Proof theory for heterogeneous logic combining formulas and diagrams —Proof normalization—

Ryo Takemura

Nihon University, Japan.

takemura.ryo@nihon-u.ac.jp

April 29, 2020

Abstract

We extend natural deduction for first-order logic (FOL) by introducing diagrams as components of formal proofs. From the viewpoint of FOL, we regard a diagram as a deductively closed conjunction of certain FOL formulas. On the basis of this observation, we first investigate basic heterogeneous logic (HL) wherein heterogeneous inference rules are defined in the styles of conjunction introduction and elimination rules of FOL. By examining what is a detour in our heterogeneous proofs, we discuss that an elimination-introduction pair of rules constitutes a redex in our HL, which is opposite the usual redex in FOL. In terms of the notion of a redex, we prove the normalization theorem for HL, and we give a characterization of the structure of heterogeneous proofs. Every normal proof in our HL consists of applications of introduction rules followed by applications of elimination rules, which is also opposite the usual form of normal proofs in FOL. Thereafter, we extend the basic HL by extending the heterogeneous rule in the style of general elimination rules to include a wider range of heterogeneous systems.

Contents

1 Introduction

Proof theory has traditionally been developed based on sentential representation of logical proofs. Formal proofs are defined as chains of sentences/formulas, and other forms of representations, such as diagrams or graphs, are regarded not as components of formal proofs, but only as auxiliary tools to construct formal proofs or to explain ideas.

However, recently, after the pioneering works of Barwise, Etchemendy, and Shin (e.g., [1, 5, 23]), diagrams have been investigated as counterparts of logical formulas, which constitute formal proofs. Diagrams are rigorously defined as syntactic objects, allowing set-theoretical semantics to be defined. Inference systems are formalized, and they are shown to be equivalent to some symbolic logical systems. Consequently, fundamental logical properties, such as soundness and completeness, have been investigated (e.g., [23, 10, 13, 26]). These diagrammatic systems are further developed into heterogeneous systems combining diagrams and formulas. These include Hyperproof [2, 5, 3], which is developed by Tarski's World; Euler and Venn systems [9, 25, 31]; and correspondence table systems [4, 28]. However, a proof theory for heterogeneous logic generally has not yet been well-developed. Based on these previous studies, we give a proof-theoretical investigation of heterogeneous logic by extending natural deduction for first-order logic (FOL).

Diagrammatic reasoning consists of construction of a diagram by unifying pieces of information contained in given premises, and extraction of a conclusion from the unified diagram. This applies to diagrammatic reasoning generally, including a proof of the Pythagorean Theorem, diagram chasing in category theory, syllogistic reasoning with Euler diagrams, and so on. Refer to Section 2.1 for an example of Euler diagrammatic reasoning. Thus, on top of the usual natural deduction for FOL, we introduce inference rules for diagrams: Construction rules (unification and application) of a diagram and extraction rules (deletion and observation) from a diagram. From the viewpoint of FOL, our diagram can be regarded as a deductively closed set of certain FOL formulas, which describes basic information contained in the diagram. We elucidate it as Postulate 2.1 in Section 2, and we investigate such diagrams in this article. Then, the above construction and extraction rules of diagrams can be considered as certain introduction and elimination rules of natural deduction for FOL, respectively.

In Section 3, we show the normalization theorem for our heterogeneous logic HL. We first define reduction rules for heterogeneous proofs by examining what is a detour, i.e., a redex in our heterogeneous proof. We discuss that, in light of the Postulate 2.1, an extraction-construction pair, i.e., an elimination-introduction pair of rules constitutes a redex in heterogeneous proofs. This is opposite the usual redex in the natural deduction for FOL, where an introduction-elimination pair constitutes a redex. Based on this notion of redex, we prove the normalization theorem for HL, and we give a characterization of the structure of heterogeneous proofs. Every normal proof in our HL consists of applications of construction rules followed by applications of extraction rules. This is also opposite the usual form of normal proofs in FOL. We further investigate a diagrammatic counterpart of the usual subformula property of FOL, that is the maximal diagram property of HL.

In Section 4, we extend our basic HL to HLe by introducing the so-called rules of exhaustive cases of [5, 2] in the style of $\vee E$-rule in the natural deduction for FOL. Although inference rules in the style of $\vee E$-rule make it difficult to grasp the structure of proofs, *general elimination rules*, including $\vee E$-rule, have been introduced by Dyckhoff [7], López-Escobar [12], and Tennant [29], and have also been independently introduced and extensively studied by von Plato and Negri [17, 16], and Schroeder-Heister [21]. In particular, translations between natural deduction with general elimination rules and sequent calculus were investigated, the normalization theorem was proved, and the structure of proofs with general elimination rules was investigated. By applying these results, we prove the normalization theorem for our extended HLe, and investigate the structure of heterogeneous proofs in HLe.

Finally, in Section 5, we summarize our results and discuss various dualities between sentential reasoning and diagrammatic reasoning.

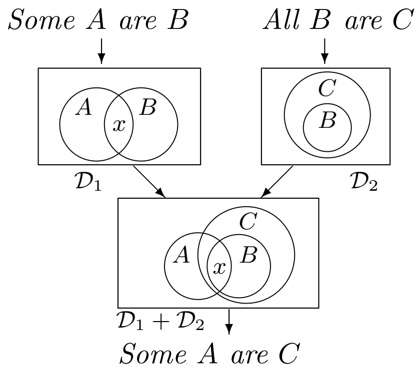
This article is a development of the research introduced in [27], and this article includes, for completeness, a partial summary of this work.

2 Syntax of heterogeneous logic

In Section 2.1, to give an insight into reasoning with diagrams, we illustrate reasoning with Euler diagrams of [13]. In Section 2.2, we review the natural deduction system for FOL, based on which we introduce our abstract heterogeneous logic HL in Section 2.3. In Section 2.4, we introduce inference rules of HL; subsequently, we also introduce term notation for our heterogeneous proofs by extending the usual λ-terms.

2.1 Euler diagrammatic reasoning

Let us consider an Euler diagrammatic representation of the following syllogistic reasoning: *Some A are B, All B are C* $\models$ *Some A are C*.



The given premise *All B are C* is represented as $\mathcal{D}_2$ in terms of the inclusion relation between circles B and C. The other premise *Some A are B* is represented as $\mathcal{D}_1$. Note that the relationship between A and B is not specified to inclusion or exclusion by the given sentence. In such a case, we represent it as in $\mathcal{D}_1$ following Venn diagrammatic conventions, and we call the relation a "crossing relation" between A and B. Furthermore, we do not assume the so-called existential import for regions, that is, we do not assume that every region in a diagram is nonempty. Thus, two crossing circles A and B do not imply anything in their own right, and to represent the nonemptiness, we need point x to represent the existence of some object explicitly. (The existential import for regions destroys the one-to-one correspondence between syllogistic sentences and diagrams. Please refer to the introduction of [13] for a brief description of the historical background of Euler and Venn diagrams.)

Then, the validity of the given inference is shown by unifying $\mathcal{D}_1$ and $\mathcal{D}_2$. In this unification, circle B in $\mathcal{D}_1$ and $\mathcal{D}_2$ is identified, and C is added to $\mathcal{D}_1$ so that B is inside C and C overlaps with A without any implication of a specific relationship (i.e., inclusion or exclusion) between A and C. From $\mathcal{D}_1 + \mathcal{D}_2$, we can read off the conclusion that *Some A are C*. Furthermore, if we have another premise *No D are C*, just by adding circle D outside C in $\mathcal{D}_1 + \mathcal{D}_2$, we can read off *No D are B*, *Some A are not D*, *Some B are not D*, and *Some C are not D* without deducing them explicitly as in sentential reasoning. This phenomenon where, by constructing a diagram, we obtain a conclusion without explicit deduction is called *free ride* in the literature of diagrammatic reasoning ([24]).

An Euler diagram of [13], called an EUL-diagram, is defined as a plane with named circles $(A, B, C, \dots)$ and points $(x, y, a, b, \dots$ collectively denoted by $p, q)$, which are called **diagrammatic objects**. EUL-diagrams can express neither disjunctive information with respect to the location of a point, nor information on contradiction. Each EUL-diagram is specified in terms of the following relations between diagrammatic objects; those are described by the corresponding FOL formulas called **diagrammatic formulas**.

$A \sqsubset B$	the interior of A is *inside of* the interior of B	$\forall x(Ax \to Bx)$
$A \vdash\!\dashv B$	the interior of A is *outside of* the interior of B	$\forall x(Ax \to \neg Bx)$
$A \bowtie B$	there is at least one *crossing* point between A and B	$\forall x(Ax \to Ax) \wedge \forall x(Bx \to Bx)$
$p \sqsubset A$	p is *inside of* the interior of A	Ap
$p \vdash\!\dashv A$	p is *outside of* the interior of A	$\neg Ap$

Note that we interpret the crossing relation so that it does not convey any specific (inclusion or exclusion) information about the relationship between circles. Thus, $A \bowtie B$ is translated into a tautology as above. Refer to [10] for the other region-based formalization of Euler diagrams in contrast to our relation-based formalization of [13]. Also refer to [14] for a comparison thereof.

For example, the above diagram $\mathcal{D}_1 + \mathcal{D}_2$ consists of relations $B \sqsubset C, A \bowtie B, A \bowtie C, x \sqsubset A, x \sqsubset B, x \sqsubset C$. We call the set of relations the **type** of the diagram and denote it by $|\mathcal{D}_1 + \mathcal{D}_2|$. The type of a diagram is the symbolic specification of that diagram. In terms of formulas, $|\mathcal{D}_1 + \mathcal{D}_2|$ is the conjunction of diagrammatic formulas $\forall x(Bx \to Cx), \forall x(Ax \to Ax) \wedge \forall x(Bx \to Bx), \forall x(Ax \to Ax) \wedge \forall x(Cx \to Cx), Ax, Bx,$ and Cx. We identify the conjunctive formula with the type of the diagram.

In this way, our diagram is specified by the set of relations that holding on the diagram. Note that $|\mathcal{D}_1| = \{A \bowtie B, x \sqsubset A, x \sqsubset B\}$, $|\mathcal{D}_2| = \{B \sqsubset C\}$, and $|\mathcal{D}_1 + \mathcal{D}_2| = |\mathcal{D}_1| \cup |\mathcal{D}_2| \cup \{x \sqsubset C\}$, where $x \sqsubset C$ is obtained from $x \sqsubset B$ and $B \sqsubset C$. Thus, $|\mathcal{D}_1 + \mathcal{D}_2|$ is not just the union of $|\mathcal{D}_1|$ and $|\mathcal{D}_2|$, but the deductive closure of $|\mathcal{D}_1| \cup |\mathcal{D}_2|$ (except for irrelevant tautologies). The notion of free ride corresponds to the deductive closedness of diagrams from the viewpoint of FOL.

Based on the above insight on Euler diagrams, we elucidate our intended diagrams, which are not restricted to Euler diagrams, in this study.

Postulate 2.1 (Diagrams)

1. We assume the set of diagrammatic formulas to be a subset of FOL formulas. Thus, every piece of basic information contained in a diagram is described by a formula of FOL. Considering FOL, we regard a diagram as the conjunction of diagrammatic formulas comprising the diagram.

2. We assume the type $|\mathcal{D}|$ of a diagram $\mathcal{D}$ is *deductively closed* with respect to diagrammatic formulas. That is, for any diagrammatic formula φ^d, if φ^d is provable from $|\mathcal{D}|$ in FOL then $\varphi^d \in |\mathcal{D}|$.

