Australasian Journal of Logic

Normalization Procedures for the Strongest Syllogistic Natural Deduction Systems

Vasiliy Shangin
Department of Logic, Faculty of Philosophy
Lomonosov MSU, Lenin hills, Moscow, 119234, Russia.
shangin@philos.msu.ru

Abstract

This paper deals with the natural deduction systems $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, and $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ for the relational syllogistics $R^{*\dagger}$, $R^{*\dagger}_{trans}$, and $RCA^{\dagger}(opp)$, respectively, studied by L.S. Moss and his coauthors. Characteristic features of $R^{*\dagger}$—the weakest of these syllogistics—are two-place relations only expressed with transitive verbs or comparative adjectives, such as 'like' or 'happier than', which, together with quantifiers, could form a nested relative clause—*set terms*—such as 'a thing that likes some women, who is happier than all men'. $R^{*\dagger}_{trans}$ extends $R^{*\dagger}$ with the transitivity property of the comparative adjectives, and $\mathbf{RCA}^{\dagger}(\mathbf{opp})$—the strongest of these syllogistics—extends $R^{*\dagger}_{trans}$ with the Boolean conjunction applied inside the set terms, such as 'a thing that likes and sees some women', and converse relations, such as 'is liked by'. Each of the syllogistics allows set term negation, such as 'nonwoman', 'dislike', or 'unhappier than', and this negation enjoys the *bar notation* consisting of the (standard) equivalences that hold true on the level of the set terms. Being propositional connective-free, the syllogistic languages employ atomic sentences only, which are composed of the set terms and constants, such as 'Angelina likes some women who do not know any puppy', with the bar notation holding true for them as well. Another specifics of $RCA^{\dagger}(opp)$ is that its formulation demands extending its language with individual variables to compose *general* sentences, such as 'x likes some women'.

The main result of this paper is the normalization theorems for the systems under consideration: each derivation in any of them is converted to a normal one via detour, permutation, and simplification conversions, where the normal derivation is efficient in the sense that it does not contain detours, hidden detours, or vacuous applications of rules canceling assumptions, respectively. The important corollary of these theorems is that each system has a syllogistic analog of the

subformula property in the case of propositional logics—the so-called *sub-set-term-property*, i.e., in a normal derivation of ϕ from Γ, each set term is a subterm of $\overline{d}$, where d is a term in ϕ and/or Γ, and each constant is a constant in ϕ and/or Γ. An overview of future plans, a conclusion, and a reference to some previous work are at the end of this paper.

Keywords natural logic, syllogistic, natural deduction, normalization, subformula, subterm

1 Introduction

The strongest syllogistics in natural logic In the literature on syllogistics, one could find some candidates for the strongest one. Of course, king of the hill depends on the underlying language and thus could not be determined once and for all in principle. However, a few syllogistics are far more common due to the tradition and teaching in logic: they are usually part of the modern trivium in humanities and (applied) mathematics. Roughly, the core of these syllogistics consists of assertoric syllogistics built over the four conventional formulae (a), (i), (e), (o) and of (a fragment of) relational syllogistics built over the conventional two-place relational formulae (for instance, 'All women like some puppies') such that it is strong enough to be suitable for the purposes of *natural logic* [Moss(2015)] and, because of these purposes, is decidable yet. If one is content herself with these boundaries—natural-friendly simplicity plus decidability—then one could find the needed candidates among the variety of syllogistics that *Moss and his coauthors*[1] have been systematically studying for the recent dozen of years [Pratt(2009), Moss(2010), Moss(2010a), Moss(2015), KruckMoss(2021), the list is very far from being exhaustive]. Then to determine the strongest of these syllogistics seems to be easy because the needed one is supposed to be the last item on their list. In [Pratt(2009), p. 669], the strongest one is $\mathcal{R}^{*\dagger}$ that is shown not to have natural deduction systems of two special kinds. In [Moss(2010), pp. 547, 557, respectively], it is a *modified* variant of $\mathcal{R}^{*\dagger}$ and its extension by the rule (*trans*). In [Moss(2015), p. 578], it is $\mathcal{RCA}^{\dagger}(opp)$. Finally, in [KruckMoss(2021), p. 33], it is $\mathcal{R}^{*\dagger}$, again. Hence, the quest for a proper candidate to be the strongest of the syllogistics in natural logic boils down to this trio in [Moss(2010), Moss(2015)].

[1]Such a collocation that we employ through this paper is not to evaluate personal contributions of a handful of the logicians cited a little lower. Rather, it reflects the obvious fact that L.S. Moss is the only one (co)author of the cited papers.

In this paper, we are going to explicate the candidates by their natural deduction systems—this is the proof system format [Gentzen(1969), Prawitz(1965)] Moss and his coauthors choose in presenting these syllogistics[2]—rather than by their languages and entailment relations. The systems for the above untitled one of the syllogistics and for its extension that we dub $\mathbf{R}^{*\dagger}$ and $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, respectively, contain 10 and 11 rules. Both systems are shown to be sound and complete [Moss(2010), pp. 549–53, 557–8]. $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ has 15 rules: the 11 $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$-rules plus 4 new ones. Without providing any reference or proof, Moss explicitly states its soundness and completeness, 'All of the proof systems in this chapter will be sound and complete: they define relations $\Gamma \vdash \phi$ based on formal proofs in the various systems, and we always have $\Gamma \vdash \phi$ *iff* $\Gamma \models \phi$, where the latter relation is semantic' [Moss(2015), pp. 563–4]. To sum it up, three natural deduction systems sorted by deductive strength are $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, $\mathbf{RCA}^{\dagger}(\mathbf{opp})$, with the latter being the needed candidate.

A syllogistic variant of the subformula property When it comes to natural deduction as one of the most popular proof system formats in modern proof theory, the normalization problem—given any derivation, obtain a *normal* form of it such that it contains no detours, or rules applications which are not efficient in some obvious sense [PelletierHazen(2023)]. Traditionally, proving the normalization theorem is just the beginning of the story. Its importance lies within its corollaries, of which the *subformula property*—each formula occurring in a normal derivation of ϕ from Γ is (a negation of) a subformula of $\{\phi, \Gamma\}$—is arguably the important one. Holding this property opens well-grounded perspectives for proof searching by confining it to an always finite set of the subformulae (in the propositional case at least).

To ease understanding, let us explicate Moss and his coauthors' original terminology that we use through this paper, '*General sentences* in (...) are what are usually called *formulas*. We prefer to change the standard terminology to make the point that here, sentences are not built from formulas by quantification. In fact, our sentences do not have variable occurrences. But general sentences do include *variables*. They are only used in our proof theory' [Moss(2015), p. 578; the *italics* original]. However, we would like to avoid more *changes of the standard terminology* by proposing the *subsentence property*, which would imply speaking of the subformula property

[2]Frankly, some syllogistics presented by them do not have proof systems because their motivation is mostly not proof-theoretic, which is in contrast to the one of this paper.

without speaking of 'formula'. Another change of the standard terminology is to use '(un)canceled' instead of the more conventional '(non)discarded', or '(non)discharged'.

In the syllogistic case without propositional connectives, one is hardly able to speak of subformulae of atomic formulae—the only sentences which $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, and $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ employ are of this kind. Rather, we present a syllogistic analog of the subformula property—*sub-set-term-constant property*[3]—in definition 11 below in order to underline a derivation where there are (negations of) the set terms and constants that its assumption(s) and conclusion contain only. Arguably it is the best one could achieve given the fact that in the derivation considered there are expressions that its assumption(s) and conclusion do not contain. To sum this chapter and make contrasts with the propositional case brighter, let us highlight the specifics of how we tackle the normalization problem for $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, and $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ below. First, detour 3.3 has a special case (detour 3.3.2) because sometimes a *genus* conversion (chapter 3 below) causes a certain unwanted effect. Second, maximal segments (definition 7) might be formed by other general sentences than an 'arbitrary' α (for example, permutation 3.9 below).

This paper is organized as follows. Section 2 presents Moss and his coauthors' natural deduction systems $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, and $\mathbf{RCA}^{\dagger}(\mathbf{opp})$. The normalization procedures consisting of detour, permutation, and simplification conversions for these systems, as well as the main results are in section 3. Section 4 outlines previous works. In section 5, one finds a conclusion and an outline for future studies.

2 Natural deduction systems $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, $\mathbf{RCA}^{\dagger}(\mathbf{opp})$

Our exposition mainly focuses on $\mathbf{RCA}^{\dagger}(\mathbf{opp})$—the most powerful system among the ones section 1 mentions above—and from time to time we highlight its parts which belong to $\mathbf{R}^{*\dagger}$ and $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$. In general, we stick to [Moss(2015)] while making some comments.[4]

In table 1, there are mutually exclusive sets of unary atoms $\{p_1, q_1, p_2, \ldots\}$, of adjective atoms $\{a_1, a_2, \ldots\}$, of tv atoms $\{t'_1, t'_2, \ldots\}$ and of constants $\{k_1, j_1, k_2, \ldots\}$ that, in English, stand for general names (nouns), such as 'woman', comparative adjectives, such as 'bigger than', transitive[5] verbs,

[3]This philological monster is, of course, due to a lack of a better name.

[4]We fix the original variable t over tv atoms to t' for the reason that the passage under table 2 below explicates, and remove the misprinting—the variable b runs over both binary atoms and set terms no more.

