{A \wedge B\}} \qquad \vee\,\frac{\Gamma\{A, B\}}{\Gamma\{A \vee B\}}$$

$$\Box_i\,\frac{\Gamma\{[A]_i\}}{\Gamma\{\Box_i A\}} \qquad \Diamond_i\,\frac{\Gamma\{\Diamond_i A, [\Delta, A]_i\}}{\Gamma\{\Diamond_i A, [\Delta]_i\}}$$

$$\boxast\,\frac{\Gamma\{\Box^k A\} \quad \text{for all } k \geq 1}{\Gamma\{\boxast A\}} \qquad \diamondast\,\frac{\Gamma\{\diamondast A, \Diamond^k A\}}{\Gamma\{\diamondast A\}}$$

Fig. 1. System $\mathsf{D_C}$

$$\mathsf{nec}\,\frac{\Gamma}{[\Gamma]_i} \qquad \mathsf{wk}\,\frac{\Gamma\{\emptyset\}}{\Gamma\{\Delta\}} \qquad \mathsf{ctr}\,\frac{\Gamma\{\Delta, \Delta\}}{\Gamma\{\Delta\}} \qquad \mathsf{cut}\,\frac{\Gamma\{A\} \qquad \Gamma\{\bar{A}\}}{\Gamma\{\emptyset\}}$$

Fig. 2. Necessitation, weakening, contraction and cut

are visualised as upward-growing trees, so the root is at the bottom. The sequent at the root is the *conclusion* and the sequents at the leaves are the *premises* of the derivation. A *proof* of a sequent Γ in a system is a derivation in this system with conclusion Γ where all leaves are axioms. We write $\mathcal{S} \vdash \Gamma$ if there is a proof of Γ in system $\mathcal{S}$.

Cut rank. The *cut rank* of an instance of cut as shown in Figure 2 is the rank of its *cut formula* A. For an ordinal γ we define the rule cut_γ which is cut with at most rank γ and the rule $\mathsf{cut}_{<\gamma}$ which is cut with a rank strictly smaller than γ. The *cut rank* of a derivation is the supremum of the cut ranks of its cuts. For a system $\mathcal{S}$ and ordinals α and γ and a sequent Γ we write $\mathcal{S} \,|^{\alpha}_{\gamma}\, \Gamma$ to say that there is a proof of Γ in system $\mathcal{S} + \mathsf{cut}_{<\gamma}$ with depth bounded by α.

Admissibility and invertibility. An inference rule ρ with premises $\Gamma_1, \Gamma_2 \ldots$ and conclusion Δ is *depth- and cut-rank-preserving admissible* for a system $\mathcal{S}$ if whenever $\mathcal{S} \,|^{\alpha}_{\gamma}\, \Gamma_i$ for each premise Γ_i then $\mathcal{S} \,|^{\alpha}_{\gamma}\, \Delta$. For each rule ρ there is its *inverse*, denoted by $\bar{\rho}$, which has the conclusion of ρ as its only premise and any premise of ρ as its conclusion. An inference rule ρ is *depth- and cut-rank-preserving invertible* for a system $\mathcal{S}$ if $\bar{\gamma}$ is depth- and cut-rank preserving admissible for $\mathcal{S}$.

In the following, we sometimes omit the "depth- and cut-rank preserving" before either admissible or invertible. Figure 2 shows the structural rules *necessitation*, *weakening* and *contraction*, which are admissible for system $\mathsf{D_C}$.

Lemma 2.2 (Admissibility of the structural rules) *For system* $\mathsf{D_C}$ *the following hold:*

(i) The necessitation rule is depth- and cut-rank-preserving admissible.
(ii) The weakening rule is depth- and cut-rank-preserving admissible.
(iii) All rules are depth- and cut-rank-preserving invertible.
(iv) The contraction rule is depth- and cut-rank-preserving admissible.