Although we mainly illustrate only one concrete Euler diagrammatic system [13] below, our investigation is independent of how concrete diagrams are defined in each concrete system, and our definition is valid for other Euler and Venn systems (e.g., [9, 23, 10]), blocks world systems ([2, 5]), and so on.

2.2 Natural deduction rules for FOL

Formulas of FOL, denoted by $\varphi, \psi, \sigma, \varphi_1, \varphi_2, \ldots$, are defined inductively as usual:

$$\varphi ::= A(e_1, \ldots, e_n) \mid \varphi \land \varphi \mid \varphi \lor \varphi \mid \varphi \to \varphi \mid \neg\varphi \mid \bot \mid \forall x\varphi \mid \exists x\varphi$$

where $A(e_1, \ldots, e_n)$ is an atomic formula comprising a predicate A and first-order terms $e_1, \ldots, e_n$. When A is a unary predicate, we usually omit parentheses and write an atomic formula such as Ae.

A proof in natural deduction is structured as a tree consisting of formulas as its nodes and inference rules as its edges. The natural deduction rules for FOL consist of dual pairs, with each pair comprising introduction (I) and elimination (E) rules, for each connective $\land, \lor, \to, \neg, \forall, \exists$, as well as $\bot E$ and RAA. In the following rules, a formula written within square brackets indicates that the assumptions of this form are closed at the inference. We call such closed assumptions **auxiliary assumptions** of the rule. In every elimination rule, the premise containing the connective being eliminated is called the **major** premise and the other premises are called **minor** premises of the rule. See [6, 19] for a detailed introduction to natural deduction.

$$\dfrac{\begin{matrix}\vdots & \vdots\\ \varphi & \psi\end{matrix}}{\varphi\land\psi}\land I \qquad \dfrac{\begin{matrix}\vdots\\ \varphi_1\land\varphi_2\end{matrix}}{\varphi_i}\land E_{(i\,=\,1,\,2)} \qquad \dfrac{\begin{matrix}\vdots\\ \varphi_i\end{matrix}}{\varphi_1\lor\varphi_2}\lor I_{(i\,=\,1,\,2)} \qquad \dfrac{\begin{matrix} & [\varphi]^n & [\psi]^n\\ \vdots & \vdots & \vdots\\ \varphi\lor\psi & \sigma & \sigma\end{matrix}}{\sigma}\lor E, n$$

$$\dfrac{\begin{matrix}[\varphi]^n\\ \vdots\\ \psi\end{matrix}}{\varphi\to\psi}\to I, n \qquad \dfrac{\begin{matrix}\vdots & \vdots\\ \varphi & \varphi\to\psi\end{matrix}}{\psi}\to E \qquad \dfrac{\begin{matrix}[\varphi]^n\\ \vdots\\ \bot\end{matrix}}{\neg\varphi}\neg I, n \qquad \dfrac{\begin{matrix}\vdots & \vdots\\ \varphi & \neg\varphi\end{matrix}}{\bot}\neg E$$

$$\dfrac{\begin{matrix}\vdots\\ \varphi\end{matrix}}{\forall x\varphi}\forall I \qquad \dfrac{\begin{matrix}\vdots\\ \forall x\varphi\end{matrix}}{\varphi[x := e]}\forall E \qquad \dfrac{\begin{matrix}\vdots\\ \varphi[x := e]\end{matrix}}{\exists x\varphi}\exists I \qquad \dfrac{\begin{matrix} & [\varphi]^n\\ \vdots & \vdots\\ \exists x\varphi & \psi\end{matrix}}{\psi}\exists E, n \qquad \dfrac{\begin{matrix}\vdots\\ \bot\end{matrix}}{\varphi}\bot E \qquad \dfrac{\begin{matrix}[\neg\varphi]^n\\ \vdots\\ \bot\end{matrix}}{\varphi}RAA, n$$

In $\forall I$, the variable x may not freely occur in any open assumption on which φ depends; in $\exists E$, x may not freely occur in ψ nor in any open assumption on which ψ depends, except in φ.

The above set of rules provides the rules of classical logic, and the set without RAA (Reductio Ad Absurdum rule) provides that of intuitionistic logic.

We write $\varphi_1, \ldots, \varphi_n \vdash \psi$ when there exists a natural deduction proof of ψ (i.e., is provable) from open assumptions $\varphi_1, \ldots, \varphi_n$. See [6] for the usual definition.

2.3 Diagrams

On top of the system of natural deduction, we introduce diagrams in this section and inference rules thereof in the next section. While concrete syntax is defined in each system, we extract common items to be specified in each system. The syntax of heterogeneous logic is defined by specifying *diagrammatic objects, diagrams*, and *diagrammatic formulas* in addition to formulas.

Diagrammatic objects are components of diagrams, and are denoted by $o, o_1, o_2, \dots$. We denote the set of diagrammatic objects that appear on a diagram $\mathcal{D}$ by $ob(\mathcal{D})$.

Diagrams are denoted by $\mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{D}_1, \mathcal{D}_2, \dots$.

Diagrammatic formulas are denoted by $\varphi^d, \psi^d, \sigma^d, \varphi_1^d, \varphi_2^d \dots$. They describe pieces of basic information contained in diagrams. When it is clear from the context, we denote a diagrammatic formula simply by φ without the superscript d.

We identify a diagrammatic object in a diagram with a predicate or a term describing the object in a diagrammatic formula. Thus, for a diagrammatic formula φ^d, we use $ob(\varphi^d)$ to denote the set of predicates and terms appearing in φ^d, each of which expresses a diagrammatic object. Furthermore, we identify a relation holding on a diagram with a diagrammatic formula describing the relation.

For a diagram $\mathcal{D}$, its *type* $|\mathcal{D}|$ is the set $\{\varphi_1^d, \varphi_2^d, \dots, \varphi_n^d\}$ of diagrammatic formulas such that the relation φ_i^d holds on $\mathcal{D}$ if and only if $\varphi_i^d \in |\mathcal{D}|$ for all $1 \leq i \leq n$. We identify $|\mathcal{D}|$ with the conjunctive formula $\varphi_1^d \wedge \varphi_2^d \wedge \cdots \wedge \varphi_n^d$.

What qualifies as a concrete diagram is defined in each concrete system. Here, we define a general operation $+$ on diagrams.

Definition 2.2 (Diagrams)

- When $\mathcal{D}$ and $\mathcal{E}$ are both diagrams, $\mathcal{D}+\mathcal{E}$ is a diagram such that $|\mathcal{D}+\mathcal{E}| = \{\psi^d \mid |\mathcal{D}|, |\mathcal{E}| \vdash \psi^d \text{ in FOL}\}$.
- When $\mathcal{D}$ is a diagram and φ^d is a diagrammatic formula, $\mathcal{D} + \varphi^d$ is a diagram such that $|\mathcal{D} + \varphi^d| = \{\psi^d \mid |\mathcal{D}|, \varphi^d \vdash \psi^d \text{ in FOL}\}$.
- The operation $+$ is associative and commutative.
- A diagram $\mathcal{E}$ is a subdiagram of $\mathcal{D}$, written as $\mathcal{E} \sqsubseteq \mathcal{D}$, when $|\mathcal{E}| \subseteq |\mathcal{D}|$ holds.

We denote a formula or a diagram by $\alpha, \beta, \gamma, \delta, \alpha_1, \alpha_2, \dots$.

$\mathcal{D} + \mathcal{E}$ is the unified diagram of $\mathcal{D}$ and $\mathcal{E}$ (cf. [10, 13]). $\mathcal{D} + \varphi^d$ is the diagram that extends from $\mathcal{D}$ by adding the information of φ^d called **app-formula** (cf. [2, 5]).

Depending on the specific definition of diagrams in each system, $\mathcal{D} + \varphi^d$ and $\mathcal{D} + \mathcal{E}$ are not always defined. Typical cases are the case of *indeterminacy*, where $\mathcal{D} + \varphi^d$ or $\mathcal{D} + \mathcal{E}$ is not uniquely determined as a single diagram because of an appearance of disjunctive information, and the case of *inconsistency*, where $\mathcal{D}$ and φ^d, or $\mathcal{D}$ and $\mathcal{E}$, are inconsistent. To avoid these cases, in the Euler diagrammatic system [13], constraints are imposed on the operation $+$, and, in [23, 10], the notion of a diagram is extended to express indeterminacy and inconsistency by introducing certain devices. In our basic HL, to make $+$ always applicable, we introduce an expression $*$, which expresses that $\mathcal{D} + \varphi^d$ or $\mathcal{D} + \mathcal{E}$ is undefined. For technical reasons, we regard $*$ as a diagram, which does not appear in given premises and conclusion. $*$ is not necessarily equivalent to $\bot$ in FOL, since, other than contradiction, there are cases wherein diagrams are not defined because of their expressive limitations. In Section 4, we extend our basic HL to HLe by introducing the rule of Cases Exhaustive [2, 5] to deal with certain disjunctive information in diagrams.

Although we do not enter into detail here, the semantics of our heterogeneous system HL is defined as the usual set-theoretic semantics for FOL because our diagram corresponds to a conjunction of appropriate formulas.

2.4 Inference rules of basic heterogeneous logic HL

Among various inference rules, as representative rules of heterogeneous systems independent of specific diagrams, we investigate the following rules, where app and obs consist of the dual pair of heterogeneous rules, and uni and del consist of the dual diagrammatic rules.

Definition 2.3 (Basic HL) Heterogeneous rules of Application (app) and Observation (obs), and diagrammatic rules of Unification (uni) and Deletion (del) have the following forms:

$$\frac{\begin{matrix}\vdots & \vdots\\ \mathcal{D} & \varphi^d\end{matrix}}{\mathcal{D}+\varphi^d}\ \mathsf{app} \qquad \frac{\begin{matrix}\vdots\\ \mathcal{D}\end{matrix}}{\psi^d}\ \mathsf{obs}\ \text{where } \psi^d \in |\mathcal{D}| \qquad \frac{\begin{matrix}\vdots & \vdots\\ \mathcal{D} & \mathcal{E}\end{matrix}}{\mathcal{D}+\mathcal{E}}\ \mathsf{uni} \qquad \frac{\begin{matrix}\vdots\\ \mathcal{D}\end{matrix}}{\mathcal{E}}\ \mathsf{del}\ \text{where } \mathcal{E} \sqsubseteq \mathcal{D}$$

In app, as well as obs, an app-formula (resp. observed formula) is restricted to being a diagrammatic formula. This is the result of our postulate regarding a diagram as a conjunction of diagrammatic formulas.

For $*$, which expresses that $\mathcal{D}+\varphi^d$ or $\mathcal{D}+\mathcal{E}$ is undefined, we introduce the following $*E$-rule similar to the usual $\bot E$-rule of FOL. When $\mathcal{D}+\varphi^d$ is not defined, we express it as follows:

$$\frac{\mathcal{D} \quad \varphi^d}{*}\ \mathsf{app} \qquad \frac{* \quad \varphi^d}{*}\ \mathsf{app} \qquad\qquad \frac{*}{\mathcal{D}}\ *E$$

Similarly for uni.

Since $*$ is obtained by an application of app or uni, we may consider $*$ as the most complex diagram. Then, $*E$ can be regarded as a special case of del.