[5]One needs transitive verbs to form binary relational terms, which is impossible in the

Table 1: Syntax of terms and sentences of $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ [Moss(2015), Table 18.2]

Expression	**Variables**	**Syntax**
unary atom	$p,\ q$	
adjective atom	a	
tv atom	t'	
binary atom	b	$a \mid t'$
constant	$k,\ j$	
unary relational term	$l,\ m$	$p \mid \bar{l} \mid l \wedge m$
binary relational term	$r,\ s$	$b \mid r^{-1} \mid \overline{r} \mid r \wedge s$
set term	$c,\ d$	$l \mid \exists(c, r) \mid \forall(c, r)$
sentence	ϕ, ψ	$\forall(c, d) \mid \exists(c, d) \mid c(k) \mid r(k,j)$

such as 'like', and proper names, such as 'Angelina', respectively. In fact, we will use the respective subscripted variables running over them (and all the expressions to follow) instead. Binary atoms are either tv atoms or adjective atoms, which reflects the idea that binary relations are formed with them only. Unary relational terms are either unary atoms or their negations,[6] such as 'nonwoman'. Binary relational terms are either binary atoms or their inverses, such as 'is liked', or their negations, such as 'dislike', or their conjunctions, such as 'is disliked and is seen'. Set terms are either unary relational terms or their existential concatenations with the binary relational terms, such as 'a thing that likes some women', or their universal concatenations with binary relational terms, such as 'a thing that likes all women'.[7] Sentences are either universal concatenations of a pair of set terms, such as 'All puppies like some women', or existential concatenations of a pair of set terms, such as 'Some puppies like no women', or concatenations of a set term and a constant, such as 'Angelina likes some puppy', or concatenations of a pair of constants, such as 'Angelina likes Basil the

case of intransitive verbs, such as 'go' that do not have a direct object. Notice the trivial truth that interpretations of the tv atoms are 'not necessary transitive. By stipulation, those of binary adjectives are' [Moss(2015), p. 574]. For example, Angelina likes Basil the puppy, and it likes to eat each stick, but it might not be the case that Angelina likes to eat each stick. Conversely, 'by stipulation' can imply that, unless, for example, The Happiness Index© is applicable, the transitive interpretation of 'happier than' is debatable.

[6] Details are in the quote on page 256 below.

[7] Notice that set terms containing constants such as 'a thing that does not like Angelina' are illegitimate, 'We should note that the relative clauses that can be obtained in this way are all "missing the subject", never "missing the object". The language is too poor to express predicates like λx.all boys see x' [Moss(2015), p. 573].

puppy'. The letter Γ with subscripts runs over sentences.

Table 2: Syntax of general sentences of $\mathbf{RCA}^{\dagger}(\mathbf{opp})$, with ϕ ranging over sentences, c over set terms, and r over relational terms. [Moss(2015), Table 18.3]

Expression	**Variables**	**Syntax**
individual variable	x, y	
individual term	t, u	$x \mid j$
general sentence	α	$\phi \mid c(x) \mid r(x,y) \mid \bot$

Notice that in the original table 1 (see footnote 4 above), t rather than t' runs over the tv atoms. Hence, the use of t in the original table 2 is inaccurate. We fix this by introducing t' instead of t in table 1. In this paper, t is used in the sense of table 2 only.

In table 2, there is added a set of individual variables $\{x_1, y_1, x_2, \ldots\}$ that, in English, stand for pronouns, such as 'she', instead of which we will use the respective subscripted variables running over them (and the expressions to follow). Individual terms are either the individual variables or the constants from table 1. General sentences are either the sentences or concatenations of a set term and a variable, such as 'She likes some puppies', concatenations of a binary relational term and a pair of the variables, such as 'She likes it', or the falsum constant.[8] Notice that no variable occurs in the set terms—the expression 'a thing that likes her' is illegitimate— therefore, there are no nested occurrences of the variables in the general sentences—the expression 'Angelina, who knows it, likes her' is illegitimate.

Definition 1. (*Bar notation* [Moss(2015), pp. 573, 577-8; clause 3 below is given implicitly by Moss]) The bar notation satisfies the following equations:

1. $\overline{\overline{p}} = p$, $\overline{\overline{r}} = r$;
2. $\overline{\exists(l,r)} = \forall(l,\overline{r})$, $\overline{\forall(l,r)} = \exists(l,\overline{r})$, $\overline{\exists(c,d)} = \forall(c,\overline{d})$, $\overline{\forall(c,d)} = \exists(c,\overline{d})$, $\overline{r^{-1}} = (\overline{r})^{-1}$, $\overline{b(x)} = \overline{b}(x)$, $\overline{r(x,y)} = \overline{r}(x,y)$;
3. $r^{-1^{-1}} = r$, $\overline{\overline{\phi}} = \phi$, $\overline{\overline{\alpha}} = \alpha$.

For instance, 'It is not the case that Angelina likes some puppy who eats any stick' is 'Angelina does not like any puppy who eats some sticks'

[8] A falsum, such as 'Angelina, who likes no puppy, likes Basil the puppy', could have been specified in the language of table 1 already, but there is no need in it.

or 'Angelina likes no puppy who eats some sticks'.[9] Thus, 'Angelina likes some puppy who eats every stick' is inconsistent with each of the following sentences: 'It is not the case that Angelina likes some puppy who eats any stick', 'Angelina does not like any puppy who eats some sticks', and 'Angelina likes no puppy who eats some sticks'.

Remark 1. Following Moss,[10] we explicitly acknowledge the bar in definition 1 not to be a first-class citizen of the syllogistics treated. On the one hand, it is applicable to terms within (general) sentences; on the other hand, its applications to (general) sentences are identifiable to certain (general) sentences such that it is applicable to terms occurring in the latter ones only.

The semantics for $\mathcal{R}^{*\dagger}$, $\mathcal{R}^{*\dagger}_{trans}$, and $\mathcal{RCA}^{\dagger}(opp)$ is in [Moss(2015)].

Cosmetically, the $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-rules in [Moss(2015), Figure 18.3.] are modified as follows.[11]

$\mathbf{R}^{*\dagger}$ consists of the rules in the top three rows, and $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$ consists of the $\mathbf{R}^{*\dagger}$-rules plus (*trans*).

$$\forall E\ \frac{c(t) \quad \forall(c,d)}{d(t)} \qquad \forall E_r\ \frac{c(u) \quad \forall(c,r)(t)}{r(t,u)} \qquad \exists I\ \frac{c(t) \quad d(t)}{\exists(c,d)} \qquad \exists I_r\ \frac{r(t,u) \quad c(u)}{\exists(c,r)(t)}$$

$$\begin{array}{ccc} [c(x)] & [c(x)] & [\overline{\phi}] \\ \vdots & \vdots & \vdots \\ \forall I\ \dfrac{d(x)}{\forall(c,d)} & \forall I_r\ \dfrac{r(t,x)}{\forall(c,r)(t)} & \overline{E}\ \dfrac{\alpha \quad \overline{\alpha}}{\bot} \qquad \overline{I}\ \dfrac{\bot}{\phi} \end{array}$$

[9] Semi-formally, $\overline{(\exists(puppy, \forall(stick, eat)), like)(Angelina)} = (\forall(puppy, \exists(stick, eat)), \overline{like})(Angelina)$. Formally, $\overline{(\exists(p, \forall(q, r)), s)(k)} = (\forall(p, \exists(q, r)), \overline{s})(k)$.

[10] 'We intend the bar notation for negation to be involutive: that is, we shall never iterate the bar notation, and so we simply will not see expressions such as $\overline{\overline{p}}$. We make the choice that our syntax does not include such a notation. However, it will be convenient to adopt a syntactic negation and to say that the syntax should be closed under the bar notation; this means that we would be *identifying* $\overline{\overline{p}}$ with p, and making similar identifications' [Moss(2015), p. 573; the *italics* original].

[11] We split the double line rules and name the eponymous rules differently, where '$_r$' refers to a binary rather than unary relational term that its counterpart rule deals with, 'Returning to the two rules with the same name, we could have chosen to use different names, say ($\forall E1$) and ($\forall E2$). But the result would have been a more cluttered notation, and it is always clear from context which rule is being used' [Moss(2015), p. 546]. In tackling the normalization problem, one has to use 'a more cluttered notation', i.e., a more explicit one. See also remark 2 about ($\overline{E}$) and ($\overline{I}$).

$$\exists E\ \frac{\exists(c,d) \qquad \begin{array}{c}[c(x)]\ \ [d(x)]\\ \vdots \\ \alpha\end{array}}{\alpha} \qquad \exists E_r\ \frac{\exists(c,r)(t) \qquad \begin{array}{c}[c(x)]\ \ [r(t,x)]\\ \vdots \\ \alpha\end{array}}{\alpha}$$

$$trans\ \frac{a(k,j) \quad a(j,l)}{a(k,l)} \qquad \wedge I\ \frac{r(k,j) \quad s(k,j)}{(r\wedge s)(k,j)} \qquad \wedge E_1\ \frac{(r\wedge s)(k,j)}{r(k,j)}$$

$$\wedge E_2\ \frac{(r\wedge s)(k,j)}{s(k,j)} \qquad invI\ \frac{r(k,j)}{r^{-1}(j,k)} \qquad invE\ \frac{r^{-1}(j,k)}{r(k,j)}$$

There are side conditions on the $(\forall I)$-rules that x must not occur free in any uncanceled assumption and on the $(\exists E)$-rules that the variable x must not occur free in the conclusion α or in any uncanceled assumption in the subderivation of α.[12]

Remark 2. We explicitly demarcate each of these rules as an introduction or an elimination. There are 9 eliminations—$(\forall E)$, $(\exists E)$, $(\forall E_r)$, $(\exists E_r)$, $(trans)$, $(\wedge E_1)$, $(\wedge E_2)$, $(\overline{E})$, and $(invE)$—and 7 introductions—$(\forall I)$, $(\exists I)$, $(\forall I_r)$, $(\exists I_r)$, $(\wedge I)$, $(\overline{I})$, and $(invI)$.[13] $(\wedge E_1)$, $(\wedge E_2)$, and $(invE)$ are *one-premiss*; the others are *two-premiss*. The original names for $(\overline{I})$ and $(\overline{E})$ are (RAA)[14] and $(\perp I)$, respectively, which we find quite misleading, especially the latter. In the proof-theoretic tradition, $(\perp I)$ is a (negation) elimination rather than an introduction,[15] which is best seen together with its counterpart (*Reductio Ad Absurdum*), or a (negation) introduction.