Proof. (i) and (ii) follow from a routine induction on the depth of the proof. The same works for the $\wedge, \vee, \Box_i$ and $\boxast$-rules in (iii). The inverses of all other rules are just weakenings. For (iv) we also proceed by induction on the depth of the proof tree, using invertibility of the rules. The cases for the propositional rules and for the $\Box_i, \boxast, \diamondast$-rules are trivial. For the $\Diamond_i$-rule we consider the formula $\Diamond_i A$ from its conclusion $\Gamma\{\Diamond_i A, [\Delta]_i\}$ and its position inside the premise of contraction $\Lambda\{\Sigma, \Sigma\}$. We have the cases 1) $\Diamond_i A$ is inside Σ or 2) $\Diamond_i A$ is inside $\Lambda\{\ \}$. We have three subcases for case 1: 1.1) $[\Delta]_i$ inside $\Lambda\{\ \}$, 1.2) $[\Delta]_i$ inside Σ, 1.3) Σ, Σ inside $[\Delta]_i$. There are two subcases of case 2: 2.1) $[\Delta]_i$ inside $\Lambda\{\ \}$ and 2.2) $[\Delta]_i$ inside Σ. All cases are either simpler than or similar to case 2.2, which is as follows:

$$\mathsf{ctr}\,\dfrac{\Diamond_i\,\dfrac{\Lambda'\{\Diamond_i A, \Sigma', [\Delta, A]_i, \Sigma', [\Delta]_i\}}{\Lambda'\{\Diamond_i A, \Sigma', [\Delta]_i, \Sigma', [\Delta]_i\}}}{\Lambda'\{\Diamond_i A, \Sigma', [\Delta]_i\}} \quad\rightsquigarrow\quad \Diamond_i\,\dfrac{\mathsf{ctr}\,\dfrac{\bar{\Diamond}_i\,\dfrac{\Lambda'\{\Diamond_i A, \Sigma', [\Delta, A]_i, \Sigma', [\Delta]_i\}}{\Lambda'\{\Diamond_i A, \Sigma', [\Delta, A]_i, \Sigma', [\Delta, A]_i\}}}{\Lambda'\{\Diamond_i A, \Sigma', [\Delta, A]_i\}}}{\Lambda'\{\Diamond_i A, \Sigma', [\Delta]_i\}} \quad ,$$

where the instance of $\bar{\Diamond}_i$ in the proof on the right is removed because it is depth-preserving admissible and the instance of contraction is removed by the induction hypothesis. $\Box$

Lemma 2.3 (Admissibility of the general identity axiom) *For all contexts* $\Gamma\{\ \}$ *and all formulas* A *we have* $\mathsf{D_C} \,|\!\overset{2\cdot rk(A)}{\underset{0}{}}\; \Gamma\{A, \bar{A}\}$.

Proof. We perform an induction on $rk(A)$ and a case analysis on the main connective of A. The cases for atoms and for the propositional connectives are obvious. For $A = \Box_i B$ and $A = \boxast B$ we respectively have

$$\Box_i\,\dfrac{\Diamond_i\,\dfrac{\Gamma\{[B, \bar{B}]_i, \Diamond_i \bar{B}\}}{\Gamma\{[B]_i, \Diamond_i \bar{B}\}}}{\Gamma\{\Box_i B, \Diamond_i \bar{B}\}} \qquad \text{and} \qquad \boxast\,\dfrac{\cdots \quad \diamondast\,\dfrac{\Gamma\{\Box^k B, \Diamond^k \bar{B}\}}{\Gamma\{\Box^k B, \diamondast \bar{B}\}} \quad \cdots_{\;1 \le k < \omega}}{\Gamma\{\boxast B, \diamondast \bar{B}\}} \quad .$$

On the left by induction hypothesis we get a proof of the premise of depth $2 \cdot rk(B)$ and thus a proof of the conclusion of depth $2 \cdot rk(B) + 2 = 2 \cdot (rk(B) + 1) = 2 \cdot rk(\Box_i B)$. On the right by Lemma 2.1 we can apply the induction hypothesis for each premise to get a proof of depth $2 \cdot rk(\Box^k B) = 2 \cdot (rk(B) + k \cdot h)$ and thus a proof of the conclusion of depth $2 \cdot (rk(B) + \omega) \le 2 \cdot (\omega + rk(B)) = 2 \cdot rk(\boxast B)$. $\Box$

3 Embedding the Hilbert System

In this section we introduce the Hilbert system $\mathsf{H_C}$ which is essentially the same as system K_h^C from the book [6]. System $\mathsf{H_C}$ is obtained from some Hilbert system for classical propositional logic by adding the axioms and rules shown in Figure 3. Soundness and completeness for $\mathsf{H_C}$ is shown in [6]. We will now embed $\mathsf{H_C}$ into $\mathsf{D_C}$ and thus establish completeness of $\mathsf{D_C}$. We omit a definition of the semantics and a proof of soundness of $\mathsf{D_C}$. We feel that it is straightforward and would not add much to the current paper.