Both app and uni are operations to unify given pieces of information into a diagram, and both obs and del (as well as $*E$) are operations to extract pieces of information from a diagram. These essentially have the same structure, and they play the same roles in a proof and in the reduction of proofs (cf. Section 3.2). Thus, we denote app and uni collectively by $+$, and obs and del (and $*E$) by $-$ as follows, where α is a diagram or a formula:

$$\frac{\begin{matrix}\vdots & \vdots\\ \mathcal{D} & \alpha\end{matrix}}{\mathcal{D}+\alpha}\ (+) \qquad\qquad \frac{\begin{matrix}\vdots\\ \mathcal{D}\end{matrix}}{\alpha}\ (-)$$

Thus, when α is a diagrammatic formula, $+$ is app and $-$ is obs. When α is a diagram, $+$ is uni and $-$ is del (or $*E$ when $\mathcal{D}$ is $*$).

A **heterogeneous proof**, denoted by $t, s, u, v, t_1, t_2, \ldots$, is inductively defined as a tree consisting of formulas and diagrams as its nodes, and inference rules as its edges. We write

$\alpha_1, \ldots, \alpha_n \vdash \alpha$, when α is provable from premises $\alpha_1, \ldots, \alpha_n$, where α, α_i is a formula or a diagram.

By extending the usual λ-terms, we introduce proof-terms for our basic HL, which describe our reduction rules compactly and precisely. Based on the usual λ-terms, we define app, uni, obs, and del-operations. We recall proof-terms for $\rightarrow$ and $\wedge$ of FOL. See, for example, [30] for terms of FOL.

Definition 2.4 **Proof-terms**, or simply **proofs**, for basic HL is inductively defined as follows.

1. A term-variable x^α is a proof of α.
2. If s is a proof of φ and t is a proof of ψ, then $\langle s, t \rangle$ is a proof of $\varphi \wedge \psi$.
3. If t is a proof of $\varphi \wedge \psi$ then $p_1(t)$ is a proof of φ and $p_2(t)$ is a proof of ψ.
4. If t is a proof of ψ then $\lambda x^\varphi.t$ is a proof of $\varphi \rightarrow \psi$.
5. If s is a proof of $\varphi \rightarrow \psi$ and t is a proof of φ, then st is a proof of ψ.
6. If s is a proof of $\mathcal{D}$ and t is a proof of φ^d, then $\mathsf{app}(s, t)$ is a proof of $\mathcal{D} + \varphi^d$.
7. If s is a proof of $\mathcal{D}$ and t is a proof of $\mathcal{E}$, then $\mathsf{uni}(s, t)$ is a proof of $\mathcal{D} + \mathcal{E}$.
8. If t is a proof of $\mathcal{D}$ and $\varphi^d \in |\mathcal{D}|$, then $\mathsf{obs}(t)$ is a proof of φ^d.
9. If t is a proof of $\mathcal{D}$ and $\mathcal{E} \sqsubseteq \mathcal{D}$, then $\mathsf{del}(t)$ is a proof of $\mathcal{E}$.
10. If t is a proof of $*$, then $*(t)$ is a proof of $\mathcal{D}$ for any diagram $\mathcal{D}$.

- We denote a proof of $\mathsf{app}(s, t)$ or $\mathsf{uni}(s, t)$ by $+(s, t)$.
- We denote a proof of $\mathsf{obs}(t), \mathsf{del}(t)$, or $*(t)$ by $-(t)$.

Although we use the same x to express a term-variable and a first-order variable, we usually denote a term-variable as x^φ using a formula φ. Although we sometimes omit the superscript to avoid notational complexity, it is possible to determine which variable we refer to by the context.

Because of our $-$-rule, the formula (type) of a given proof-term is not uniquely determined by the form of the term. To avoid this situation, we may describe our proof-term with its formula or diagram α as t^α, but we omit the superscript to avoid notational complexity.

3 Normalization of heterogeneous proofs

We recall the notions of *detour*, *reduction*, and *normal proof* in the usual natural deduction for FOL in Section 3.1. Then, we discuss their counterparts for heterogeneous logic in Section 3.2. In Section 3.3, we prove our normalization theorem for basic HL. Based on the theorem, we investigate a characterization of the structure of heterogeneous proofs in Section 3.4.

3.1 Normal proof and normalization in FOL

In general, a natural deduction proof may contain some redundant steps and formulas called maximal formulas, i.e., formulas that simultaneously stands as the conclusion of an introduction rule and as the major premise of an elimination rule. For example, the formula $\varphi_1 \wedge \varphi_2$

and the pair of applications of $\wedge I$ and $\wedge E$ rules on the left in the following proof are redundant. This is because without them, we already have a proof t_1 of φ_1 as illustrated on the right.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\, t_1\\ \varphi_1\end{array} \quad \begin{array}{c}\vdots\, t_2\\ \varphi_2\end{array}}{\varphi_1 \wedge \varphi_2}\wedge I}{\varphi_1}\wedge E \qquad\qquad\qquad \begin{array}{c}\vdots\, t_1\\ \varphi_1\end{array}$$

A maximal formula along with its related pair of applications of an introduction and an elimination rule are called **detour** in a proof, and it is possible to remove such a detour as illustrated above. This rule of rewriting a given proof by removing a detour is called the **reduction rule**. It is defined for every pair of dual introduction and elimination rules. We call such a pair of rules a **redex**. In addition to the above $\wedge$-reduction rule, the reduction rules for $\to, \vee, \forall, \exists$ are defined. We recall only $\wedge$ and $\to$ reduction rules in term notation.

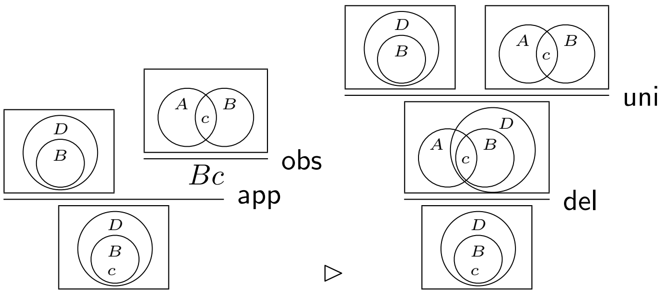
$\wedge$-reduction $\quad p_i\langle t_1, t_2\rangle \quad \rhd \quad t_i$ for $i = 1, 2$
$\to$-reduction $\quad (\lambda x^{\varphi}.s)t \quad \rhd \quad s[x^{\varphi} := t]$

A natural deduction proof is said to be in **normal form** when it does not contain any redex. Then, the following normalization theorem holds for the natural deduction system: *If φ is provable from $\varphi_1, \ldots, \varphi_n$, then there is a normal proof of φ from $\varphi_1, \ldots, \varphi_n$.* Normal proofs have an essential property called the **subformula property**: *A normal proof t of φ from $\varphi_1, \ldots, \varphi_n$ contains only subformulas of φ and $\varphi_1, \ldots, \varphi_n$.* (See [6, 30] for the precise formulation of these theorems.) Normalization theorem makes various proof-theoretical analyses possible. For example, the notion of normal proofs enables us to characterize the structure of proofs in a formal system. Prawitz [18, 19] revealed that each normal proof consists of two parts: An analytical part in which premises are decomposed into their components by using elimination rules; and a synthetic part in which the final components obtained in the analytical part are put together to construct the conclusion by using introduction rules.

3.2 Reduction rules for heterogeneous proofs

Let us define reduction rules for our heterogeneous proofs. Our dual pair of rules are app and obs, as well as uni and del. It is possible to define reduction rules thereof by applying the reduction rules of FOL straightforwardly through a translation of diagrams into formulas of FOL. However, as discussed in [26], the resulting "normal" proof does not reflect the characteristics of diagrammatic inference, and it may be called the "sentential normal" proof. This article investigates reduction rules inherent in our diagrammatic inference.

The notion of reduction in natural deduction for FOL is explained as the removal of a *detour*, i.e., a redundant maximal concept (formula) as well as its introduction and elimination rules. Let us consider what is a detour inherent in our diagrammatic inference. Based on our Postulate 2.1, in which we regard a diagram as the deductively closed conjunction of diagrammatic formulas, the use of redundant subconcepts (diagrams/formulas) may be considered to be a detour. This detour is part of a proof where one infers by deducing subconcepts, even though one can directly infer by using a superior concept. For example, Bc and the pair of obs-app in the following proof on the left are redundant, i.e., a detour, since the information of Bc is already contained in the diagram above the obs, and we can obtain the same conclusion by directly unifying two premise diagrams without deducing the formula Bc.



$$\dfrac{\overset{\vdots\; u}{\mathcal{E}} \qquad \overset{\vdots\; s}{\dfrac{\overset{\vdots\; t}{\mathcal{D}}}{\varphi}\,\mathsf{obs}}\;\psi}{\mathcal{E}+\psi}\,\mathsf{app} \quad \rhd \quad \dfrac{\dfrac{\overset{\vdots\; u}{\mathcal{E}} \qquad \overset{\vdots\; t}{\mathcal{D}}}{\mathcal{E}+\mathcal{D}}\,\mathsf{uni}}{\mathcal{E}+\psi}\,\mathsf{del}$$

$$\mathsf{app}(u, s[x^{\varphi} := \mathsf{obs}(t)]) \;\rhd\; \mathsf{del}(\mathsf{uni}(u,t))$$

This kind of a detour and its reduction can be formulated by the above rule on the right. When ψ depends only on φ (i.e., $\varphi \vdash \psi$), φ and ψ as well as the formulas/diagrams between them are redundant subconcepts. This is because our diagrams are deductively closed, and we are able to infer $\mathcal{E}+\psi$ by directly unifying $\mathcal{E}$ and $\mathcal{D}$ without deducing φ nor ψ as illustrated on the right. Note that this reduction is possible when ψ depends only on φ ($\varphi \vdash \psi$). That is, s is expressed as $s[x^{\varphi} := \mathsf{obs}(t)]$, where x^{φ} is fresh variable for s, and no variable other than x^{φ} appears in $s[x^{\varphi}]$.

Remark 3.1 Note that the notion of "detour" is conceptual, and it is not necessarily related to the length of proofs. It is well known that normal proofs may be more lengthy and complex than non-normal proofs, but they are conceptually simpler in the sense that no detour is contained.

The above obs-app-reduction may be generalized as follows: When $\varphi_1, \dots, \varphi_n, \sigma_1^d, \dots, \sigma_l^d \vdash \psi$, the following part in a proof on the left is reduced to the part on the right, where $\vec{\mathcal{D}}$ is the unified diagram $\mathcal{D}_1 + \mathcal{D}_2 + \cdots + \mathcal{D}_n$:

$$\dfrac{\overset{\vdots}{\mathcal{E}} \qquad \overset{\vdots}{\dfrac{\overset{\vdots}{\mathcal{D}_1}}{\varphi_1}\,\mathsf{obs} \cdots \dfrac{\overset{\vdots}{\mathcal{D}_n}}{\varphi_n}\,\mathsf{obs} \quad \overset{\vdots}{\sigma_1^d} \cdots \overset{\vdots}{\sigma_l^d}}\;\psi}{\mathcal{E}+\psi}\,\mathsf{app} \quad \rhd \quad \dfrac{\dfrac{\overset{\vdots}{\mathcal{E}} \qquad \dfrac{\dfrac{\overset{\vdots}{\mathcal{D}_1} \cdots \overset{\vdots}{\mathcal{D}_n}}{\vec{\mathcal{D}}}\,\mathsf{uni} \quad \overset{\vdots}{\sigma_1^d} \cdots \overset{\vdots}{\sigma_l^d}}{\vec{\mathcal{D}} + \vec{\sigma^d}}\,\mathsf{app}}{\mathcal{E} + \vec{\mathcal{D}} + \vec{\sigma^d}}\,\mathsf{uni}}{\mathcal{E}+\psi}\,\mathsf{del}$$

In the above reduced proof on the right, the double line ═══ uni refers to repeated applications of uni to $\mathcal{D}_1, \dots, \mathcal{D}_n$; the same applies to ═══ app. Each σ_i^d is a diagrammatic formula independent of obs. Since ψ may depend not only on diagrams $\mathcal{D}_1, \dots, \mathcal{D}_n$ but also on formulas $\sigma_1^d, \dots, \sigma_l^d$, the similar structure of our detour may occur in such a part. Thus, we generalize our reduction by including σ_i^d to reduce such a part; cf. Fig. 1 of Example 3.9.