Once remark 2 demarcates the eliminations, we are ready for the following

Definition 2. (An adaptation of 1.2.3. Major and minor premisses [Prawitz(1971), p. 241]) The *major* premiss is the occurrences of $\forall(c,d)$, $\forall(c,r)(t)$, $\exists(c,d)$, $\exists(c,r)(t)$, and $\overline{\alpha}$ in $(\forall E)$, $(\forall E_r)$, $(\exists E)$, $(\exists E_r)$, and $(\overline{E})$, respectively, as well

[12]Let us draw the reader's attention to the fact that the standard dichotomy of a *free/bound occurrence of a variable in a formula* does not apply here because all occurrences of each variable in a general sentence are free, or bound by no quantifier. In table 2, neither sentences nor $\perp$ contain variables, whilst in the general sentences *c(x)* and *r(x,y)*, the set term c and the binary relational term r, respectively, contain no occurrences of any variable. Hence, one could have dropped the specification 'free' as redundant.

[13]With regard to the rules that [Moss(2015)] does not demarcate, applications of $(trans)$, the $\wedge$-eliminations, and $(invE)$ eliminate the corresponding terms from a derivation (for instance, r^{-1}, in the case of $(invE)$). Applications of $(invI)$ and $(\wedge I)$, in turn, introduce r^{-1} and $r\wedge s$ into a derivation, respectively.

[14]See its formulation on page 274 below.

[15]Generally, this rule allows one to derive an arbitrary ψ—including $\perp, \phi, \overline{\phi}$—from its premisses [Kurbis(2022)].

as the occurrences of $a(k,j)$ and $a(j,l)$ in (*trans*). The *minor* premiss is the occurrences of *c(t)*, *c(u)*, and α in $(\forall E)$, $(\forall E_r)$, $(\exists E)$, $(\exists E_r)$, and $(\overline{E})$, respectively, as well as the above-line occurrence of α in $(\exists E)$ and $(\exists E_r)$.

Definition 2 is analogical to the standard one in the case of natural deduction for propositional logics, excepting (*trans*), both premisses of which are major by analogy to the rule of *cut* in the case of the sequent calculus there.

The notion of a derivation in each system in question is defined in the standard Gentzen–Prawitz tree-like way with updates [Gentzen(1969), Prawitz(1965), Prawitz(1971), NegriPlato(2001), Kurbis(2022)]. By default, we follow [Prawitz(1971)]. Hence, we skip its definition as well as the ones of the most basic notions related to it: a node, to be above (below), or succeeding (preceding) in a branch (path), etc., while highlighting the following

Definition 3. (An adaptation of 3.2.1. Branches and paths [Prawitz(1971), p. 249]) A *branch* in a derivation is a sequence of (general) sentence occurrences $\alpha_1, \alpha_2, \ldots, \alpha_n$ in the derivation, where α_1 is an assumption not canceled by $(\exists E)$ or $(\exists E_r)$, α_{i+1} is the (general) sentence occurrence immediately below α_i, α_n is either the conclusion of the derivation or the minor premiss of $(\forall E)$, $(\forall E_r)$, and $(\overline{E})$, but no α_i with $i < n$ is such a minor premiss. A *path* is like a branch except that the (general) sentence occurrence immediately succeeding the major premiss of $(\exists E)$ or $(\exists E_r)$ is not the (general) sentence occurrence immediately below, but one of the assumptions canceled by this application of $(\exists E)$ or $(\exists E_r)$.[16]

$\Gamma \vdash_{\mathbf{R}} \phi$ means the existence of a $\mathbf{R}$-derivation of ϕ from a set of uncanceled assumptions Γ, where $\mathbf{R} \in \{\mathbf{R}^{*\dagger}, \mathbf{R}^{*\dagger}_{\mathbf{trans}}, \mathbf{RCA}^{\dagger}(\mathbf{opp})\}$.[17] [Moss(2015)] contains examples of $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-derivations, and

Theorem 1. *([Moss(2015), pp. 563-4])* $\Gamma \vdash_{\mathbf{R}} \phi$ *iff* $\Gamma \models_{\mathcal{R}} \phi$*, where* Γ *is a set of uncanceled assumptions and* $\mathcal{R} \in \{\mathcal{R}^{*\dagger}, \mathcal{R}^{*\dagger}_{trans}, \mathcal{RCA}^{\dagger}(opp)\}$.

The letter Σ with super- and subscripts runs over $\mathbf{R}$-derivations. By default, we mean $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-derivations. Sometimes space reasons make us use the 'horizontal' notation, not the usual 'vertical' one, in presenting derivations (remark 4 below).

If there is an occurrence of a term from the above tables 1 and 2 in ϕ or α, then the term is said to be *from/in* ϕ or α. While using the variable denotations from tables 1 and 2 above, let us present

[16] Sometimes *path* is called *thread* [Plato(2012), p. 208]. Sometimes *branch* in the meaning of *path*, or *thread* is introduced only [Kurbis(2022), p. 123].

[17] Let us remind the readers that Γ is a set of sentences rather than general sentences.

Definition 4. (Subterm *st* of a term, of a sentence, of a general sentence)

- *st* of p, a, t', k, x is only p, a, t', k, x, respectively.
- *sts* of $\bar{l}$ are $\bar{l}$ and *sts* of l.
- *sts* of $l \wedge m$ are $l \wedge m$, *sts* of l and of m.
- *sts* of $r \wedge s$ are $r \wedge s$, *sts* of r and of s.
- *sts* of r^{-1}, $\overline{r}$ are r^{-1}, $\overline{r}$, respectively, and *sts* of r.
- *sts* of $\overline{r \wedge s}$ are $\overline{r \wedge s}$, *sts* of $\overline{r}$ and of $\overline{s}$.
- *sts* of $\exists(c, r)$, $\forall(c, r)$ are $\exists(c, r)$, $\forall(c, r)$, respectively, *sts* of c and of r.
- *sts* of $\exists(c, d)$, $\forall(c, d)$ are *sts* of c and of d.
- *sts* of *c(k)*, *r(k,j)* are *sts* of c, of r, j, respectively, and k.
- *sts* of *c(x)*, *r(x,y)* are *sts* of c, of r, y, respectively, and x.
- *st* of $\bot$ is only $\bot$.

Remark 3. First of all, let us remind the readers that the identification in definition 1 does not allow 'double negation' expressions. While making the 'passive voice construction' more complex, from the standpoint of the subterm property, we are governed by its plain higher syntactic complexity that $(InvI)$ and $(InvE)$ above reflect rather than by certain grammatical arguments. Due to the fact that definition 1 does not work *inside* $r \wedge s$, we are to 'apply' the respective De Morgan law to it. Third, the bar notation makes certain (general) sentences equal, which means they have the same subterms.[18] At last, neither *c(x)* nor *d(x)* are subterms of $\exists(c, d)$ because x is not in the latter.[19]

Definition 5. (Degree *dg* of a term, of a sentence, of a general sentence)

- $dg(p) = dg(a) = dg(t') = dg(k) = dg(x) = 1$.
- $dg(\bar{l}) = dg(l) + 1$.
- $dg(l \wedge m) = dg(l) + dg(m) + 1$.
- $dg(\overline{r}) = dg(r) + 1$.
- $dg(r^{-1}) = dg(r) + 1$.

[18] It is unlike *FOL*, where $\overline{\forall x P(x)}$ and $\exists x \overline{P(x)}$ do not have the same subformulae.

[19] In the other 'quantifier' rules, the situation is analogous. This is in quite a contrast with the following situation, 'In a step of indirect inference in natural deduction, a negative temporary assumption $\neg A$ is closed and A concluded. When assumptions are closed otherwise, they are subformulas of the conclusion, as in implication introduction, or of the major premiss, as in those elimination rules that close assumptions, and the subformula property of normal derivations follows' [Plato(2012), p. 205].

- $dg(\exists(c,r)) = dg(\forall(c,r)) = dg(c) + dg(r) + 1$.
- $dg(r \wedge s) = dg(r) + dg(s) + 1$.
- $dg(\exists(c,d)) = dg(\forall(c,d)) = dg(c) + dg(d) + 1$.
- $dg(c(k)) = dg(c) + dg(k) + 1$.
- $dg(r(k,j)) = dg(r) + dg(k) + dg(j) + 1$.
- $dg(\overline{\phi}) = dg(\phi) + 1$.
- $dg(c(x)) = dg(c) + dg(x) + 1$.
- $dg(r(x,y)) = dg(r) + dg(x) + dg(y) + 1$.
- $dg(\overline{\alpha}) = dg(\alpha) + 1$.
- $dg(\bot) = 0$.

Let us remind the readers again about the identification in definition 1.

Definition 6. (An adaptation of 3.1.1. Maximum formulas [Prawitz(1971), p. 248]) A (general) sentence occurrence in a derivation that stands at the same time as the conclusion of an introduction and as the major premiss of an elimination is called a *maximal (general) sentence*.

Definition 7. (An adaptation of 3.1.2. Maximum segments [Prawitz(1971), p. 248]) A sequence of (general) sentence occurrences $\alpha_1, \alpha_2, \ldots, \alpha_n$ in a derivation is said to be a *maximal segment* in the derivation when they are of the same form, α_{i+1} stands immediately below α_i, α_1 is the conclusion of an introduction rule, and α_n is the major premiss of an elimination rule (α_i, for $i < n$, is then a minor premiss of $(\forall E)$, $(\exists E)$, $(\forall E_r)$, $(\exists E_r)$, or $(\overline{E})$).

The *length* of a maximal segment is the parameter n. The degree of α is the *degree of a maximal segment* that α forms, or consists of α.

Definition 8. (An adaptation of 3.1.3. Normal derivations [Prawitz(1971), p. 248]) A derivation in any of the systems considered above is now defined as *normal* or said to be in *normal form* when it contains no maximal (general) sentence and no maximal segment.