$$\text{(K)} \quad \Box_i A \wedge \Box_i (A \supset B) \supset \Box_i B \qquad \text{(CCL)} \quad \circledast A \supset (\Box A \wedge \Box \circledast A)$$

$$\text{(IND)} \ \frac{B \supset (\Box A \wedge \Box B)}{B \supset \circledast A} \qquad \text{(MP)} \ \frac{A \quad A \supset B}{B} \qquad \text{(NEC)} \ \frac{A}{\Box_i A}$$

Fig. 3. System $\mathsf{H_C}$

Theorem 3.1 *For each formula A if $\mathsf{H_C} \vdash A$ then there are $m, n < \omega$ such that $\mathsf{D_C} \mid_{\omega \cdot n}^{\omega \cdot m} A$.*

Proof. The proof is by induction on the length of the derivation in $\mathsf{H_C}$. If A is a propositional axiom of $\mathsf{H_C}$ then there is a finite derivation of A in system $\mathsf{D_C}$ such that all premises are instances of the general identity axiom. Thus we obtain $\mathsf{D_C} \mid_{0}^{\omega \cdot m} A$ for some $m < \omega$ by admissibility of the general identity axiom (Lemma 2.3).

If A is an instance of (K), then we obtain $\mathsf{D_C} \mid_{0}^{\omega \cdot m} A$ for some $m < \omega$ from the following derivation and admissibility of the general identity axiom to take care of the premises.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Diamond_i \bar{A}, \Diamond_i (A \wedge \bar{B}), [B, A, \bar{A}]_i}{\Diamond_i \bar{A}, \Diamond_i (A \wedge \bar{B}), [B, A]_i}\Diamond_i \qquad \Diamond_i \bar{A}, \Diamond_i (A \wedge \bar{B}), [B, \bar{B}]_i}{\Diamond_i \bar{A}, \Diamond_i (A \wedge \bar{B}), [B, A \wedge \bar{B}]_i}\wedge}{\Diamond_i \bar{A}, \Diamond_i (A \wedge \bar{B}), [B]_i}\Diamond_i}{\Diamond_i \bar{A}, \Diamond_i (A \wedge \bar{B}), \Box_i B}\Box_i}{\Box_i A \wedge \Box_i (A \supset B) \supset \Box_i B}\vee^2}{}$$

If A is an instance of (CCL), then we obtain $\mathsf{D_C} \mid_{0}^{\omega \cdot m} A$ for some $m < \omega$ from the following derivation and again admissibility of the general identity axiom to take care of the premises. An argument similar to the one used to derive the general identity axiom guarantees that all premises of the $\circledast$ rule are derivable with depth smaller than $rk(\circledast A)$.

$$
\dfrac{\dfrac{\dfrac{\Diamond\bar{A},\Box A}{\diamondast\bar{A},\Box A}\;\diamondast,\mathsf{wk}
\qquad
\dfrac{\vdots \qquad \dfrac{\dfrac{\vdots \quad \dfrac{\dfrac{\dfrac{\dfrac{[\Diamond^k\bar{A},\Box^k A]_i}{\Diamond_i\Diamond^k\bar{A},[\Box^k A]_i}\;\Diamond_i,\mathsf{wk}}{\Diamond^{k+1}\bar{A},[\Box^k A]_i}\;\vee,\mathsf{wk}}{\diamondast\bar{A},[\Box^k A]_i}\;\diamondast,\mathsf{wk} \quad \vdots\;{}_{1\le k<\omega}}{}}{\diamondast\bar{A},[\boxast A]_i}\;\boxast}{\diamondast\bar{A},\Box_i\boxast A}\;\Box_i \qquad \vdots\;{}_{1\le i\le h}}{\diamondast\bar{A},\Box\boxast A}\;\wedge}{\diamondast\bar{A},\Box A\wedge\Box\boxast A}\;\wedge}{\boxast A\supset(\Box A\wedge\Box\boxast A)}\;\vee
$$