The same applies to another dual pair of del and uni. Moreover, since del and obs share a similar structure (i.e., extraction of information), a pair of del and app may form the same detour as before although del and app are not in duality. Thus, for removing the detours of the same kind in a proof, we define our reduction for pairs of + (app and uni) and − (obs, del, and $*E$).

We abbreviate a set of proofs $\overset{\vdots\; t_1}{\alpha_1}, \dots, \overset{\vdots\; t_n}{\alpha_n}$ collectively as $\overset{\vdots\; \vec{t}}{\vec{\alpha}}$ with an appropriate number n. On the one hand, we denote a set of proofs by $\vec{t}$, on the other hand, we denote a unification of diagrams and formulas $\alpha_1 + \cdots + \alpha_n$ by $\vec{\alpha}$.

In the expression of substitution, say $s[x^\alpha := t]$, we always assume that t is free for s as usual. We denote simultaneous substitution $s[x_1^{\alpha_1} := t_1, \dots, x_n^{\alpha_n} := t_n]$ by $s[\overrightarrow{x^\alpha := t}]$, and a sequence of simultaneously substituted terms $s_1[\overrightarrow{x^\alpha := t}], \dots, s_n[\overrightarrow{x^\alpha := t}]$ by $\vec{s}\,[\overrightarrow{x^\alpha := t}]$.

Definition 3.2 **$\pm$-reduction rule** in basic HL is of the following form.

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \qquad \begin{array}{c}\dfrac{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array}}{\beta}(-) \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}\\ \vdots\, s\\ \alpha\end{array}}{\mathcal{E}+\alpha}(+) \quad \text{where } \vec{t_1} \neq \emptyset \qquad \rhd \qquad \dfrac{\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \quad \genfrac{}{}{1.5pt}{}{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}}{\vec{\mathcal{D}}+\vec{\gamma}}(+)}{\mathcal{E}+\vec{\mathcal{D}}+\vec{\gamma}}(+)}{\mathcal{E}+\alpha}(-)$$

$$+(u, s\{\overrightarrow{x^\beta := -(t_1)}, \overrightarrow{y^\gamma := t_2}\}) \quad \rhd \quad -(+(u, +(\vec{t_1}, \vec{t_2})))$$

Here, in the term notation $s\{\overrightarrow{x^\beta := -(t_1)}, \overrightarrow{y^\gamma := t_2}\}$, the term s contains only $\vec{x^\beta}, \vec{y^\gamma}$ as its term-variables, although not all of them should appear in s. We express this condition using braces { } that is different from the usual square brackets [ ] expressing substitution. The above part on the left in a given proof is called a **redex**.

Although the following is the other form of $\pm$-reduction, since we assume the operation $+$ to be commutative, we mainly consider the above reduction rule in what follows.

$$\dfrac{\begin{array}{c}\dfrac{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array}}{\beta}(-) \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}\\ \vdots\, s\\ \mathcal{E}\end{array} \qquad \begin{array}{c}\vdots\, u\\ \alpha\end{array}}{\mathcal{E}+\alpha}(+) \text{where } \vec{t_1} \neq \emptyset \qquad \rhd \qquad \dfrac{\dfrac{\genfrac{}{}{1.5pt}{}{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}}{\vec{\mathcal{D}}+\vec{\gamma}}(+) \quad \begin{array}{c}\vdots\, u\\ \alpha\end{array}}{\vec{\mathcal{D}}+\vec{\gamma}+\alpha}(+)}{\mathcal{E}+\alpha}(-)$$

$$+(s\{\overrightarrow{x^\beta := -(t_1)}, \overrightarrow{y^\gamma := t_2}\}, u) \quad \rhd \quad -(+(+(\vec{t_1}, \vec{t_2}), u))$$

Example 3.3 $\pm$-reduction is performed as follows. See also Example 3.9.

$$\dfrac{\mathcal{E}^x \qquad \dfrac{\dfrac{\mathcal{D}^y}{\sigma}(-) \quad \varphi^z}{\sigma \wedge \varphi}}{\mathcal{E}+(\sigma \wedge \varphi)}(+) \qquad \rhd \qquad \dfrac{\dfrac{\mathcal{E}^x \quad \dfrac{\mathcal{D}^y \quad \varphi^z}{\mathcal{D}+\varphi}(+)}{\mathcal{E}+(\mathcal{D}+\varphi)}}{\mathcal{E}+(\sigma \wedge \varphi)}(-)$$

$$+(x, \langle -(y), z\rangle) \quad \rhd \quad -(+(x, +(y, z)))$$

Example 3.4 The following proof is expressed as $+(u, \lambda x. -(+(t_1, t_2[x])))$ in our term notation, where variable x appears in t_2. These applications of $-$-rule and $+$-rule do not constitute a redex because $-(+(t_1, t_2[x]))$ is not free for $\lambda x.z$ when we regard $+(u, (\lambda x.z)[z := -(+(t_1, t_2[x]))])$.

$$\dfrac{\begin{matrix}\vdots\; u\\ \mathcal{E}\end{matrix} \qquad \dfrac{\dfrac{\dfrac{\begin{matrix}\vdots\; t_1\\ \mathcal{D}\end{matrix}\quad \begin{matrix}[\varphi_1]^x\\ \vdots\; t_2\\ \varphi_2\end{matrix}}{\mathcal{D}+\varphi_2}\,(+)}{\varphi_3}\,(-)}{\varphi_1\to\varphi_3}\,x}{\mathcal{E}+(\varphi_1\to\varphi_3)}\,(+)$$

Such a case, where the "free for condition" is not satisfied exclusively occurs by applications of FOL-rules to construct app-formulas. (Cf. Proposition 3.13.)

In a given proof, our redex is not uniquely determined. There are several redexes with respect to the same application of app/uni. For example, in the following form of a proof, each of del and obs induces a redex with respect to the same app.

$$\dfrac{\begin{matrix}\vdots\; u\\ \mathcal{E}\end{matrix}\qquad \begin{matrix}\vdots\; t\\ \dfrac{\dfrac{\mathcal{D}}{\mathcal{F}}\,\mathsf{del}}{\varphi}\,\mathsf{obs}\\ \vdots\; s\\ \psi\end{matrix}}{\mathcal{E}+\psi}\,\mathsf{app}$$

Thus, to enumerate redexes in a given proof, we count every possible redex, such as pairs of del-app (i.e., $s\{x^{\mathcal{F}} := -(t)\}$) and obs-app (i.e., $s\{y^{\varphi} := -(-(t))\}$) in the above proof. Thus, the above proof contains two redexes (explicitly).

Definition 3.5 (Normal form) A (heterogeneous) proof that does not contain any redex is called a **normal (heterogeneous) proof**.

In accordance with our Postulate 2.1, in which we regard our diagram as the deductively closed conjunction of formulas, our ±-reduction rule can be understood by decomposing it into two operations of deductive closure and ±-permutation. Based on this consideration, we show the subject reduction property as follows.

Lemma 3.6 (Deductively closed) *If* $s\{\overrightarrow{x^{\beta} := -(t_1)}, \overrightarrow{y^{\gamma} := t_2}\}$ *is a proof of* α*, then so is* $-(+(\vec{t_1}, \vec{t_2}))$.

Proof. Assume that given s is the following form on the left. Then, we obtain the following proof on the right, where the last application of $-$ is valid since the diagram $\vec{\mathcal{D}} + \vec{\gamma}$ is deductively closed.

$$\begin{matrix}\begin{matrix}\vdots\; \vec{t_1}\\ \dfrac{\mathcal{D}}{\beta}(-)\end{matrix}\quad \begin{matrix}\vdots\; \vec{t_2}\\ \gamma\end{matrix}\\ \vdots\; s\\ \alpha\end{matrix} \qquad\qquad \dfrac{\dfrac{\begin{matrix}\vdots\; \vec{t_1}\\ \mathcal{D}\end{matrix}\quad\begin{matrix}\vdots\; \vec{t_2}\\ \gamma\end{matrix}}{\vec{\mathcal{D}}+\vec{\gamma}}\,(+)}{\alpha}\,(-)$$

■

Lemma 3.7 (±-permutation) *If* $+(u, -(t))$ *is a proof of* $\mathcal{E}+\alpha$*, then so is* $-(+(u,t))$.

Proof. Assume that $+(u, -(t))$ is the following form on the left. We can then obtain the following proof on the right, since $\alpha \in |\mathcal{D}|$.

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \quad \dfrac{\begin{array}{c}\vdots\, t\\ \mathcal{D}\end{array}}{\alpha}(-)}{\mathcal{E}+\alpha}(+) \qquad\qquad \dfrac{\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \quad \begin{array}{c}\vdots\, t\\ \mathcal{D}\end{array}}{\mathcal{E}+\mathcal{D}}(+)}{\mathcal{E}+\alpha}(-)$$

∎

Proposition 3.8 (Subject reduction) *If* $+(u, s\{\overrightarrow{x^\beta := -(t_1)}, \overrightarrow{y^\gamma := t_2}\})$ *is a proof of* $\mathcal{E}+\alpha$*, then so is* $-(+(u, +(\vec{t_1}, \vec{t_2})))$*.*

Proof. Assume that the given proof is the following form on the left. Then, we obtain the proof in the middle by Lemma 3.6. We further obtain the proof on the right by Lemma 3.7.

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \quad \begin{array}{c}\dfrac{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array}}{\beta}(-) \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}\\ \vdots\, s\\ \alpha\end{array}}{\mathcal{E}+\alpha}(+) \qquad \dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \quad \dfrac{\dfrac{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}}{\vec{\mathcal{D}}+\vec{\gamma}}(+)}{\alpha}(-)}{\mathcal{E}+\alpha}(+) \qquad \dfrac{\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{E}\end{array} \quad \dfrac{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array}}{\vec{\mathcal{D}}+\vec{\gamma}}(+)}{\mathcal{E}+\vec{\mathcal{D}}+\vec{\gamma}}(+)}{\mathcal{E}+\alpha}(-)$$

∎

Example 3.9 (Euler diagrams) We obtain the normal proof on the right by reducing the obs-app pair of the following proof on the left in Fig. 1.