3 Detour and permutation conversions for $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, $\mathbf{RCA}^{\dagger}(\mathbf{opp})$

In section 3, we consecutively consider detour and permutation conversions for $\mathbf{R}^{*\dagger}$, then for $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, and finally for $\mathbf{RCA}^{\dagger}(\mathbf{opp})$.

Detour conversions for $\mathbf{R}^{*\dagger}$ 5 introduction and 5 elimination $\mathbf{R}^{*\dagger}$-rules make 25 detours, at most. Detours $(\exists I)-(\forall E)$, $(\exists I)-(\forall E_r)$, and $(\exists I)-(\exists E_r)$ are impossible because no major premisses of $(\forall E)$, $(\forall E_r)$, and $(\exists E_r)$, respectively, are derivable via $(\exists I)$. The analogy holds for their $(\exists I_r)$-counterparts. $(\forall I)-(\exists E)$, $(\forall I)-(\forall E_r)$, and $(\forall I)-(\exists E_r)$ are impossible because no major premisses of $(\exists E)$, $(\forall E_r)$, and $(\exists E_r)$, respectively, are derivable via $(\forall I)$. The analogy holds for their $(\forall I_r)$-counterparts. There are 13 detours left. We start with 8 detours without $(\overline{E})$ and then consider 5 with it.

3.1 $(\forall I)-(\forall E)$. $(\forall I_r)-(\forall E_r)$ is analogous.

$$\dfrac{\dfrac{\begin{array}{c}[c(x)]^1\\ \Sigma_1\\ d(x)\end{array}}{\forall(c,d)}\forall I^1 \qquad \begin{array}{c}\Sigma_2\\ c(t)\end{array}}{d(t)}\forall E \quad \dashrightarrow \quad \begin{array}{c}\Sigma_2\\ c(t)\\ \Sigma_1^{x/t}\\ d(t)\end{array}$$

Remark 4. $\Sigma_1^{x/t}$ results from Σ_1 via replacing the occurrences of x with t. Let us show the correctness of $\langle c(t), \Sigma_1^{x/t}, d(t)\rangle$. It is obvious in the case of applications of $(\forall E)$, $(\exists I)$, and $(\overline{E})$. Any application of $(\forall I)$, $(\exists E)$, and $(\overline{I})$ remains correct because the latter applies to sentences only, whilst applying the former ones needs other variables than x. At last, there are no occurrences of x in the uncanceled assumptions in Σ_1, by theorem 1.

3.2 $(\exists I)-(\exists E)$. $(\exists I_r)-(\exists E_r)$ is analogous.

$$\dfrac{\dfrac{\begin{array}{cc}\Sigma_1 & \Sigma_2\\ c(t) & d(t)\end{array}}{\exists(c,d)}\exists I \qquad \dfrac{\begin{array}{c}[c(x)]^1\\ \Sigma_3\\ \alpha\end{array}}{\alpha}}{\alpha}\exists E^1 \quad \dashrightarrow \quad \begin{array}{cc}\Sigma_1^{t/x} & \Sigma_2^{t/x}\\ c(x) & d(x)\\ \multicolumn{2}{c}{\Sigma_3}\\ \multicolumn{2}{c}{\alpha}\end{array}$$

The correctness of $\Sigma_1^{t/x}$ and $\Sigma_2^{t/x}$ is shown by analogy with that in remark 4.

3.3 $(\overline{I}) - (\forall E)$. $(\overline{I}) - (\forall E_r)$ is analogous.

3.3.1 General case of $(\overline{I}) - (\forall E)$.

$$\dfrac{\dfrac{[\overline{\forall(c,d)}]^1}{\Sigma_1}}{\overline{I}^1\,\dfrac{\bot}{\forall(c,d)}} \qquad \dfrac{\Sigma_2}{c(t)}$$
$$\forall E\,\dfrac{\qquad\qquad}{d(t)}$$

$$\dashrightarrow$$

$$\forall E\,\dfrac{\dfrac{\Sigma_2}{[\forall(c,d)]^1 \qquad c(t)}}{\overline{E}\,\dfrac{d(t) \qquad\qquad [\overline{d(t)}]^2}{\overline{I}^1\,\dfrac{\bot}{\overline{\forall(c,d)}}}}$$
$$\dfrac{\Sigma_1}{\overline{I}^2\,\dfrac{\bot}{d(t)}}$$

Remark 5. There are three things worth mentioning. First, having been proposed in [Plato(2012)] to solve the notoriously hard normalization problem for the Gentzen–Prawitz natural deduction system [Gentzen(1969), Prawitz(1965)], the technique applied in detour 3.3 is known under the name of *classicizing* [PelletierHazen(2023)]. The cost of classicizing solutions is that the *subformula property* must be understood more broadly in order to include negations of the formulae under consideration. Second, having been thus classicized, the normal derivation obtains the property desired (theorem 2 and corollary 11 below) in propositional and modal cases.[20] Third, it is possible that, in the initial derivation, the application of $(\overline{I})$ cancels more than one occurrence of the assumption $\overline{\forall(c,d)}$, i.e., the ones in Σ_1, which the uppermost application of $(\overline{I})$ cannot cancel in the derivation converted. In such a case, one replaces each occurrence of $\overline{\forall(c,d)}$ in Σ_1 in the derivation converted by its derivation there. The strategy from definition 10 below guarantees that in Σ_1, there is no maximal α such that $dg(\alpha) \geq dg(\overline{\forall(c,d)})$.

3.3.2 Special case of $(\overline{I}) - (\forall E)$.

A special case of detour 3.3 announced in chapter 1 above happens if in the derivation converted, the occurrence of $\overline{\forall(c,d)}$ is the major premiss of an application of $(\overline{E})$—the only elimination rule it could be the major premiss of. Hence, $\overline{\forall(c,d)}$—a sentence that is not maximal in the initial derivation—becomes maximal in the derivation converted, and an unwanted situation happens that detour 3.3 increases the rank of a derivation (definition 9

[20] 'For normal derivability and the subformula property, it is sufficient to permute down instances of indirect proof whenever they have been used for concluding a major premiss of an elimination rule' [Plato(2012), abstract]. The Plato–Siders rule of indirect proof cited is $(\overline{I})$, by and large.

below).[21] In the special case, the unwanted variant of detour 3.3 looks as follows, where Σ_1 is $\langle \Sigma_1^*, \forall(c,d), \bot, \Sigma_1^{**} \rangle$.

$$\forall E\,\dfrac{\overline{I}^1\,\dfrac{\begin{matrix}\overline{E}\,\dfrac{[\overline{\forall(c,d)}]^1 \quad \begin{matrix}\Sigma_1^*\\ \forall(c,d)\end{matrix}}{\bot}\\ \Sigma_1^{**}\\ \bot\end{matrix}}{\forall(c,d)} \quad \begin{matrix}\Sigma_2\\ c(t)\end{matrix}}{d(t)} \quad \dashrightarrow$$

$$\overline{I}^2\,\dfrac{\begin{matrix}\overline{E}\,\dfrac{\overline{I}^1\,\dfrac{\overline{E}\,\dfrac{\forall E\,\dfrac{[\forall(c,d)]^1 \quad \begin{matrix}\Sigma_2\\ c(t)\end{matrix}}{d(t)} \quad [\overline{d(t)}]^2}{\bot}}{\overline{\forall(c,d)}} \quad \begin{matrix}\Sigma_1^*\\ \forall(c,d)\end{matrix}}{\bot}\\ \Sigma_1^{**}\\ \bot\end{matrix}}{d(t)}$$

A straightforward solution to get rid of the *unwanted* variant of the derivation converted would have been its *shortened* variant:

$$\forall E\,\dfrac{\begin{matrix}\Sigma_1^*\\ \forall(c,d)\end{matrix} \quad \begin{matrix}\Sigma_2\\ c(t)\end{matrix}}{d(t)}$$

However, Σ_1^* might have the occurrences of the assumption $\overline{\forall(c,d)}$ such that the application of $(\overline{I})$ discharges them in the initial derivation, which implies the incorrectness of the shortened variant as well. So, an updated version of the special case is as follows, where Σ_1^* is $\langle \overline{\forall(c,d)}, \Sigma_1^{***} \rangle$, and there are no occurrences of the assumption $\overline{\forall(c,d)}$ in Σ_1^{***}.

[21] $dg(\forall(c,d)) < dg(\overline{\forall(c,d)})$, according to definition 5.

$$\dfrac{\overline{E}\;\dfrac{\dfrac{\dfrac{}{\overline{\forall(c,d)}}^{1}}{\Sigma_1^{***}}}{\dfrac{\dfrac{\dfrac{}{\overline{\forall(c,d)}}^{1}}{\bot}\;\forall(c,d)}{\Sigma_1^{**}}}}{\forall E\;\dfrac{\overline{I}^{1}\;\dfrac{\bot}{\forall(c,d)}}{d(t)}\qquad\dfrac{\Sigma_2}{c(t)}}$$

Now, the permutation proceeds in two steps. First, one gets rid of the occurrences of the assumption $\overline{\forall(c,d)}$ in Σ_1^* in the unwanted variant of the initial derivation. Second, the shortened variant with Σ_1^{***} of the derivation converted is obtained.

$$\dfrac{\overline{E}\;\dfrac{\dfrac{\dfrac{}{\overline{\forall(c,d)}}^{1}}{\Sigma_1^{***}}}{\dfrac{\dfrac{}{\overline{\forall(c,d)}}^{1}\qquad\overline{I}^{1}\;\dfrac{\forall(c,d)}{\forall(c,d)}}{\bot}}{\Sigma_1^{**}}}{\forall E\;\dfrac{\overline{I}^{1}\;\dfrac{\bot}{\forall(c,d)}}{d(t)}\qquad\dfrac{\Sigma_2}{c(t)}}\quad\dashrightarrow\quad\forall E\;\dfrac{\dfrac{\Sigma_1^{***}}{\forall(c,d)}\qquad\dfrac{\Sigma_2}{c(t)}}{d(t)}$$

3.4 $(\overline{I}) - (\exists E)$. $(\overline{I}) - (\exists E_r)$ is analogous.

$$\dfrac{\exists E^{2}\;\dfrac{\dfrac{\dfrac{}{\overline{\exists(c,d)}}^{1}}{\Sigma_1}\qquad[c(x)]^{2}\qquad[d(x)]^{2}}{\overline{I}^{1}\;\dfrac{\bot}{\exists(c,d)}\qquad\dfrac{\Sigma_2}{\alpha}}}{\alpha}\quad\dashrightarrow$$

$$\dfrac{\exists E^{2}\;\dfrac{[\exists(c,d)]^{1}\qquad\dfrac{\Sigma_2}{\alpha}}{\alpha}}{\overline{E}\;\dfrac{\alpha\qquad\qquad[\overline{\alpha}]^{3}}{\overline{I}^{1}\;\dfrac{\bot}{\overline{\exists(c,d)}}}}\qquad\dfrac{[c(x)]^{2}\qquad[d(x)]^{2}}{\Sigma_1}$$

$$\overline{I}^{3}\;\dfrac{\bot}{\alpha}$$

See remark 5, detours 3.3.1 and 3.3.2 above.