If the last rule in the derivation is an instance of (MP), then by the induction hypothesis there are $m_1, m_2, n_1, n_2 < \omega$ such that $\mathsf{D_C} \,|^{\omega\cdot m_1}_{\omega\cdot n_1}\; A$ and $\mathsf{D_C} \,|^{\omega\cdot m_2}_{\omega\cdot n_2}\; A \supset B$. Thus we get $\mathsf{D_C} \,|^{\omega\cdot m_1}_{\omega\cdot n_1}\; A, B$ by weakening admissibility and $\mathsf{D_C} \,|^{\omega\cdot m_2}_{\omega\cdot n_2}\; \bar{A}, B$ by invertibility. An application of cut yields $\mathsf{D_C} \,|^{\omega\cdot m}_{\omega\cdot n}\; B$ for $m = max(m_1, m_2) + 1$ and $n = max(n_1, n_2, rk(B) + 1)$.

If the last rule in the derivation is an instance of (NEC), then the claim follows from the induction hypothesis, the fact that nec is cut-rank- and depth-preserving admissible, and an application of $\Box_i$.

If the last rule in the derivation is an instance of (IND), then by the induction hypothesis there are $m_1, n_1 < \omega$ such that $\mathsf{D_C} \,|^{\omega\cdot m_1}_{\omega\cdot n_1}\; B \supset (\Box A \wedge \Box B)$. Then by invertibility of the $\wedge$- and $\vee$-rules we obtain

$$
1)\ \mathsf{D_C} \,|^{\omega\cdot m_1}_{\omega\cdot n_1}\; \bar{B}, \Box B \qquad \text{and} \qquad 2)\ \mathsf{D_C} \,|^{\omega\cdot m_1}_{\omega\cdot n_1}\; \bar{B}, \Box A.
$$

Let n_2 be such that $rk(\Box B) < \omega \cdot n_2$. We set $n = max(n_1, n_2)$. By induction on k we show that for all $k \geq 1$ there is an $m_2 < \omega$ such that $\mathsf{D_C} \,|^{\omega\cdot m_1+m_2}_{\omega\cdot n}\; \bar{B}, \Box^k A$. The case $k = 1$ is given by 2) and the induction step is as follows:

$$
\dfrac{\bar{B},\Box B \qquad \dfrac{\vdots \qquad \dfrac{\dfrac{\dfrac{\dfrac{\bar{B},\Box^k A}{[\bar{B},\Box^k A]_i}\;\mathsf{nec}}{\Diamond_i\bar{B},[\Box^k A]_i}\;\Diamond_i,\mathsf{wk}}{\Diamond_i\bar{B},\Box_i\Box^k A}\;\Box_i}{\Diamond\bar{B},\Box_i\Box^k A}\;\vee,\mathsf{wk} \qquad \vdots\;{}_{1\le i\le h}}{\Diamond\bar{B},\Box^{k+1}A}\;\wedge}{\bar{B},\Box^{k+1}A}\;\mathsf{cut}
$$

,

where the premise on the left is 1) and the premise on the right follows by induction hypothesis. The claim follows by applications of $\boxast$ and $\vee$. $\Box$

4 Cut-Elimination

We write $\alpha \mathbin{\#} \beta$ for the *natural sum of α and β*. The natural sum is commutative because it does not cancel additive components, in contrast to the ordinary ordinal sum Intuitively, the natural sum is the most natural generalisation of addition from natural numbers to ordinal number if one wants to keep commutativity. For an introduction to ordinals, and a definition of the natural sum in particular, we refer to Schütte [17]. The *binary Veblen function* φ is generated inductively as follows:

(i) $\varphi_0\beta := \omega^\beta$,

(ii) if $\alpha > 0$, then $\varphi_\alpha\beta$ denotes the βth common fixpoint of the functions $\lambda\xi.\varphi_\gamma\xi$ for $\gamma < \alpha$.

The Veblen function allows us to name ordinal numbers. First, notice that by using terms built from natural numbers and ω by addition, multiplication and exponentiation we cannot name any ordinal α such that $\omega^\alpha = \alpha$. All of them are greater than any ordinal that we can name. To name them, the ε-function has been introduced, where ε_β is the β-th ordinal α such that $\omega^\alpha = \alpha$. Gentzen's famous result is that the strength of Peano arithmetic is ε_0. For our purposes we have to go a little bit further than the ε-function. We need the Veblen function, which is a generalisation of the ε-function: $\varphi_1\alpha = \varepsilon_\alpha$.