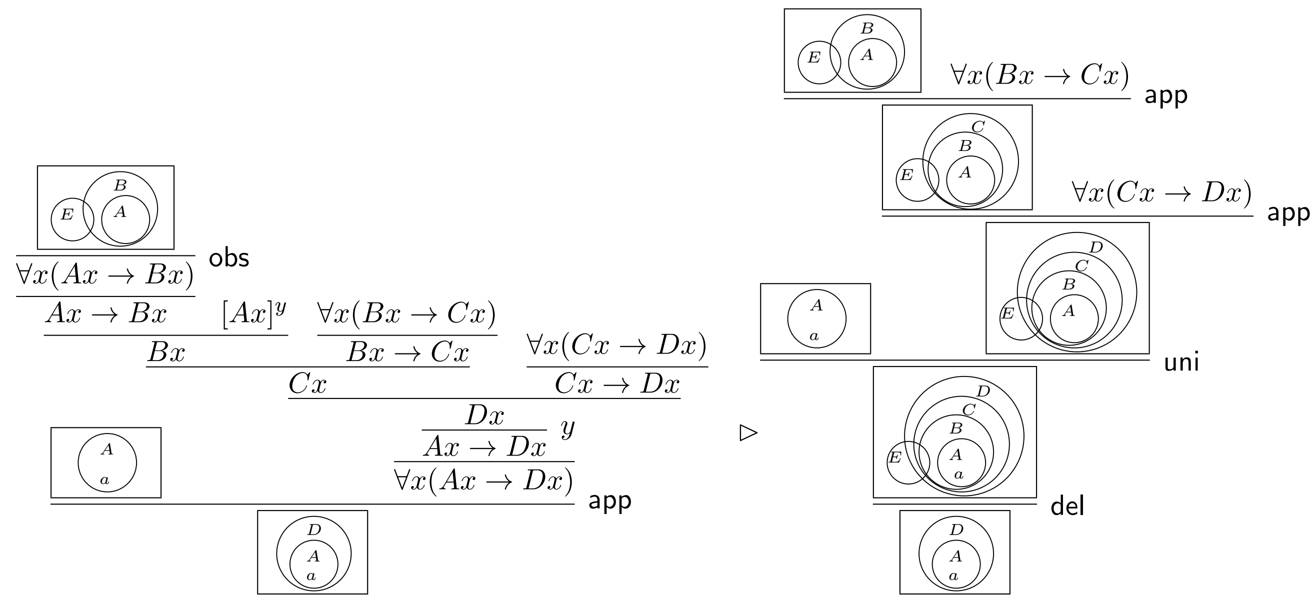


Fig. 1 ±-reduction

3.3 Normalization for basic HL

We show the normalization theorem for our basic HL. Note that the usual FOL-reductions and our $\pm$-reduction are performed separately, that is, $\pm$-reduction does not provide any FOL-redex, and FOL-reductions do not provide any $\pm$-redex. Thus, it is sufficient to prove the normalization theorem only for the diagrammatic part, since that for FOL part is already established.

Our $\pm$-reduction provides numerous new applications of $+$-rule; hence, by reducing a redex, numerous new redexes may be provided in general (when there are applications of $-$-rule above the redex in question). However, by successively applying $\pm$-reduction to the topmost-leftmost redex in a given proof, we can prove the weak normalization theorem for our basic HL (cf. [27]). In this article, instead of fixing a reduction strategy, we generalize our $+$-rule and consider a kind of parallel $+$-rule, which blocks the generation of new applications of $+$-rule, and we show essentially the strong normalization theorem.

Theorem 3.10 (Normalization for basic HL) *Let α, α_i be a diagram or a formula. Any proof of α from $\alpha_1, \ldots, \alpha_n$ is reduced to a normal proof of α from $\alpha_1, \ldots, \alpha_n$.*

Proof. We generalize our $+$-rule to a parallel rule $\oplus$ as follows, and we abbreviate the rule as it appears on the right.

$$\dfrac{\begin{array}{c}\vdots\, t_1\\ \mathcal{D}_1\end{array}\quad\cdots\quad\begin{array}{c}\vdots\, t_n\\ \mathcal{D}_n\end{array}\quad\begin{array}{c}\vdots\, s_1\\ \sigma_1\end{array}\quad\cdots\quad\begin{array}{c}\vdots\, s_m\\ \sigma_m\end{array}}{\vec{\mathcal{D}}+\vec{\sigma}}\,(\oplus)\qquad\qquad \dfrac{\begin{array}{c}\vdots\, \vec{t}\\ \mathcal{D}\end{array}\quad\begin{array}{c}\vdots\, \vec{s}\\ \sigma\end{array}}{\vec{\mathcal{D}}+\vec{\sigma}}\,(\oplus)$$

$$\oplus(t_1, \ldots, t_n, s_1, \ldots, s_m)\qquad\qquad\qquad \oplus(\vec{t}, \vec{s})$$

The case $n \geq 2$ and $m = 0$ is the generalization of uni, and the case $n \geq 1$ and $m \geq 1$ is the generalization of app. The $\pm$-reduction rule is generalized as follows:

$$\dfrac{\begin{array}{c}\vdots\, \vec{u}\\ \mathcal{E}\end{array}\quad\begin{array}{c}\dfrac{\begin{array}{c}\vdots\, \vec{t}\\ \mathcal{D}\end{array}}{\beta}\,(-)\quad\begin{array}{c}\vdots\, \vec{s}\\ \gamma\end{array}\\ \vdots\, \vec{v}\\ \alpha\end{array}}{\vec{\mathcal{E}}+\vec{\alpha}}\,(\oplus)\quad\rhd\quad\dfrac{\dfrac{\begin{array}{c}\vdots\, \vec{u}\\ \mathcal{E}\end{array}\quad\begin{array}{c}\vdots\, \vec{t}\\ \mathcal{D}\end{array}\quad\begin{array}{c}\vdots\, \vec{s}\\ \gamma\end{array}}{\vec{\mathcal{E}}+\vec{\mathcal{D}}+\vec{\gamma}}\,(\oplus)}{\vec{\mathcal{E}}+\vec{\alpha}}\,(-)$$

$$\oplus(\vec{u}, \vec{v}\{\overrightarrow{x^{\beta} := -(t)}, \overrightarrow{y^{\gamma} := s}\})\quad\rhd\quad -(\oplus(\vec{u}, \vec{t}, \vec{s}))$$

Note that $\vec{v}$ is a sequence of proofs such that $v_1\{\overrightarrow{x^{\beta} := -(t)}, \overrightarrow{y^{\gamma} := s}\}, \ldots, v_n\{\overrightarrow{x^{\beta} := -(t)}, \overrightarrow{y^{\gamma} := s}\}$.

By applying the above generalized $\pm$-reduction, the number of applications of $-$ above the $\oplus$ is reduced; accordingly, the number of redexes is reduced. Therefore, by induction on the number of redexes in the given proof, it is shown that any reduction sequence terminate in finite steps.

Thus, for a given proof, we repeatedly apply the generalized $\pm$-reduction rule and obtain a proof with some applications of $\oplus$ without any redex. The proof with $\oplus$ is easily transformed into the usual normal proof without $\oplus$, and we obtain the theorem. ∎

3.4 Structure of normal heterogeneous proofs in basic HL

Let us investigate how diagrammatic inference and FOL inference appear, and are related in our heterogeneous proofs. Following [6], let us define the notion of tracks. To make our discussion simple, we concentrate on the diagrammatic part by ignoring the FOL part. (Thus, the usual subformula property of FOL is not obtained from the following definition of tracks.)

Definition 3.11 (Track) A **path** in a proof is a sequence of formulas and diagrams $\alpha_1, \ldots, \alpha_n$ such that (1) α_1 is an open or closed assumption; (2) α_n is the conclusion of the proof; (3) α_i is a premise immediately above α_{i+1} for $1 \leq i \leq n-1$.
A **track** is an initial part of a path which stops at the first app-formula or at the conclusion.

Lemma 3.12 (Track) *In every track in a normal proof in basic* HL*, no application of* $-$*-rule can precede an application of* $+$*-rule.*

Proof. Assume that there is an application of $-$-rule in the given track.
(1) When the given application of $-$ is obs, assume to the contrary that there is an application of $+$ under the obs. We assume the $+$ is the first one under the obs. Since the given track does not end at the app-formula of the given $+$, it is the following form of u_1, u_2:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\, u_1\\ \mathcal{D}_1\end{array}}{\begin{array}{c}\varphi_1\\ \vdots\, u_2\\ \mathcal{D}_2\end{array}}\,\mathsf{obs} \qquad \begin{array}{c}\vdots\\ \alpha\end{array}}{\mathcal{D}_2 + \alpha}\,(+)$$

Since formula φ_1 should be transformed into diagram $\mathcal{D}_2$ in u_2, there should be another application of $+$ in u_2, which contradicts to the assumption that the above $+$ is the first one under the obs. Thus, no application of obs can precede an application of $+$.
(2) When the given application of $-$ is del, assume that there is an application of $+$ under the del. We assume the $+$ is the first one under the del. Since there exists no obs before the $+$, the given track is the following form of u_1, u_2:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\, u_1\\ \mathcal{D}_1\end{array}}{\begin{array}{c}\mathcal{D}_2\\ \vdots\, u_2\\ \mathcal{D}_3\end{array}}\,\mathsf{del} \qquad \begin{array}{c}\vdots\\ \alpha\end{array}}{\mathcal{D}_3 + \alpha}\,(+)$$

Since the above $+$ is the first one under the del, u_2 consists only of applications of del. Thus, no variable of u_1 becomes bound in the process to construct u_2; hence the above pair of del and $+$ comprises a redex, which contradicts to the assumption that given proof is normal.

The same applies to the case wherein the given $-$ is $*E$.

Therefore, no application of $-$-rule can precede an application of $+$-rule. ∎

Proposition 3.13 (Normal proof in basic HL) *In any normal proof t, every track in t has the form (1) $\varphi_1, \ldots, \varphi_k$ (when the track starts from a formula), or (2) $\mathcal{D}_1, \ldots, \mathcal{D}_l, \mathcal{E}_1, \ldots, \mathcal{E}_m$, $\psi_1, \ldots, \psi_n$ (when the track starts from a diagram), where*

- $\varphi_2, \ldots, \varphi_k$ *and* $\psi_2, \ldots, \psi_n$ *are obtained by FOL-rules (FOL part);*
- $\mathcal{D}_2, \ldots, \mathcal{D}_l$ *are obtained by +-rule (+ part);*
- $\mathcal{E}_1, \ldots, \mathcal{E}_m, \psi_1$ *are obtained by −-rule (− part).*

Proof. (1) When the given track starts from a formula φ_1, only FOL-rules are applicable to obtain $\varphi_2, \ldots, \varphi_k$, where φ_k is an app-formula or the conclusion of t. This is because every track stops at an app-formula or the conclusion; hence, no formulas are transformed into a diagram in any track starting from a formula.
(2) When the given track starts from a diagram $\mathcal{D}_1$, rules of app, uni, obs, del are applicable. By Lemma 3.12, we assume without loss of generality, that $+$ is applied certain times to $\mathcal{D}_1, \ldots, \mathcal{D}_{l-1}$, and then $-$ is applied to $\mathcal{D}_l$. $\mathcal{D}_2, \ldots, \mathcal{D}_l$ may be $*$, and if it appears once, $*$ continues to appear until $\mathcal{D}_l$.

obs and del are applicable to $\mathcal{D}_l$. Since obs transforms a diagram into a formula, we assume without loss of generality, that del (or $*E$ when $\mathcal{D}_l$ is $*$) is applied certain times to $\mathcal{D}_l, \mathcal{E}_1, \ldots, \mathcal{E}_{m-1}$, and then obs is applied to $\mathcal{E}_m$.

Then, ψ_1 is obtained from $\mathcal{E}_m$ by obs, and by the same reason as (1) above, only FOL-rules are applied to $\psi_1, \ldots, \psi_{n-1}$ to obtain ψ_n. ■

Thus, in our basic HL, every normal proof has the following form.

$$
\begin{array}{c}
\dfrac{\dfrac{\mathcal{D}_1 \quad \begin{array}{c}\varphi_1\\ \vdots\ \text{FOL}\\ \varphi_k\end{array}}{\mathcal{D}_2}\,(+) \qquad \begin{array}{c}\vdots\\ \alpha\end{array}}{\mathcal{D}_3}\,(+)\\
\vdots\ (+)\\
\dfrac{\mathcal{D}_l}{\mathcal{E}_1}\ \mathsf{del}/*E\\
\vdots\ \mathsf{del}\\
\dfrac{\mathcal{E}_m}{\psi_1}\ \mathsf{obs}\\
\vdots\ \text{FOL}\\
\psi_n
\end{array}
$$

When ψ_n is an app-formula, the whole structure as above appears as a part of a normal proof.