Let us consider 5 detours with $(\overline{E})$.

3.5 $(\overline{I}) - (\overline{E})$

$$\dfrac{\dfrac{\begin{array}{c}[\phi]^1\\ \Sigma_1\\ \bot\end{array}}{\overline{\overline{\phi}}}\overline{I}^1 \qquad \Sigma_2}{\phi}\overline{E} \quad \dashrightarrow \quad \begin{array}{c}\Sigma_2\\ \phi\\ \Sigma_1\\ \bot\end{array}$$

Notice that, in the initial derivation, $\overline{\overline{\phi}} = \phi$, according to definition 1 (see also remark 1).[22] Notice also that if, in the initial derivation, the application of $(\overline{I})$ cancels more than one occurrence of ϕ in Σ_1, then, in the converted derivation, all of them are replaced with the subderivation $\langle \Sigma_2, \phi \rangle$.

3.6 $(\forall I) - (\overline{E})$**.** $(\forall I_r) - (\overline{E})$ **is analogous.**

$$\dfrac{\dfrac{\begin{array}{c}[c(x)]^1\\ \Sigma_1\\ \overline{d}(x)\end{array}}{\forall(c,\overline{d})}\forall I^1 \qquad \Sigma_2}{\bot}\overline{E} \quad \dashrightarrow$$

$$\dfrac{\dfrac{\begin{array}{c}\Sigma_2\\ \exists(c,d)\end{array} \qquad \begin{array}{c}[c(x)]^1\\ c(x)\end{array} \qquad \begin{array}{c}[d(x)]^1\\ d(x)\end{array}}{\bot}\exists E^1 \qquad \begin{array}{c}\Sigma_2\\ \exists(c,d)\end{array} \qquad \begin{array}{c}[c(x)]^2\\ c(x)\end{array} \qquad \begin{array}{c}[d(x)]^2\\ d(x)\end{array}}{\bot}\exists E^2$$

Notice that $\overline{\exists(c,d)} = \forall(c,\overline{d})$, according to definition 1.

3.7 $(\exists I) - (\overline{E})$**.** $(\exists I_r) - (\overline{E})$ **is analogous.**

$$\dfrac{\dfrac{\begin{array}{c}\Sigma_1\\ c(t)\end{array} \qquad \begin{array}{c}\Sigma_2\\ \overline{d}(t)\end{array}}{\exists(c,\overline{d})}\exists I \qquad \begin{array}{c}\Sigma_3\\ \forall(c,d)\end{array}}{\bot}\overline{E} \quad \dashrightarrow \quad \dfrac{\dfrac{\begin{array}{c}\Sigma_3\\ \forall(c,d)\end{array} \qquad \begin{array}{c}\Sigma_1\\ c(t)\end{array}}{d(t)}\forall E \qquad \begin{array}{c}\Sigma_2\\ \overline{d}(t)\end{array}}{\bot}\overline{E}$$

Notice that $\exists(c,\overline{d}) = \overline{\forall(c,d)}$, according to definition 1.

[22] Notice also that $(\overline{I})$ applies to sentences only.

Permutation conversions for $\mathbf{R}^{*\dagger}$ Traditionally, permutation conversions remove a maximal segment (definition 7), or a *hidden* detour that relates to applications of the elimination rules, such as $(\exists E)$ and $(\exists E_r)$, whose conclusions are arbitrary formulae.[23] However, one of the specifics of the present systems announced in chapter 1 above is that maximal segments could relate to applications of some of the other eliminations as well.

5 introduction $\mathbf{R}^{*\dagger}$-rules make 25 permutations, at most. $(\forall E) - (\exists E)$ and $(\forall E) - (\forall E)$ are impossible because there is no (general) sentence such that, in an application of $(\forall E)$, one of its occurrences is its minor premiss and another one is its conclusion. Their $(\forall E_r)$-counterparts are analogous. $(\overline{E}) - (\forall E)$, $(\overline{E}) - (\exists E)$, $(\overline{E}) - (\exists E_r)$, and $(\overline{E}) - (\forall E_r)$ are impossible because $\perp$ cannot be major premisses of $(\forall E)$, $(\forall E_r)$, $(\exists E)$, and $(\exists E_r)$, respectively. So, there are 17 permutations left.

3.8 $(\overline{E}) - (\overline{E})$

$$\overline{E}\,\dfrac{\overline{E}\,\dfrac{\begin{matrix}\Sigma_2\\ \perp\end{matrix}\qquad\begin{matrix}\Sigma_1\\ \overline{\perp}\end{matrix}}{\perp}\qquad\begin{matrix}\Sigma_3\\ \overline{\perp}\end{matrix}}{\perp}\quad\dashrightarrow\quad\begin{matrix}\Sigma_2\\ \perp\end{matrix}$$

Notice that in both conversions, $\overline{\overline{\perp}} = \perp$, according to definition 1, and the valid $\overline{\perp}$ is a theorem in the syllogistics treated and therefore is skipped. (See also the derivation below corollary 2.2 to follow.)

3.9 $(\forall E) - (\forall E_r)$

$$\forall E\,\dfrac{\forall E_r\,\dfrac{\begin{matrix}\Sigma_2\\ \forall(\forall(c,r),\forall(c,r))\end{matrix}\qquad\begin{matrix}\Sigma_1\\ \forall(c,r)(t)\end{matrix}}{\forall(c,r)(t)}\qquad\begin{matrix}\Sigma_3\\ c(u)\end{matrix}}{r(t,u)}\quad\dashrightarrow\quad\forall E_r\,\dfrac{\begin{matrix}\Sigma_1\\ \forall(c,r)(t)\end{matrix}\qquad\begin{matrix}\Sigma_3\\ c(u)\end{matrix}}{r(t,u)}$$

$(\forall E) - (\overline{E})$, $(\forall E) - (\exists E_r)$, $(\forall E_r) - (\overline{E})$, $(\forall E_r) - (\forall E_r)$, and $(\forall E_r) - (\exists E_r)$ are analogous.

[23] A powerful line of research concerns natural deduction systems with the so-called *general*—both elimination and introduction—rules only, whose conclusions are just arbitrary formulae [Segerberg(1983), Prawitz(1971), Plato(2012), Milne(2015), Kurbis(2022), Shangin(2023), Geuvers(2024)].

3.10 $(\exists E) - (\exists E)$

$$\dfrac{\dfrac{\Sigma_3 \quad \exists(c_1,d_1) \qquad \dfrac{[c_1(y)]^1 \quad [d_1(y)]^1}{\Sigma_1 \quad \exists(c,d)}}{\exists E^1 \dfrac{\exists(c,d)}{\alpha}} \qquad \dfrac{[c(x)]^2 \quad [d(x)]^2}{\Sigma_2 \quad \alpha}}{\exists E^2 \dfrac{}{\alpha}} \quad \dashrightarrow$$

$$\dfrac{\Sigma_3 \quad \exists(c_1,d_1) \qquad \exists E^2 \dfrac{\dfrac{[c_1(y)]^1 \quad [d_1(y)]^1}{\Sigma_1 \quad \exists(c,d)} \qquad \dfrac{[c(x)]^2 \quad [d(x)]^2}{\Sigma_2 \quad \alpha}}{\alpha}}{\exists E^2 \dfrac{}{\alpha}}$$

$(\exists E) - (\forall E)$, $(\exists E) - (\overline{E})$, $(\exists E) - (\forall E_r)$, $(\exists E) - (\exists E_r)$, $(\exists E_r) - (\forall E)$, $(\exists E_r) - (\exists E)$, $(\exists E_r) - (\overline{E})$, $(\exists E_r) - (\forall E_r)$, and $(\exists E_r) - (\exists E_r)$ are analogous.

Considering the detour and permutation conversions for $\mathbf{R}^{*\dagger}$—the weakest syllogistic in this paper—ends here. Now, we start doing the ones for $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$.

Detour conversions for $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$ The new elimination $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$-rule (*trans*) adds 5 detours to the above ones, at most. 4 of them are impossible because no premiss of (*trans*) can be derived via $(\forall I)$, $(\exists I)$, $(\forall I_r)$, or $(\exists I_r)$. Besides, let us remind the readers of the specifics of (*trans*) on page 258—both premisses of it are major. Hence, there are 3 variants for detour 3.11 below: both occurrences of its premisses are maximal, and only one occurrence of them is maximal.