Given a proof π we denote its depth by $|\pi|$. We write $|^{\alpha}_{\beta}\ \Gamma$ for $\mathsf{D_C} \,|^{\alpha}_{\beta}\ \Gamma$.

Lemma 4.1 (Reduction Lemma) *If there is a proof*

$$\mathsf{cut}_\gamma\ \dfrac{\begin{array}{c}\pi_1\\ \Gamma\{A\}\end{array} \qquad \begin{array}{c}\pi_2\\ \Gamma\{\bar{A}\}\end{array}}{\Gamma\{\emptyset\}}$$

with π_1 and π_2 in $\mathsf{D_C} + \mathsf{cut}_{<\gamma}$ then $|^{|\pi_1| \# |\pi_2|}_{\gamma}\ \Gamma\{\emptyset\}$.

Proof. By induction on $|\pi_1| \# |\pi_2|$. We perform a case analysis on the two lowermost rules in the given proofs. If one of the two rules is passive and an axiom then $\Gamma\{\emptyset\}$ is axiomatic as well. If one is active and an axiom then we have

$$\mathsf{cut}_0\ \dfrac{\Gamma\{a,\bar{a}\} \qquad \begin{array}{c}\pi_2\\ \Gamma\{\bar{a},\bar{a}\}\end{array}}{\Gamma\{\bar{a}\}} \quad \leadsto \quad \mathsf{ctr}\ \dfrac{\begin{array}{c}\pi_2\\ \Gamma\{\bar{a},\bar{a}\}\end{array}}{\Gamma\{\bar{a}\}} \ ,$$

and by contraction admissibility we have $\vdash^{|\pi_2|}_{\gamma} \Gamma\{\bar{a}\}$ and thus $\vdash^{|\pi_1| \,\#\, |\pi_2|}_{\gamma} \Gamma\{\bar{a}\}$. If some rule ρ is passive then we have

$$\mathsf{cut}_\gamma \dfrac{\dfrac{\pi_1}{\Gamma\{A\}} \qquad \rho\,\dfrac{\vdots \qquad \dfrac{\pi_{2i}}{\Gamma_i\{\bar{A}\}} \qquad \vdots}{\Gamma\{\bar{A}\}}}{\Gamma\{\emptyset\}} \quad \rightsquigarrow$$

$$\rho\,\dfrac{\vdots \qquad \mathsf{cut}_\gamma\,\dfrac{\bar{\rho}\,\dfrac{\dfrac{\pi_1}{\Gamma\{A\}}}{\Gamma_i\{A\}} \qquad \dfrac{\pi_{2i}}{\Gamma_i\{\bar{A}\}}}{\Gamma_i\{\emptyset\}} \qquad \vdots}{\Gamma\{\emptyset\}} \quad ,$$

where i ranges from 1 to the number of premises of ρ. By invertibility of ρ we get $\vdash^{|\pi_1|}_{\gamma} \Gamma_i\{A\}$, thus by induction hypothesis $\vdash^{|\pi_1| \,\#\, |\pi_{2i}|}_{\gamma} \Gamma_i\{\emptyset\}$ for all i and by ρ we get $\vdash^{|\pi_1| \,\#\, |\pi_2|}_{\gamma} \Gamma\{\emptyset\}$.

This leaves the case that both rules are active and not axioms. We have:

$(\wedge - \vee)$:

$$\mathsf{cut}_{\sigma+1}\,\dfrac{\wedge\,\dfrac{\dfrac{\pi_{11}}{\Gamma\{B\}} \qquad \dfrac{\pi_{12}}{\Gamma\{C\}}}{\Gamma\{B \wedge C\}} \qquad \vee\,\dfrac{\dfrac{\pi_{21}}{\Gamma\{\bar{B}, \bar{C}\}}}{\Gamma\{\bar{B} \vee \bar{C}\}}}{\Gamma\{\emptyset\}} \quad \rightsquigarrow$$

$$\mathsf{cut}_\sigma\,\dfrac{\dfrac{\pi_{11}}{\Gamma\{B\}} \qquad \mathsf{cut}_\sigma\,\dfrac{\mathsf{wk}\,\dfrac{\dfrac{\pi_{12}}{\Gamma\{C\}}}{\Gamma\{\bar{B}, C\}} \qquad \dfrac{\pi_{21}}{\Gamma\{\bar{B}, \bar{C}\}}}{\Gamma\{\bar{B}\}}}{\Gamma\{\emptyset\}} \quad ,$$