Normal proofs of FOL have the subformula property, and we find that sentential inference comprises decomposition of given assumptions and construction of a conclusion. In contrast, from Proposition 3.13, we find that diagrammatic inference comprises construction of a diagram by unifying given assumptions and extracting a conclusion. Thus, we have the following maximal diagram property, which may be considered to be a diagrammatic counterpart of the subformula property.

Corollary 3.14 (Maximal diagrams) *Let t be a track $\mathcal{D}_1, \ldots, \mathcal{D}_l, \mathcal{E}_1, \ldots, \mathcal{E}_m, \psi_1, \ldots, \psi_n$ in a normal proof. Then, there exists a diagram $\mathcal{D}$ in t such that $\mathcal{D}_i, \mathcal{E}_j \sqsubseteq \mathcal{D}$ for $1 \leq i \leq l, 1 \leq j \leq m$, or there exists a sequence of $*$ in t.*

Proof. First observe that premises of +-rule are subdiagrams of its conclusion, and the conclusion of del is a subdiagram of its premise.

By Proposition 3.13, $\mathcal{D}_2, \ldots, \mathcal{D}_l$ are obtained by +-rule, and hence $\mathcal{D}_i \sqsubseteq \mathcal{D}_l$ for all $1 \leq i \leq l-1$. $\mathcal{E}_1, \ldots, \mathcal{E}_m$ are obtained by del, and hence $\mathcal{E}_j \sqsubseteq \mathcal{D}_l$ for all $1 \leq j \leq m$. Therefore, $\mathcal{D}_l$ is the maximal diagram of the given track.

As shown in Proposition 3.13, $*$ appears, if any, at $\mathcal{D}_2, \ldots \mathcal{D}_l$, and if it appears once, $*$ continues to appear until $\mathcal{D}_l$. ■

Let us call a track that reaches to the given conclusion a *main track*; and a track that ends with an app-formula a *minor track*. A main track comprises the main part of diagrammatic reasoning, where a maximal diagram is constructed and the conclusion is extracted; and every minor track constructs an app-formula.

There exists a maximal diagram in every track, and it is impossible in general to compare the complexities of such maximal diagrams in a proof.

4 Extension of basic heterogeneous logic

In Section 4.1, we extend our basic HL to HLe by extending the +-rule in the style of general elimination rules in FOL. In Section 4.2, we prove the normalization theorem for extended HLe. In Section 4.3, we investigate the structure of normal proofs, and show the maximal diagram property of HLe.

4.1 Extended HLe

Other than the case of contradiction, our +-rule (app and uni) in basic HL may provide $*$. One of the typical cases is the case wherein the result of unification between $\mathcal{D}$ and φ is not uniquely determined as a single diagram. For example, in our Euler diagrammatic system [13], we cannot apply our app-rule to the following diagram $\mathcal{D}$ in Fig. 2 and a formula $\forall x(Ax \to Bx)$. This is because whether point a is included by A is not determined in the unified diagram. In such a case, the result of unification of $\mathcal{D}$ and φ becomes $*$ in basic HL.

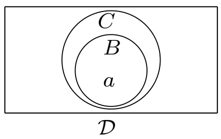


Fig. 2 Indeterminacy

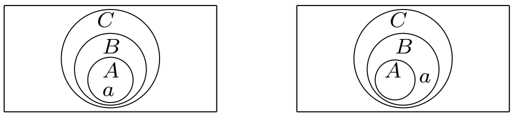
There are several options to avoid such a difficulty: (i) we may prohibit such an application of +-rule (e.g., [13]); (ii) we may extend the notion of a diagram so that disjunctive information can be expressed by introducing certain devices such as linking between points and diagrams (e.g., [23, 10]); (iii) we may extend inference rules, instead of diagrams, to deal with disjunctive information (e.g., [5]). This article adopts (iii) with our Postulate 2.1. To this end, we first define exhaustive cases, which refer to a disjunction of diagrams equivalent to the deductive closure of $\mathcal{D}$ and α at the level of types.

Definition 4.1 (Exhaustive cases) Let $\mathcal{D}, \mathcal{E}_1, \ldots, \mathcal{E}_n$ be diagrams, and α be a diagram or a formula. When $ob(\mathcal{D}) \cup ob(\alpha) = ob(\mathcal{E}_i)$ for $1 \leq i \leq n$ and $\{\varphi^d \mid |\mathcal{D}|, |\alpha| \vdash \varphi^d$ in intuitionistic FOL$\} \leftrightarrow |\mathcal{E}_1| \vee \cdots \vee |\mathcal{E}_n|$ holds in intuitionistic FOL, we call $\mathcal{E}_1, \ldots, \mathcal{E}_n$ **exhaustive cases** for $\mathcal{D}$ plus α. We denote $\mathcal{E}_i$ by $(\mathcal{D} + \alpha)_i$.

In particular, when there is no exhaustive cases for $\mathcal{D}$ plus α, we express it by the special diagram $*$.

Note that we restrict the FOL provability relation to that in "intuitionistic logic." This is because the disjunction property of intuitionistic logic is applied in the proof of Lemma 4.5 (Subject reduction). In what follows, for the FOL provability in our extended HLe, we mean the intuitionistic provability.

Example 4.2 Exhaustive cases for $\mathcal{D}$ in Fig. 2 plus $\forall x(Ax \to Bx)$ are the following two diagrams.



$(\mathcal{D} + \forall x(Ax \to Bx))_1$ $(\mathcal{D} + \forall x(Ax \to Bx))_2$

Using the exhaustive cases, we extend +-rule as follows.

Definition 4.3 Extended +-rule (app and uni) as well as its term notation in HLe are of the following forms. Let $(\mathcal{D}+\alpha)_1, \ldots, (\mathcal{D}+\alpha)_n$ be exhaustive cases for $\mathcal{D}$ plus α.

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, t\\ \alpha\end{array} \quad \begin{array}{c}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_1\\ \beta\end{array} \quad \cdots \quad \begin{array}{c}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_n\\ \beta\end{array}}{\beta}\,(+_x) \qquad +_x\,(u, t; s_1, \ldots, s_n)$$

where every x in s_i $(1 \leq i \leq n)$ is bound by $+_x$. In this rule, $\mathcal{D}$ and α are major premises and $\beta, \ldots, \beta$ are minor premises. In our term notation, we divide proofs of major premises (u and t) and minor premises ($s_1, \ldots, s_n$) by a semicolon ; . Furthermore, we call exhaustive cases *auxiliary assumptions* in a given proof in accordance with the $\vee E$-rule of FOL.

The +-rule in basic HL, where the conclusion diagram is uniquely determined and hence exhaustive cases is a singleton, is the following special case of the extended +-rule:

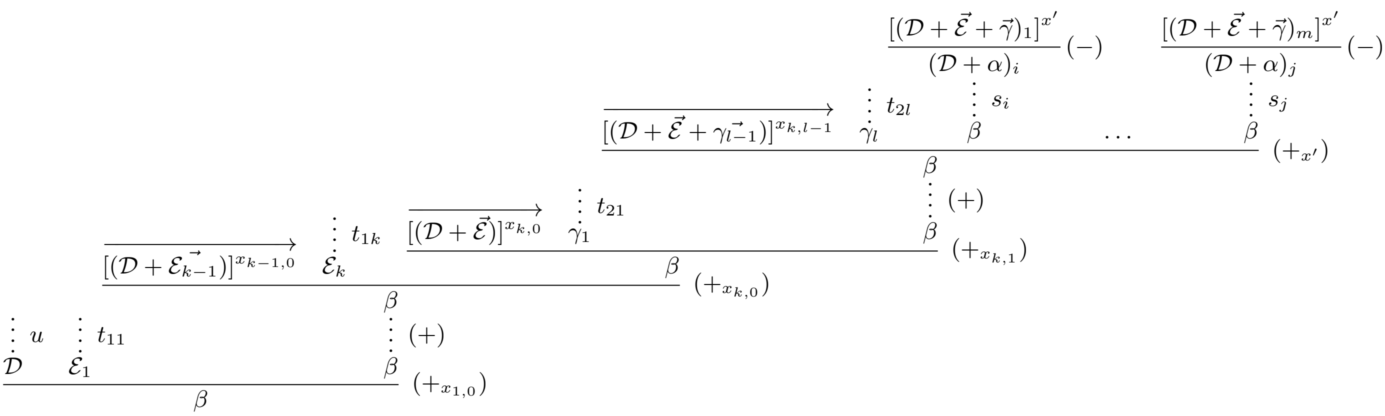
$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, t\\ \alpha\end{array} \quad [\mathcal{D}+\alpha]^x}{\mathcal{D}+\alpha}\,(+_x) \qquad +_x\,(u, t; x)$$

$\pm$-reduction is also extended as follows:

Definition 4.4 Extended $\pm$-reduction in HLe is of the following form. The following part in a proof is called **redex**:

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{D}\end{array} \quad \dfrac{\begin{array}{c}\vdots\, \vec{t_1}\\ \mathcal{E}\end{array}}{\delta}\,(-) \;\; \begin{array}{c}\vdots\, \vec{t_2}\\ \gamma\end{array} \atop {\vdots\, t \atop \alpha} \quad \begin{array}{c}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_1\\ \beta\end{array} \quad \cdots \quad \begin{array}{c}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_n\\ \beta\end{array}}{\beta}\,(+_x)$$

is reduced to the following form, where we assume $\vec{\mathcal{E}} = \mathcal{E}_1 + \cdots + \mathcal{E}_k$ and $\vec{\gamma} = \gamma_1 + \cdots + \gamma_l$. $s_i, \ldots, s_j \in \{s_1, \ldots, s_n\}$. In the following $x_{p,q}$, p ranges from 1 to k of $\vec{\mathcal{E}}$ and q ranges from 0 to l of $\vec{\gamma}$:



The term notation of $\pm$-reduction in HLe is as follows:

$$+_x(u, t\{\overrightarrow{y := -(t_1)}, \overrightarrow{z := t_2}\}; s_1, \ldots, s_n)$$

$$\rhd \quad +_{x_{1,0}}(u, t_{11}; \cdots; +_{x_{k,0}}(\overrightarrow{x_{k-1,0}}, t_{1k}; +_{x_{k,1}}(\overrightarrow{x_{k,0}}, t_{21}; \cdots$$
$$; +_{x'}(\overrightarrow{x_{k,l-1}}, t_{2l}; s_i[x := -(x')], \ldots, s_j[x := -(x')]) \cdots)) \cdots)$$

where $\vec{t_1}, \vec{t_2}$ are free for t. Every bound variable x in the original s_i is replaced by x' after the reduction, and x' in $s_i, \ldots, s_j$ are bound by $+_{x'}$. $+_{x'}(\overrightarrow{x_{k,l-1}}, t_{2l}; s_i[x := -(x')], \ldots, s_j[x := -(x')])$ means that the whole structure $+_{x'}(x_{k,l-1}, t_{2l}; s_i[x := -(x')], \ldots, s_j[x := -(x')])$ is repeated by the number of times depending on the number of exhaustive cases for $\mathcal{D} + \vec{\mathcal{E}} + \overrightarrow{\gamma_{l-2}}$ plus γ_{l-1}. In the expression $s_i[x := -(x')]$, proof s_i may contain variables other than x.