3.11 $(\overline{I}) - (trans)$**. The other two variants are analogous.**

$$\dfrac{\dfrac{\dfrac{[\overline{a(k,j)}]^1}{\Sigma_1}}{\overline{I}^1\,\dfrac{\bot}{a(k,j)}} \quad \dfrac{\dfrac{[\overline{a(j,l)}]^2}{\Sigma_2}}{\overline{I}^2\,\dfrac{\bot}{a(j,l)}}}{trans\,\dfrac{}{a(k,l)}} \dashrightarrow$$

$$\dfrac{\dfrac{[a(k,j)]^1 \quad [a(j,l)]^2}{trans\,\dfrac{}{a(k,l)}} \quad [\overline{a(k,l)}]^3}{\overline{E}\,\dfrac{}{\overline{I}^1\,\dfrac{\bot}{\overline{a(k,j)}}}}$$

$$\dfrac{\dfrac{\dfrac{\Sigma_1}{\overline{I}^2\,\dfrac{\bot}{\overline{a(j,l)}}}}{\Sigma_2}}{\overline{I}^3\,\dfrac{\bot}{a(k,l)}}$$

Detour 3.11 is essentially detour 3.3 above. See also remark 5 above.

Permutation conversions for $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$ The new elimination $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$-rule $(trans)$ adds 11 permutations to the ones above, at most. $(trans) - (\forall E)$, $(trans) - (\forall E_r)$, $(trans) - (\exists E)$, and $(trans) - (\exists E_r)$ are impossible because no major premisses of $(\forall E)$, $(\forall E_r)$, $(\exists E)$, or $(\exists E_r)$, respectively, can be derived via $(trans)$. $(\forall E) - (trans)$ is impossible because no premiss of $(trans)$ can be derived via $(\forall E)$. $(\forall E_r) - (trans)$ is impossible because there is no sentence such that one of its occurrences is a premiss of $(\forall E_r)$ and another one is a premiss of $(trans)$. $(\overline{E}) - (trans)$ is impossible because $\bot$ cannot be a premiss of $(trans)$. $(trans) - (\overline{E})$ is impossible because no premiss of $(trans)$ can be of the form $\overline{a}(k,j)$. At last, $(trans) - (trans)$ is impossible because $(trans)$ does not have the minor premiss. So, there are 2 permutations left.

As in detour 3.11 above, due to the specifics of $(trans)$, there are 3 variants for permutation 3.12 below: both premiss occurrences in $(trans)$ form different maximal segments, and only one of them does.

3.12 $(\exists E) - (trans)$**. The other two variants as well as** $(\exists E_r) - (trans)$ **are analogous.**

$$\exists E^1\,\dfrac{\exists(c_1,d_1)\quad\dfrac{[c_1(y)]^1\quad[d_1(y)]^1}{\Sigma_3}\qquad a(k,j)}{trans\,\dfrac{a(k,j)\qquad\dfrac{\exists(c,d)\quad\dfrac{[c(x)]^2\quad[d(x)]^2}{\Sigma_4}\qquad a(j,l)}{\exists E^2\,\dfrac{}{a(j,l)}}\Sigma_2}{a(k,l)}\Sigma_1}{a(k,l)}$$

$$\dashrightarrow$$

$$\exists E^1\,\dfrac{\exists(c_1,d_1)\quad\dfrac{\exists(c,d)\quad\dfrac{[c_1(y)]^1\quad[d_1(y)]^1}{\Sigma_4}\qquad\dfrac{[c(x)]^2\quad[d(x)]^2}{\Sigma_2}}{\exists E^2\,\dfrac{}{a(k,l)}}\Sigma_3\qquad trans\,\dfrac{a(k,j)\qquad a(j,l)}{a(k,l)}\Sigma_1}{a(k,l)}$$

Considering the detour and permutation conversions for $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$—the intermediate system in this paper—ends here. Now, we start doing the ones for $\mathbf{RCA}^{\dagger}(\mathbf{opp})$.

Detour conversions for $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ 2 new introduction and 3 new elimination $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-rules add 33 detours to the ones above, at most. 12 detours with $(\exists I)$, $(\forall I)$, $(\exists I_r)$, and $(\forall I_r)$ are impossible because their conclusions cannot be major premisses of 3 new elimination $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-rules. There are 9 detours with $(\wedge I)$, of which 2 are possible only because the major premisses of $(\wedge E_1)$ and $(\wedge E_2)$ can be derived via $(\wedge I)$, trivially. There are 9 detours with $(invI)$, of which 3 are possible only because the major premisses of $(\wedge E_1)$, $(\wedge E_2)$, and $(invE)$ can be derived via $(invI)$, where $(invI) - (invE)$ is trivial. So, there are 8 detours left, of which 3 are already said to be trivial and therefore skipped below.

3.13 $(\mathbf{invI}) - (\wedge\mathbf{E_1})$**.** $(invI) - (\wedge E_2)$ **is analogous**

$$invI\,\dfrac{\wedge E_1\,\dfrac{\dfrac{(r\wedge s)(j,k)}{\Sigma_1}}{(r\wedge s)^{-1}(k,j)}}{r^{-1}(k,j)}\quad\dashrightarrow\quad\wedge E_1\,\dfrac{\dfrac{(r\wedge s)(j,k)}{\Sigma_1}}{invI\,\dfrac{r(j,k)}{r^{-1}(k,j)}}$$

3.14 $(\overline{I})-(invE)$**.** $(\overline{I})-(\wedge E_1)$**,** $(\overline{I})-(\wedge E_2)$ **are analogous.**

$$\begin{array}{c}[(\overline{r})^{-1}(k,j)]^1\\ \Sigma_1\\ \overline{I}^1\,\dfrac{\bot}{r^{-1}(k,j)}\\ invE\,\dfrac{}{r(j,k)}\end{array}\quad\dashrightarrow\quad\begin{array}{c}\overline{E}\,\dfrac{[\overline{r}(j,k)]^1\qquad invE\,\dfrac{[r^{-1}(k,j)]^2}{r(j,k)}}{\overline{I}^2\,\dfrac{\bot}{(\overline{r})^{-1}(k,j)}}\\ \Sigma_1\\ \overline{I}^1\,\dfrac{\bot}{r(j,k)}\end{array}$$

Detour 3.14 is essentially detour 3.3 above. See also remark 5 above.

Permutation conversions for $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ 3 new elimination $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-rules add 45 permutations to the ones above. To save some space, we employ abbreviations for them in this passage. $(\wedge E_1)-$, $(\wedge E_2)-$, $(invE)-$, and $(trans)-$ are 30 impossible permutations because these rules have no minor premisses. $(\overline{E})-$ are 3 impossible permutations because no elimination has the premiss $\bot$. $(\forall E)-$ and $(\forall E_r)-$ are 6 impossible permutations because their minor premisses cannot be premisses of 3 new elimination $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-rules. So, $(\exists E)-$ and $(\exists E_r)-$ are 6 possible permutations, which are particular cases of permutation 3.10 above and therefore are skipped.

Considering the detour and permutation conversions for $\mathbf{RCA}^{\dagger}(\mathbf{opp})$—the strongest system in this paper—ends here.

Simplification conversions for each of the syllogistics treated Traditionally, one also considers simplification conversions, or immediate simplifications [Prawitz(1971), p. 254], whose purpose is to 'get rid of redundant applications' of rules canceling assumptions. Due to the fact that such rules in our systems are conventional, we left these conversions to the readers.

The main results of this paper

Definition 9. (Rank of a derivation [Kurbis(2022), definition 8]) The *rank* of a derivation is the pair $\langle e,l\rangle$, where e is the highest degree of a maximal sentence or maximal segment in it or 0 if there is none, and l is the sum of the lengths of maximal segments of the highest degree and the number of maximal sentences in the derivation. $\langle e,l\rangle < \langle e',l'\rangle$ iff either (i) $e<e'$ or (ii) $e=e'$ and $l<l'$.

While tackling the normalization problem, we stick to the following

Definition 10. (An updated adaptation of the *strategy* in [Kurbis(2022), p. 122])[24] To reduce the rank of the derivation, the detour (permutation) conversions to follow apply to a maximal sentence (segment) of highest degree such that no maximal sentence (segment) of highest degree stands above it in the derivation. The permutation conversions are of priority.

Theorem 2. $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ *normalize.*

Proof. It follows by induction on the rank of a derivation (definition 9). First, being governed by the strategy above, the detour conversions from section 3 imply $e < e'$.[25] Second, in the case $e = e'$, the detour conversions from section 3 either imply $l < l'$ or simplify a derivation (for instance, detour 3.2). The same reasoning holds for the permutation conversions from section 3 as well. □

Both corollaries to follow are anticipated and therefore are stated without proofs.

Corollary 2.1. *(Form of a normal* $\mathbf{R}^{*\dagger}$*-,* $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$*-,* $\mathbf{RCA}^{\dagger}(\mathbf{opp})$*-derivation) In a normal* $\mathbf{R}^{*\dagger}$*-,* $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$*-,* $\mathbf{RCA}^{\dagger}(\mathbf{opp})$*-derivation, all major premises of eliminations are (canceled or uncanceled) assumptions of it.*

Corollary 2.2. *Any branch in a normal derivation consists of three consecutive parts: (possibly, empty) E-part that consists of applications of the eliminations only; the minimal sentence; (possibly, empty) I-part that consists of applications of the introductions only.*

Corollary 2.2 implies that no path in a normal derivation has more than one application of $(\overline{I})$ because $\overline{\bot}$ is $R^{*\dagger}$-, $R^{*\dagger}_{trans}$-, $RCA^{\dagger}(opp)$-valid, and therefore is a $\mathbf{R}^{*\dagger}$-, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$-, $\mathbf{RCA}^{\dagger}(\mathbf{opp})$-theorem (theorem 1). Hence, it is skipped.

$$\dfrac{\dfrac{[\overline{\phi}]^1}{\Sigma_1} \quad \dfrac{[\overline{\bot}]^2}{\Sigma_2}}{\dfrac{\dfrac{\bot}{\overline{I}^2}}{\dfrac{\bot}{\overline{I}^1}}} \dashrightarrow \dfrac{[\overline{\phi}]^1}{\dfrac{\Sigma_1}{\dfrac{\Sigma_2}{\dfrac{\bot}{\overline{I}^1}}}}$$

[24] To put it with Prawitz, 'Although it is true that new maximal formulas and maximal segments may arise by a reduction, it can be seen that the reductions always can be *chosen* so that the formulas occurring as new maximal formulas or occurring in new maximal segments are of less degree than the removed one. Thus, we can use an induction over these measures' [Prawitz(1971), p. 256; the *italics* originally 'choosen'].