where by weakening admissibility we get $\vdash^{|\pi_{12}|}_{\gamma} \Gamma\{\bar{B}, C\}$, and since $\sigma < \sigma + 1 = \gamma$ we get $\vdash^{\alpha}_{\gamma} \Gamma\{\emptyset\}$ for $\alpha = max(|\pi_{11}|, max(|\pi_{12}|, |\pi_{21}|) + 1) + 1$. It is easy to check that $\alpha \leq |\pi_1| \,\#\, |\pi_2|$.

$(\Box_i - \Diamond_i)$:

$$\mathsf{cut}_{\sigma+1}\,\dfrac{\Box_i\,\dfrac{\overset{\pi_{11}}{\Gamma\{[\Delta]_i,[A]_i\}}}{\Gamma\{[\Delta]_i,\Box_i A\}} \qquad \Diamond\,\dfrac{\overset{\pi_{21}}{\Gamma\{[\Delta,\bar{A}]_i,\Diamond_i\bar{A}\}}}{\Gamma\{[\Delta]_i,\Diamond_i\bar{A}\}}}{\Gamma\{[\Delta]_i\}} \qquad \rightsquigarrow$$

$$\mathsf{cut}_{\sigma}\,\dfrac{\mathsf{ctr}\,\dfrac{\mathsf{wk}^2\,\dfrac{\overset{\pi_{11}}{\Gamma\{[\Delta]_i,[A]_i\}}}{\Gamma\{[\Delta,A]_i,[\Delta,A]_i\}}}{\Gamma\{[\Delta,A]_i\}} \qquad \mathsf{cut}_{\sigma+1}\,\dfrac{\mathsf{wk},\Box_i\,\dfrac{\overset{\pi_{11}}{\Gamma\{[\Delta]_i,[A]_i\}}}{\Gamma\{[\Delta,\bar{A}]_i,\Box_i A\}} \qquad \overset{\pi_{21}}{\Gamma\{[\Delta,\bar{A}]_i,\Diamond_i\bar{A}\}}}{\Gamma\{[\Delta,\bar{A}]_i\}}}{\Gamma\{[\Delta]_i\}} \quad ,$$

where the premises of the upper cut have been derived by use of weakening admissibility with depth $|\pi_{11}|+1$ and $|\pi_{21}|$, the natural sum of which is smaller than $|\pi_1|\,\#\,|\pi_2|$. The induction hypothesis thus yields $\vdash^{(|\pi_{11}|+1)\,\#\,|\pi_{21}|}_{\gamma} \Gamma\{[\Delta,\bar{A}]_i\}$ and since $\sigma<\sigma+1=\gamma$ we get $\vdash^{|\pi_1|\,\#\,|\pi_2|}_{\gamma} \Gamma\{[\Delta]_i\}$ by the lower cut.

$(\boxast - \diamondast)$:

$$\mathsf{cut}_{\omega+\sigma}\,\dfrac{\boxast\,\dfrac{\cdots \quad \overset{\pi_{1k}}{\Gamma\{\Box^k A\}} \quad \cdots \quad {\scriptstyle k<\omega}}{\Gamma\{\boxast A\}} \qquad \diamondast\,\dfrac{\overset{\pi_{21}}{\Gamma\{\diamondast\bar{A},\Diamond^j\bar{A}\}}}{\Gamma\{\diamondast\bar{A}\}}}{\Gamma\{\emptyset\}} \qquad \rightsquigarrow$$

$$\mathsf{cut}_{\sigma+(j\cdot h)}\,\dfrac{\overset{\pi_{1j}}{\Gamma\{\Box^j A\}} \qquad \mathsf{cut}_{\omega+\sigma}\,\dfrac{\boxast\,\dfrac{\cdots \quad \mathsf{wk}\,\dfrac{\overset{\pi_{1k}}{\Gamma\{\Box^k A\}}}{\Gamma\{\Box^k A,\Diamond^j\bar{A}\}} \quad \cdots \quad {\scriptstyle k<\omega}}{\Gamma\{\boxast A,\Diamond^j\bar{A}\}} \qquad \overset{\pi_{21}}{\Gamma\{\diamondast\bar{A},\Diamond^j\bar{A}\}}}{\Gamma\{\Diamond^j\bar{A}\}}}{\Gamma\{\emptyset\}} \quad ,$$

where the induction hypothesis applied on the upper cut gives us $\vdash^{|\pi_1|\,\#\,|\pi_{21}|}_{\gamma} \Gamma\{\Diamond^j\bar{A}\}$ and since by Lemma 2.1 we have $\sigma+j\cdot h<\omega+\sigma=\gamma$ the lower cut yields $\vdash^{|\pi_1|\,\#\,|\pi_2|}_{\gamma} \Gamma\{\emptyset\}$.