The above $\pm$-reduction is essentially the same as that in basic HL. When α is deduced from $\vec{\mathcal{E}}$ and $\vec{\gamma}$ by using $-$-rule, and α is unified with $\mathcal{D}$, we obtain the same result without deducing α, but by directly unifying $\vec{\mathcal{E}}, \vec{\gamma}$, and $\mathcal{D}$. Although the form of extended $+$-rule makes the appearance of proof complicated, the above proof after the reduction just expresses successive applications of $+$-rule to $\mathcal{D}, \mathcal{E}_1, \ldots, \mathcal{E}_k$ and $\gamma_1, \ldots, \gamma_l$ one by one.

After the reduction, which proofs $s_i, \ldots, s_j$ of $s_1, \ldots, s_n$ in the original proof appear depends on the exhaustive cases for $\mathcal{D}$, $\vec{\mathcal{E}}$, plus $\vec{\gamma}$. This is shown in the following proof of the subject reduction property. (See also Example 4.6.)

Lemma 4.5 (Subject reduction) *The extended $\pm$-reduction in* HLe *preserves the conclusion of a proof.*

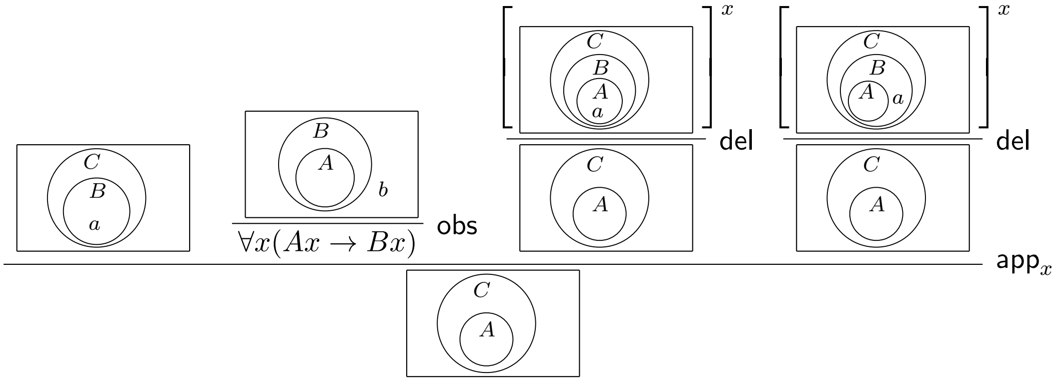
Proof. To show that the applications of $-$-rule in the reduced proof are valid, we prove for all $1 \leq i \leq m$, there exists $1 \leq j \leq n$ such that $|(\mathcal{D} + \vec{\mathcal{E}} + \vec{\gamma})_i| \vdash |(\mathcal{D} + \alpha)_j|$.

By the assumption on the original proof, we have $|\mathcal{E}_1|, \ldots, |\mathcal{E}_k|, |\gamma_1|, \ldots, |\gamma_l| \vdash |\alpha|$, which implies $|\mathcal{D}|, |\mathcal{E}_1|, \ldots, |\mathcal{E}_k|, |\gamma_1|, \ldots, |\gamma_l| \vdash |\mathcal{D}| \wedge |\alpha|$.

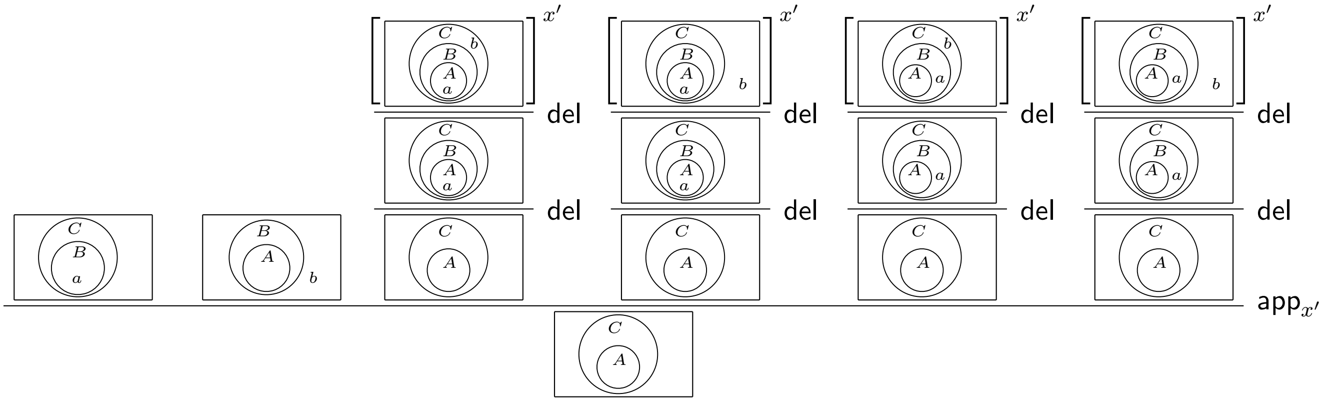
Let $Cl(\varphi_1, \ldots, \varphi_n)$ be the deductive closure of $\varphi_1, \ldots, \varphi_n$. Then, by the definition of exhaustive cases, we have $\bigvee_i |(\mathcal{D} + \vec{\mathcal{E}} + \vec{\gamma})_i| = Cl(|\mathcal{D}|, |\mathcal{E}_1|, \ldots, |\mathcal{E}_k|, |\gamma_1|, \ldots, |\gamma_l|)$ and $\bigvee_j |(\mathcal{D} + \alpha)_j| = Cl(|\mathcal{D}|, |\alpha|)$. Hence, we have $\bigvee_i |(\mathcal{D} + \vec{\mathcal{E}} + \vec{\gamma})_i| \vdash \bigvee_j |(\mathcal{D} + \alpha)_j|$, which implies $|(\mathcal{D} +$

$\vec{\mathcal{E}}+\vec{\gamma})_i| \vdash \bigvee_j |(\mathcal{D}+\alpha)_j|$ for all $1 \le i \le m$. Then, by the disjunction property of intuitionistic logic, there exists $1 \le j \le n$ such that $|(\mathcal{D}+\vec{\mathcal{E}}+\vec{\gamma})_i| \vdash |(\mathcal{D}+\alpha)_j|$. ■

Example 4.6 (Euler diagrams) Our extended ±-reduction in HLe can be applied in our Euler diagrammatic system as follows:



is reduced to the following:



Remark 4.7 The following applications of $-$ and $+$ ($+_x$ and $+_y$) do not constitute any redex.

$$\dfrac{\dfrac{\overset{\vdots\, u}{\mathcal{D}} \quad \overset{\vdots\, t_1}{\alpha} \quad \dfrac{[\mathcal{D}+\alpha]^x}{\mathcal{E}}\,(-)}{\mathcal{E}}\,(+_x) \quad \overset{\vdots\, t_2}{\gamma} \quad \overset{\overset{[\mathcal{E}+\gamma]^y}{\vdots\, s}}{\beta}}{\beta}\,(+_y)$$

In our term notation, the above proof is $+_y(+_x(u, t_1; -(x)), t_2; s)$. If we consider $+_x(u, t_1; z)[z := -(x)]$ then $-(x)$ is not free for $+_x(u, t_1; z)$; hence, it is not a redex. In contrast, the following proof contains a redex (marked by boxes).

$$\dfrac{\overset{\vdots\, u}{\mathcal{D}} \quad \overset{\vdots\, t_1}{\alpha} \quad \dfrac{\dfrac{[\mathcal{D}+\alpha]^x}{\mathcal{E}}\,\boxed{(-)} \quad \overset{\vdots\, t_2}{\gamma} \quad \overset{\overset{[\mathcal{E}+\gamma]^y}{\vdots\, s}}{\beta}}{\beta}\,\boxed{(+_y)}}{\beta}\,(+_x)$$

The above proof is $+_x(u, t_1; +_y(-(x), t_2; s))$; if we consider $+_y(z, t_2; s)[z := -(x)]$ then $-(x)$ is free for $+_y(z, t_2; s)$. Hence, it constitutes a redex.

4.2 Normalization for HLe

Our extended ±-reduction may provide a new redex. Let us consider the following proof, where we assume that every exhaustive case is single for simplicity.

$$\dfrac{\begin{matrix}\vdots\, t_1\\ \mathcal{D}\end{matrix} \quad \dfrac{\begin{matrix}\vdots\, t_2\\ \mathcal{E}\end{matrix}}{\varphi}\boxed{(-)} \quad \dfrac{[\mathcal{D}+\varphi]^x \quad \begin{matrix}\vdots\, t_3\\ \sigma\end{matrix} \quad [\mathcal{D}+\varphi+\sigma]^y}{\mathcal{D}+\varphi+\sigma}\,(+_y)}{\mathcal{D}+\varphi+\sigma}\,\boxed{(+_x)} \qquad +_x\,(t_1, -(t_2); +_y(x, t_3; y))$$

By reducing the above redex, we obtain the following proof, where a new redex is produced:

$$\dfrac{\begin{matrix}\vdots\, t_1\\ \mathcal{D}\end{matrix} \quad \begin{matrix}\vdots\, t_2\\ \mathcal{E}\end{matrix} \quad \dfrac{\dfrac{[\mathcal{D}+\mathcal{E}]^{x'}}{\mathcal{D}+\varphi}\boxed{(-)} \quad \begin{matrix}\vdots\, t_3\\ \sigma\end{matrix} \quad [\mathcal{D}+\varphi+\sigma]^y}{\mathcal{D}+\varphi+\sigma}\,\boxed{(+_y)}}{\mathcal{D}+\varphi+\sigma}\,(+_{x'}) \qquad +_{x'}\,(t_1, t_2; +_y(-(x'), t_3; y))$$

When we further reduce the above new redex, we obtain the following proof without any redex (explicitly). Note that the application of $-$ in the minor premise does not constitute a redex.

$$\dfrac{\begin{matrix}\vdots\, t_1\\ \mathcal{D}\end{matrix} \quad \begin{matrix}\vdots\, t_2\\ \mathcal{E}\end{matrix} \quad \dfrac{[\mathcal{D}+\mathcal{E}]^{x'} \quad \begin{matrix}\vdots\, t_3\\ \sigma\end{matrix} \quad \dfrac{[\mathcal{D}+\mathcal{E}+\sigma]^{y'}}{\mathcal{D}+\varphi+\sigma}\,(-)}{\mathcal{D}+\varphi+\sigma}\,(+_{y'})}{\mathcal{D}+\varphi+\sigma}\,(+_{x'}) \qquad +_{x'}\,(t_1, t_2; +_{y'}(x', t_3; -(y')))$$

In general, the more exhaustive cases for an application of +, the more new redexes are provided after an application of ±-reduction. However, those new redexes are provided at the minor premises of the original application of +, that is, redexes are moved to the right upper part of the proof, and the height of proofs remains essentially unchanged by ±-reduction. Thus, we prove our normalization theorem by the following procedure: (1) we choose the rightmost-uppermost redex; (2) after an application of ±-reduction, we further reduce new redexes one by one, which are located at the right upper part. Thus, our theorem is essentially the weak normalization theorem.