[25] Notice that it also holds in the special cases of detours 3.3 and 3.4, when the usual respective detour conversions lead to the unwanted $e' > e$.

Notice that in contrast to a path, a branch might have more than one application of $(\overline{I})$.[26]

It is hopeless to expect each term in an (even normal) derivation to be a subterm (definition 4) of its uncanceled assumptions or its conclusion simply because in order to prove that a sentence follows from a set of sentences, one could use general sentences—the peculiar *scaffolding* of the syllogistic languages considered. The individual variables, the individual terms composed from those variables, and the general sentences composed of them both, including $\bot$, which is not a sentence—all these expressions from table 2 are absent in the entailments one could prove in the systems considered (the syntax of these entailments is in table 1 only). Yet, normal derivations are closed with respect to subterms and constants, which we state as a particular case of the subterm property.

Definition 11. (Sub-set-term-constant property) A normal $\mathbf{R}^{*\dagger}$-, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$-, $\mathbf{RCA}^{\dagger}$(**opp**)-derivation of ϕ from Γ has the *sub-set-term-constant property* iff in it, each set term c is a subterm of $\overline{d}$, where a set term d is a subterm of ϕ and/or Γ, and each constant k is a constant of ϕ and/or Γ.

Sub-set-term-constant is not to be understood as related to the meaningless *subset of a term* or *subset of a constant*.

Definition 12. (An adaptation of definition 11. Order of branches [Kurbis(2022), p. 123]) A path is *of order 0* if its last sentence is the conclusion of the derivation; it is *of order* $n+1$ if its last sentence is the minor premiss of $(\forall E)$, $(\forall E_r)$, $(\exists E)$, $(\exists E_r)$, or $(\overline{E})$ such that its major premiss is on a path of order n. A path of order 0 is also called a *main path* in the derivation.

Corollary 2.3. $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, $\mathbf{RCA}^{\dagger}$(**opp**) *have the sub-set-term-constant property.*

Proof. It goes by induction on the order of paths in a normal derivation and starts with a main path. If its E-part starts with a canceled assumption, then its I-part has the conclusion of the respective application of an introduction, whose subterms and constants satisfy definition 11, by introspection

[26] 'There is at most one instance of rule *DN* (i.e., our $(\overline{I})$—the author) along any thread of a normal derivation, for otherwise there would be an instance of DN with the conclusion $\bot$. The assumption closed by *DN*, however, would then be the provable formula $\bot \supset \bot$ (i.e., our $\overline{\bot}$—the author). Therefore, as in [Prawitz(1965)], steps of indirect proof separate a part with E-rules (i.e., our eliminations—the author) from a successive part with I-rules (i.e., our introductions—the author), save in threads instead of derivation branches' [Plato(2012), p. 208].

of the introductions. If its E-part starts with an uncanceled assumption, then it is the major one, by definition 3. Then its E-part has the conclusion of the respective application of an elimination, whose set terms and constants satisfy definition 11, by introspection of the eliminations excepting $(\overline{E})$, $(\exists E)$, and $(\exists E_r)$.

In the former case, $\bot$—which is not a sentence—is the minimal sentence of the main path in question, and the main path continues with an application of $(\overline{I})$, as is said above.[27] In the latter cases, the major premisses of $(\exists E)$ and $(\exists E_r)$ might not share set terms or constants with the conclusion α of these rules. However, the Prawitz trick of definition 3 is to guarantee that the set terms and constants in α satisfy definition 11 with respect to the assumptions that the applications of these rules cancel. And these assumptions, in turn, satisfy definition 11 with respect to the major premisses of the applications considered.

At last, according to the bar-notation (definition 1, remark 3), $\overline{\exists(l, r)} = \forall(l, \overline{r})$, which means the subterms of the former are the ones of the latter, and vice versa. The same applies to the other equations there. □

4 Related Work

In section 4, we confine ourselves to some works that combine syllogistics and the normalization problem rather than outline the colossal massive of the literature concerning these two topics separately. To the best of our knowledge, Moss and his coauthors have never discussed the normalization problem for the syllogistic natural deduction systems they propose. However, there are believed to be two *glimpses* worth relating here.

The first one is in [Moss(2010a), pp. 12–18], where Moss *explicitly* discusses a normal form of a derivation and *implicitly* discusses the normalization problem. The language he uses is a fragment of the present one. The main purpose of the normalization procedure proposed is, in unified notation,

Theorem 3. *(Theorem 4.8 [Moss(2010a)]) If* $\Gamma \vdash_{RAA} \phi$*, then* $\Gamma \vdash \phi$*.*

$\vdash_{RAA}$, $\vdash$ are derivability relations of the natural deduction systems that both contain the seven rules in [Moss(2010a), p. 3, Figure 1] plus the rule (RAA) and the rule $(Contrad)$, respectively, in unified notation.

[27] Otherwise, one would have a derivation whose conclusion is $\bot$, which is impossible by definition 1.

$$RAA^1\ \frac{\begin{array}{cc}[\phi]^1 & [\phi]^1\\ \vdots & \vdots\\ \psi & \neg\psi\end{array}}{\neg\phi} \qquad Contrad\ \frac{\forall(c,d) \quad \exists(c,\overline{d})}{\phi}$$

(RAA) is discussed on page 258 above (notice that the bar notation from definition 1 is in effect as well). $(Contrad)$—a syllogistic variant of $(\overline{E})$—derives an arbitrary sentence because its premisses are inconsistent.

Whilst its opposite is trivial, theorem 3 is surely not. To this end, Moss presents the normalization procedure whose aim is to obtain a derivation of the three variants of a normal form defined in [Moss(2010a), p. 13, Figure 3]. Below, in unified notation, we formulate the rules it employs in order to ease understanding and then make comments on the second variant of the normal form. The readers, nevertheless, are referred to [Moss(2010a)] for more details.

$$Zero\ \frac{\forall(c,\overline{c})}{\forall(c,d)} \qquad One\ \frac{\forall(\overline{c},c)}{\forall(d,c)} \qquad Barbara\ \frac{\forall(c,c_1) \qquad \forall(c_1,d)}{\forall(c,d)}$$

The second variant of the normal form is as follows:

$$\begin{array}{rc} & \begin{array}{c}\Sigma_1\\ \forall(\overline{c},c)\end{array} \qquad \begin{array}{c}\Sigma_2\\ \forall(c,\overline{c})\end{array}\\ Barbara & \dfrac{One\ \dfrac{}{\forall(d,c)} \qquad Zero\ \dfrac{}{\forall(c,c_1)}}{\forall(d,c_1)}\end{array}$$

With regard to (One) and $(Zero)$, both are introductions due to the fact that their application introduces d into a derivation. The premiss of $(Zero)$, as its name suggests, states that c is empty, and therefore it is included in any d.[28] Symmetrically, the premiss of (One), as its name suggests, states that c is universal, and therefore it includes any d. With regard to $(Barbara)$, its name obviously refers to the most famous syllogistic mood, which is a syllogistic variant of the (cut), or the transitivity rule. It is an elimination, due to the fact that its application eliminates c_1 from a derivation.

With regard to the second variant of the normal form (and the other ones that we skip here), such normal derivations do not satisfy corollary 2.1 or 2.2 above. In it, the major premisses of $(Barbara)$—as in the case of $(trans)$, both premisses are major—are not assumptions, as well as the application of $(Barbara)$ succeeding the applications of (One) and $(Zero)$. It

[28] Notice that semantically, set terms are understood as sets.

does not follow, however, that normal derivations in the system in question do not enjoy a variant of the subterm property—the syllogistic analog of the subformula property. This topic is not of Moss' concern there. Hence, it might belong to future studies.

The second glimpse is shorter than the previous one. Without explicitly mentioning a normal form or normalization, [KruckMoss(2021), pp. 4–5; example 1.2] presents the following derivation of $\forall(c,d), \forall(c_1,c) \vdash \forall(\forall(d,r), \forall(c_1,r))$.[29]

$$\dfrac{\dfrac{\forall(c,d)}{\forall(\forall(d,r),\forall(c,r))}\,ANTI \qquad \dfrac{\forall(c_1,c)}{\forall(\forall(c,r),\forall(c_1,r))}\,ANTI}{\forall(\forall(d,r),\forall(c_1,r))}\,Barbara$$

Notice that $\forall(c,r)$ occurs neither in the assumptions nor in the conclusion. From the Kruckman–Moss standpoint, it is a shortcoming because it does not allow one to show the PTime complexity of the entailment relation in question [KruckMoss(2021), theorem 1.5]. To get rid of the shortcoming, they set up another derivation.[30]

$$\dfrac{\dfrac{\forall(c_1,c) \qquad \forall(c,d)}{\forall(c_1,d)}\,Barbara}{\forall(\forall(d,r),\forall(c_1,r))}\,ANTI$$

From the proof-theoretic standpoint, the uppermost derivation is nothing but a detour, where both major premisses of the elimination are derived via the introduction. So, a term that does not occur in its assumptions or conclusion appears in the derivation. The Kruckman–Moss solution is standard—to convert the derivation by permuting the respective applications of these rules.