$\square$

From the reduction lemma we obtain the first and the second elimination lemma as usual, see for instance Pohlers [14,15] or Schütte [17].

Lemma 4.2 (First Elimination Lemma) *If* $\vdash^{\alpha}_{\gamma+1} \Gamma$ *then* $\vdash^{2^\alpha}_{\gamma} \Gamma$.

Lemma 4.3 (Second Elimination Lemma) *If* $\vdash^{\alpha}_{\beta+\omega^\gamma} \Gamma$ *then* $\vdash^{\varphi_\gamma \alpha}_{\beta} \Gamma$.

The embedding of the Hilbert system into the deep sequent system together with the second elimination lemma gives us the cut elimination theorem.

Theorem 4.4 (Cut Elimination) *If A is a valid formula, then* $\vdash^{\varphi_2 0}_{0} A$.

Proof. Let A be a valid formula. By the embedding of the Hilbert system into the deep sequent system, there are natural numbers m, n such that $\mathsf{D_C} \vdash^{\omega\cdot m}_{\omega\cdot n} A$. By the second elimination lemma we obtain $\mathsf{D_C} \vdash^{\alpha}_{0} A$ where $\alpha = \varphi_1(\ldots(\varphi_1(\omega \cdot m))\ldots)$. We know $\varphi_{\beta_1}\gamma_1 < \varphi_{\beta_2}\gamma_2$ if $\beta_1 < \beta_2$ and $\gamma_1 < \varphi_{\beta_2}\gamma_2$. Thus $\alpha < \varphi_2 0$. $\Box$

5 Conclusion

We have introduced an infinitary deep sequent system for common knowledge and a syntactic cut-elimination procedure for it. We embedded the Hilbert style system and obtained $\varphi_2 0$ as upper bound on the length of cut-free proofs for valid formulas.

To draw some more conclusions, let us look at the problem of cut elimination in the ordinary sequent calculus, for example in the one by Alberucci and Jäger. It has the following $\Box_i$-rule:

$$\Box_i \; \dfrac{A, \Gamma, \circledast\Delta}{\Box_i A, \Diamond_i \Gamma, \circledast\Delta, \Sigma} ,$$

where Γ, Δ and Σ are sets of formulas and $\Diamond_i\Gamma$ is $\{\Diamond_i A \mid A \in \Gamma\}$. The problem here is the context restriction. Consider the following proof, where the cut is multiplicative (context-splitting)

$$\text{cut}\;\dfrac{\Box_i\,\dfrac{\dfrac{\pi_1}{A, \Gamma, \circledast\bar{B}}}{\Box_i A, \Diamond_i \Gamma, \Sigma, \circledast\bar{B}} \qquad \boxast\,\dfrac{\vdots \quad \dfrac{\pi_{2k}}{\Box^k B, \Delta} \quad \vdots_{\,1 \le k < \omega}}{\boxast B, \Delta}}{\Box_i A, \Diamond_i \Gamma, \Sigma, \Delta}$$

Here the inference rule above the cut on the left does not apply to the cut formula while the inference rule on the right does. The typical transformation would push the left rule instance below the cut, as follows:

$$
\Box_i\,\dfrac{\mathsf{cut}\,\dfrac{\dfrac{\pi_1}{A,\Gamma,\circledast\bar{B}} \qquad \boxast\,\dfrac{\vdots \qquad \dfrac{\pi_{2k}}{\Box^k B,\Delta} \qquad \vdots_{\;1\leq k<\omega}}{\boxast B,\Delta}}{A,\Gamma,\Delta}}{\Box_i A,\Diamond_i\Gamma,\Sigma,\Diamond_i\Delta}
$$

However, this transformation introduces the $\Diamond_i$ in $\Diamond_i\Delta$, and thus it does not yield a proof of the original conclusion. This is caused by the context restriction in the $\Box_i$-rule. Such a context restriction also occurs in the standard sequent calculus for the modal logic K. While it is hardly elegant, at least it does not cause any difficulties for syntactic cut-elimination for K. However, we see that the context restriction poses a genuine problem for logics with more modalities like in the logic of common knowledge. Our more general format for sequents and inference rules solves the problem since it does not require context restrictions.

The first item on the list of future work is of course to embed our cut-free deep sequent system into the ordinary cut-free sequent system by Alberucci and Jäger. This would yield a syntactic cut-elimination procedure for their system, since the embeddings with cut are straightforward. We think we know how to do this, but we still have to check the details. The second item on the list is cut-elimination for a system for S5-based common knowledge. After all, S5 is *the* system for knowledge, and deep sequents easily handle S5. Generalising contexts to allow two holes, the rule to add would be

$$
\mathsf{S5}\ \dfrac{\Gamma\{\Diamond A\}\{A\}}{\Gamma\{\Diamond A\}\{\emptyset\}} \quad .
$$

After that, questions become more speculative. What is the mathematical meaning of the upper bound on the depth of cut-free proofs? Is there a kind of boundedness lemma in modal logic similar to the one used in the analysis of set theories and second order arithmetic? Is $\varphi_2 0$ the best possible upper bound on the depth of proofs? What would be the equivalent of a well-ordering proof in modal logic? And finally, how could one syntactically eliminate cuts in a finitary system?

References

[1] Pietro Abate, Rajeev Goré, and Florian Widmann. Cut-free single-pass tableaux for the logic of common knowledge. In *Workshop on Agents and Deduction at TABLEAUX 2007*, 2007.

[2] Luca Alberucci and Gerhard Jäger. About cut elimination for logics of common knowledge. *Annals of Pure and Applied Logic*, 133:73–99, 2005.

[3] Evangelia Antonakos. Justified and common knowledge: Limited conservativity. In *LFCS*, pages 1–11. Springer, 2007.

[4] Sergei Artemov. Justified common knowledge. *Theoretical Computer Science*, 357(1):4–22, 2006.

[5] Kai Brünnler. Deep sequent systems for modal logic. In Guido Governatori, Ian Hodkinson, and Yde Venema, editors, *Advances in Modal Logic*, volume 6, pages 107–119. College Publications, 2006.

[6] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. The MIT Press, Boston, 1995.

[7] Alessio Guglielmi. A system of interaction and structure. *ACM Transactions on Computational Logic*, 8(1):1–64, 2007.

[8] Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.

[9] Gerhard Jäger, Mathis Kretz, and Thomas Studer. Cut-free common knowledge. To appear in the Journal of Applied Logic, 2007.

[10] Ryo Kashima. Cut-free sequent calculi for some tense logics. *Studia Logica*, 53:119–135, 1994.

[11] Mathis Kretz and Thomas Studer. Deduction chains for common knowledge. *Jornal of Applied Logic*, 4:331–357, 2006.

[12] John-Jules Meyer and Wiebe van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, 1995.

[13] Regimantas Pliuskevicius. Investigation of finitary calculus for a discrete linear time logic by means of infinitary calculus. In *Baltic Computer Science, Selected Papers*, pages 504–528, London, UK, 1991. Springer-Verlag.

[14] Wolfram Pohlers. *Proof Theory - An introduction*. Springer, 1989.

[15] Wolfram Pohlers. Subsystems of set theory and second order number theory. In Sam Buss, editor, *Handbook of Proof Theory*, pages 209–335. Elsevier, 1998.

[16] Kurt Schütte. Schlussweisen-Kalküle der Prädikatenlogik. *Mathematische Annalen*, 122:47–65, 1950.

[17] Kurt Schütte. *Proof Theory*. Springer-Verlag, 1977.

[18] Charles Stewart and Phiniki Stouppa. A systematic proof theory for several modal logics. In Renate Schmidt, Ian Pratt-Hartmann, Mark Reynolds, and Heinrich Wansing, editors, *Advances in Modal Logic*, volume 5 of *King's College Publications*, pages 309–333, 2005.

[19] Yoshihito Tanaka. Some proof systems for predicate common knowledge logic. *Reports on mathematical logic*, 37:79–100, 2003.