Notice that $\forall(c,d)$, $\forall(c_1,c)$ could not be assumptions, generally. Hence, in another derivation, the occurrences of c might be a detour, which is not the case in the uppermost derivation, i.e. there is an unwanted situation when getting rid of this detour gives rise to another. However, it does not cause a problem because $dg(c) < dg(\forall(c,r))$, by definition 5, which secures the inductive hypothesis of the normalization theorem.

Both glimpses above concern syllogistics without propositional connectives. Mineshima–Okada–Takemura's paper, which we briefly mention, proposes 'a simple inference system for the *syllogistic* fragment of natural lan-

[29] We refer to [KruckMoss(2021)] for the formulation of the introduction ($ANTI$), while unifying the notation.

[30] Notice that Kruckman–Moss do not elaborate on the way they find another derivation, and they do not consider the other shortcomings of this kind in the system.

guage, thereby making a connection between proof theory and natural language semantics' [Mineshima(2012), p. 753; the *italics* original]. Without using relational terms, they employ just one sentential connective—a generalization of the conventional conjunction, and it is the conjunction only that has eliminations and introductions. In other words, the syllogistic connective-free rules of their system are neither introductions nor eliminations. (Strictly, all those rules are eliminations, but we consider the distinction introductions/eliminations not to hold in this case.) Hence, in contrast to our approach, posing the normalization problem stays, so to speak, on the standard level of propositional connectives without going down to the 'syllogistic' one.[31] As a result, Mineshima–Okada–Takemura obtain the standard normalization theorem for their system.

5 Conclusion and Plans for Future

The paper gives a solution to the normalization problem for the sound and complete natural deduction systems $\mathbf{R}^{*\dagger}$, $\mathbf{R}^{*\dagger}_{\mathbf{trans}}$, and $\mathbf{RCA}^{\dagger}(\mathbf{opp})$ which formalize the relational syllogistics $R^{*\dagger}$, $R^{*\dagger}_{trans}$, and $RCA^{\dagger}(opp)$, respectively, that Moss introduces with his coauthors [Moss(2010), Moss(2010a), Moss(2015), KruckMoss(2021), Pratt(2009), to mention but a few]. Their sophisticated languages and syntax make $R^{*\dagger}$, $R^{*\dagger}_{trans}$, and $RCA^{\dagger}(opp)$ (ordered by expressive power) a proper candidate to be the most powerful of syllogistics without propositional connectives. As a result, it is of much importance to solve the normalization problem—to obtain detour-free derivations in the syllogistics considered—for them. The specifics of tackling the problem in question is that the variety of methods that have been successfully applied in the case of propositional and quantifier logics [Prawitz(1965), Prawitz(1971), Plato(2012), Kurbis(2022), PelletierHazen(2023), Parigot(1997), Geuvers(2024), Plato(2017)] fail the specifics of the syllogistics, and therefore could not be mechanically applied to them. Hence, instead of the subformula property—one of the main corollaries from the normalization theorem in the propositional case—one needs to speak of a particular case of the subterm property, due to the fact that the syllogistics considered deal with atomic sentences only and due to the fact that the proof theory for $RCA^{\dagger}(opp)$ needs extending the respective language with *general* sentences, which do not occur in valid, or provable entailments. This specificity implies that the normalization theorem 2 proven for each system in question leads to the so-called *sub-set-term-constant property* **only** (corollary 2.3).

[31] We do not consider it to be a shortcoming of their approach, by any means.

With regard to the future plans, the present approach could be developed extensively and intensively. The former opens perspectives of its applications to the syllogistics with [Mineshima(2012)] and without propositional connectives [Moss(2010a), KruckMoss(2021)] that chapter 4 reviews, as well as to the syllogistics related to generalized quantifiers [Westerstaahl(1989)]. The latter concerns tackling the proof search problem for the systems in this paper and would definitely need a stronger result on the subterm property than the sub-set-term-constant property. We conjecture that proof search procedures would affect the current presentation of the systems in question: dropping down the ad-hoc syntax of general sentences introduced—especially for $\mathbf{RCA}^{\dagger}(\mathbf{opp})$—would go hand-in-hand with introducing some machinery in order to avoid losing its deductive power, as is done for *FOL* in [Quine(1950)], for instance.

Acknowledgments Special thanks go to Dag Westerståhl for drawing the author's attention to the results by Larry S. Moss and his coauthors. Angelina R. Baidina is thanked by the author for the *good vibrations*. This work was supported by Russian Science Foundation, grant number 24-28-00672 (https://rscf.ru/project/24-28-00672/).

Declaration on the use of AI The author used Quillbot at https://quillbot.com/ to assist with language editing. All generated text was reviewed, edited, and verified by the author, who takes full responsibility for the content.

References

[Gentzen(1969)] G. Gentzen. *The collected papers of Gerhard Gentzen*, M.E. Szabo (ed.), North-Holland Pub. Co., Amsterdam, 1969. https://doi.org/10.2307/2272429

[Geuvers(2024)] H. Geuvers and T. Hurkens. Normalization for classical natural deduction. In *30th International conference on types for proofs and programs, TYPES 2024—Abstracts, June 10-14, 2024, Copenhagen, Denmark*, 114–117, 2024. https://doi.org/10.4230/LIPIcs.TYPES.2022.2

[KruckMoss(2021)] A. Kruckman and L.S. Moss. Exploring the landscape of relational syllogistic logics. *Review of Symbolic Logic*, 14(3), 728–765, 2021. https://doi.org/10.1017/S1755020320000386

[Kurbis(2022)] N. Kürbis. Normalisation and subformula property for a system of classical logic with Tarski's rule. *Archive for Mathematical Logic*, 61, 105–129, 2022. https://doi.org/10.1007/s00153-021-00775-6 Correction: N. Kürbis. Normalisation and subformula property for a system of classical logic with Tarski's rule, and a correction, 2023. https://doi.org/10.48550/arXiv.2108.03939

[Milne(2015)] P. Milne. Inversion principles and introduction rules. In H. Wansing (ed.). *Dag Prawitz on proofs and meaning*, 1st ed., Springer, Cham, 189–224, 2015. https://doi.org/10.1007/978-3-319-11041-7_8

[Mineshima(2012)] K. Mineshima, M. Okada, and R. Takemura. A generalized syllogistic inference system based on inclusion and exclusion relations. *Studia Logica*, 100(4), 753–785, 2012. https://doi.org/10.1007/s11225-012-9428-5

[Moss(2010)] L.S. Moss. Syllogistic logics with verbs. *Journal of Logic and Computation*, 20(4), 947–967, 2010. https://doi.org/10.1093/logcom/exn086

[Moss(2010a)] L.S. Moss. Syllogistic logic with complements, in J. van Benthem, A. Gupta, and E. Pacuit (eds.), Games, norms and reasons: logic at the crossroads, Springer Synthese Library Series, 2010, 185–203. https://doi.org/10.1007/978-94-007-0714-$6_1$1

[Moss(2015)] L.S. Moss. Natural Logic. In *Handbook of contemporary semantic theory*, second edition, chapter 18, John Wiley & Sons, 2015. https://doi.org/10.1002/9781118882139

[NegriPlato(2001)] S. Negri and J. von Plato. *Structural proof theory*, Cambridge University Press, 2001. https://doi.org/10.1017/CBO9780511527340

[Parigot(1997)] M. Parigot. Proofs of strong normalization for second order classical natural deduction. *Journal of Symbolic Logic*, 62, 1461–1479, 1997. https://doi.org/10.2307/2275652

[PelletierHazen(2023)] F.J. Pelletier and A. Hazen. Natural deduction systems in logic. In E.N. Zalta and U. Nodelman (eds.). *The Stanford Encyclopedia of Philosophy (spring 2023 edition)*, 2023. https://plato.stanford.edu/archives/spr2023/entries/natural-deduction/.

[Plato(2017)] J. von Plato. *Sources and studies in the history of mathematics and physical sciences. Saved from the cellar: Gerhard Gentzen's shorthand notes on logic and foundations of mathematics*, Cham, Switzerland, Springer, 2017. https://doi.org/10.1007/978-3-319-42120-9

[Plato(2012)] J. von Plato and A. Siders. Normal derivability in classical natural deduction. *Review of Symbolic Logic*, 5(2), 205–211, 2012. https://doi.org/10.1017/S1755020311000311

[Pratt(2009)] I. Pratt-Hartmann and L.S. Moss. Logics for the relational syllogistic. *Review of Symbolic Logic*, 2(4), 647–683, 2009. https://doi.org/10.1017/S1755020309990086

[Prawitz(1965)] D. Prawitz. *Natural deduction: a proof-theoretical study*, Stockholm, Sweden, Dover Publications, 1965. https://doi.org/10.2307/2271676

[Prawitz(1971)] D. Prawitz. *Ideas and results in proof theory.* In J.E. Fenstad (ed.). *Studies in logic and the foundations of mathematics*, 63, Elsevier, 235–307, 1971. https://doi.org/10.1016/S0049-237X(08)70849-8

[Quine(1950)] Quine, W.V. (1950). On natural deduction. *Journal of Symbolic Logic*, 15(2), 93–102. https://doi.org/10.2307/2266969 Correction: (1950). JSL volume 15 issue 2 Cover and front matter and errata. *Journal of Symbolic Logic*, 15(2), f1–f6. https://doi.org/10.1017/s0022481200103913

[Segerberg(1983)] K. Segerberg. Arbitrary truth-value functions and natural deduction. *Mathematical Logic Quarterly*, 29(11), 557–564, 1983. https://doi.org/10.1002/malq.19830291102

[Shangin(2023)] V. Shangin. A classical first-order normalization procedure with $\forall$ and $\exists$ based on the Milne–Kürbis approach. *Synthese*, 202(48), 1–24, 2023. https://doi.org/10.1007/s11229-023-04276-2

[Westerstaahl(1989)] D. Westerståhl. Aristotelian syllogisms and generalized quantifiers. *Studia Logica*, 48(4), 577–585, 1989. https://doi.org/10.1007/BF00370209