Theorem 4.8 (Normalization for HLe) *Let α, α_i be a diagram or a formula. Any proof of α from $\alpha_1, \ldots, \alpha_n$ is reduced to a normal proof of α from $\alpha_1, \ldots, \alpha_n$.*

Proof. For the same reason as the proof of the theorem for basic HL, to block the generation of inessential redexes, we generalize our +-rule to a parallel rule $\oplus$, where the number of major premises is not restricted to two. Let $(\mathcal{D}+\vec{\alpha})_1, \ldots, (\mathcal{D}+\vec{\alpha})_n$ be exhaustive cases for $\mathcal{D}$ plus $\vec{\alpha}$:

$$\dfrac{\begin{matrix}\vdots\, u\\ \mathcal{D}\end{matrix} \quad \begin{matrix}\vdots\, \vec{t}\\ \alpha\end{matrix} \quad \begin{matrix}[(\mathcal{D}+\vec{\alpha})_1]^x\\ \vdots\, s_1\\ \beta\end{matrix} \quad \cdots \quad \begin{matrix}[(\mathcal{D}+\vec{\alpha})_n]^x\\ \vdots\, s_n\\ \beta\end{matrix}}{\beta}\,(\oplus_x) \qquad \oplus_x\,(u, \vec{t}; s_1, \ldots, s_n)$$

where every x in s_i is bound by $\oplus_x$.

Our $\pm$-reduction is also generalized as follows.

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{D}\end{array}\qquad \begin{array}{c}\begin{array}{c}\vdots\,\vec{t_1}\\ \dfrac{\vec{\mathcal{E}}}{\vec{\delta}}(-)\end{array}\quad\begin{array}{c}\vdots\,\vec{t_2}\\ \vec{\gamma}\end{array}\\ \vdots\,t_3\\ \alpha\end{array}\qquad \begin{array}{c}[(\mathcal{D}+\vec{\alpha})_1]^x\\ \vdots\, s_1\\ \beta\end{array}\quad \cdots \quad \begin{array}{c}[(\mathcal{D}+\vec{\alpha})_n]^x\\ \vdots\, s_n\\ \beta\end{array}}{\beta}\,(\oplus_x)$$

is reduced to the following:

$$\dfrac{\begin{array}{c}\vdots\, u\\ \mathcal{D}\end{array}\quad \begin{array}{c}\vdots\,\vec{t_1}\\ \vec{\mathcal{E}}\end{array}\quad \begin{array}{c}\vdots\,\vec{t_2}\\ \vec{\gamma}\end{array}\qquad \begin{array}{c}\dfrac{[(\mathcal{D}+\vec{\mathcal{E}}+\vec{\gamma})_1]^{x'}}{(\mathcal{D}+\vec{\alpha})_i}(-)\\ \vdots\, s_i[x:=-(x')]\\ \beta\end{array}\quad \cdots \quad \begin{array}{c}\dfrac{[(\mathcal{D}+\vec{\mathcal{E}}+\vec{\gamma})_m]^{x'}}{(\mathcal{D}+\vec{\alpha})_j}(-)\\ \vdots\, s_j[x:=-(x')]\\ \beta\end{array}}{\beta}\,(\oplus_{x'})$$

where we assume $(\mathcal{D}+\vec{\mathcal{E}}+\vec{\gamma})_1,\ldots,(\mathcal{D}+\vec{\mathcal{E}}+\vec{\gamma})_m$ be exhaustive cases for $\mathcal{D},\vec{\mathcal{E}}$, plus $\vec{\gamma}$, and $s_i,\ldots,s_j \in \{s_1,\ldots,s_n\}$. The term notation is as follows:

$$\oplus_x(u, t_3\{\overrightarrow{y^\delta := -(t_1)}, \overrightarrow{z^\gamma := t_2}\}; s_1,\ldots,s_n) \;\rhd\; \oplus_{x'}(u,\vec{t_1},\vec{t_2}; s_i[x:=-(x')],\ldots,s_j[x:=-(x')])$$

We choose the rightmost-uppermost redex, which does not contain any redex in its minor premises. Thus, such redex has the above form, where $s_1,\ldots,s_n$ does not contain any redex, although $u,\vec{t_1},\vec{t_2},\vec{t_3}$ may contain some redexes. We show that the redex is removable by the induction on the maximal number K of applications of $\oplus$ in every $s_1,\ldots,s_n$.
(1) When $K=0$, none of $s_1,\ldots,s_n$ contain any application of $\oplus$. By reducing the given redex, the original redex is reduced and no redex is generated in this part because $s_i,\ldots,s_j$ does not contain any application of $\oplus$.
(2) When $K>0$, there exists at most K applications of $\oplus$ in every $s_1,\ldots,s_n$. Assume that s_i contains K applications of $\oplus$, and s_i appears after the reduction. Although these K applications of $\oplus$ do not constitute any redex in the original proof since we choose the rightmost-uppermost redex, they may constitute numerous new redexes in $s_i[x:=-(x')]$ after the reduction. However, for every redex generated after the reduction, the number of applications of $\oplus$ in its minor premises is reduced to less than K. The same applies to proofs other than s_i. Thus, by the induction hypothesis, we can reduce case $K>0$ to case $K=0$, and we are able to reduce the rightmost-uppermost redex.

Therefore, by the induction on the number of redexes, we obtain the theorem. ∎

4.3 Structure of normal heterogeneous proofs in HLe

As is the case with $\vee E$-rule in the usual natural deduction for FOL, our extended $+$-rule in HLe makes it difficult to grasp the structure of proofs. When we examine the maximal diagram property of HLe, we find that it is difficult to compare the complexities of diagrams $\mathcal{D}+\alpha$ and $\mathcal{E}+\gamma$ in the proof of Remark 4.7. To overcome this difficulty, we consider other reduction rules related to our $+$-rule. These include simplification reduction and permutation reduction as investigated by [19, 17, 16].

We first consider the simplification reduction. When one of the minor premises does not depend on an auxiliary assumption of an application of $+$, we already have a proof without the application of $+$.

Definition 4.9 (Simplification) Simplification reduction is of the following form:

$$\dfrac{\begin{matrix}\vdots\, u\\ \mathcal{D}\end{matrix}\quad \begin{matrix}\vdots\, t\\ \alpha\end{matrix}\quad \begin{matrix}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_1\\ \beta\end{matrix}\quad \cdots \quad \begin{matrix}\vdots\, s_i\\ \beta\end{matrix}\quad\cdots\quad \begin{matrix}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_n\\ \beta\end{matrix}}{\beta}\,(+_x) \quad \rhd \quad \begin{matrix}\vdots\, s_i\\ \beta\end{matrix}$$

where s_i does not contain the term-variable x.

As indicated by [17, 16], simplification reduction may provide new $\pm$-redex. Thus, we need to apply simplification reduction first before applications of $\pm$-reduction.

Next, we investigate the permutation reduction. We extend the notion of major premises to apply FOL introduction rules and $-$-rule, too.

Definition 4.10 (Permutation) **Major premises** of introduction rules of FOL, as well as $-$-rule, are premises of those rules.

Let R be an inference rule. When a major premise of R is obtained by $+$-rule, we apply the following permutation reduction:

- The case where R is not $+$-rule is as follows:

$$\dfrac{\dfrac{\begin{matrix}\vdots\, u\\ \mathcal{D}\end{matrix}\quad \begin{matrix}\vdots\, t\\ \alpha\end{matrix}\quad \begin{matrix}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_1\\ \beta\end{matrix}\quad \cdots \quad \begin{matrix}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_n\\ \beta\end{matrix}}{\beta}\,(+_x)}{\gamma}\,R \quad \rhd \quad \dfrac{\begin{matrix}\vdots\, u\\ \mathcal{D}\end{matrix}\quad \begin{matrix}\vdots\, t\\ \alpha\end{matrix}\quad \begin{matrix}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_1\\ \dfrac{\beta}{\gamma}\,R\end{matrix}\quad \cdots \quad \begin{matrix}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_n\\ \dfrac{\beta}{\gamma}\,R\end{matrix}}{\gamma}\,(+_x)$$

- The case where R is $+$-rule ($+_y$ below) is as follows:

$$\dfrac{\begin{matrix}\vdots\, u_1\\ \mathcal{E}\end{matrix}\quad \dfrac{\begin{matrix}\vdots\, u_2\\ \mathcal{D}\end{matrix}\quad \begin{matrix}\vdots\, t\\ \alpha\end{matrix}\quad \begin{matrix}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_{11}\\ \beta\end{matrix}\quad \cdots \quad \begin{matrix}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_{1n}\\ \beta\end{matrix}}{\beta}\,(+_x)\quad \begin{matrix}[(\mathcal{E}+\beta)]^y\\ \vdots\, \vec{s_2}\\ \gamma\end{matrix}}{\gamma}\,(+_y)$$

is reduced to the following:

$$\dfrac{\begin{matrix}\vdots\, u_2\\ \mathcal{D}\end{matrix}\quad \begin{matrix}\vdots\, t\\ \alpha\end{matrix}\quad \dfrac{\begin{matrix}\vdots\, u_1\\ \mathcal{E}\end{matrix}\quad \begin{matrix}[(\mathcal{D}+\alpha)_1]^x\\ \vdots\, s_{11}\\ \beta\end{matrix}\quad \begin{matrix}[(\mathcal{E}+\beta)]^y\\ \vdots\, \vec{s_2}\\ \gamma\end{matrix}}{\gamma}\,(+_y)\quad \cdots \quad \dfrac{\begin{matrix}\vdots\, u_1\\ \mathcal{E}\end{matrix}\quad \begin{matrix}[(\mathcal{D}+\alpha)_n]^x\\ \vdots\, s_{1n}\\ \beta\end{matrix}\quad \begin{matrix}[(\mathcal{E}+\beta)]^y\\ \vdots\, \vec{s_2}\\ \gamma\end{matrix}}{\gamma}\,(+_y)}{\gamma}\,(+_x)$$

Observe that permutation reduction moves an application of $+$ in a major premise to the positions of minor premises.

Example 4.11 Permutation reduction is applied as follows:

$$
\dfrac{
\dfrac{\begin{array}{c}\vdots\, u_1\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, t_1\\ \alpha\end{array} \quad \begin{array}{c}[\mathcal{D}+\alpha]^x\\ \vdots\, s_1\\ \beta\end{array}}{\beta}\,(+_x)
\quad
\dfrac{\begin{array}{c}\vdots\, u_2\\ \mathcal{E}\end{array} \quad \begin{array}{c}\vdots\, t_2\\ \alpha'\end{array} \quad \begin{array}{c}[\mathcal{E}+\alpha']^y\\ \vdots\, s_2\\ \gamma\end{array}}{\gamma}\,(+_y)
\quad
\begin{array}{c}[\beta+\gamma]^z\\ \vdots\, s_3\\ \delta\end{array}
}{\delta}\,(+_z)
$$

is reduced to the following:

$$
\dfrac{
\begin{array}{c}\vdots\, u_1\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, t_1\\ \alpha\end{array} \quad
\dfrac{
\begin{array}{c}[\mathcal{D}+\alpha]^x\\ \vdots\, s_1\\ \beta\end{array}
\quad
\dfrac{\begin{array}{c}\vdots\, u_2\\ \mathcal{E}\end{array} \quad \begin{array}{c}\vdots\, t_2\\ \alpha'\end{array} \quad \begin{array}{c}[\mathcal{E}+\alpha']^y\\ \vdots\, s_2\\ \gamma\end{array}}{\gamma}\,(+_y)
\quad
\begin{array}{c}[\beta+\gamma]^z\\ \vdots\, s_3\\ \delta\end{array}
}{\delta}\,(+_z)
}{\delta}\,(+_x)
$$

The above proof is further reduced to the following:

$$
\dfrac{
\begin{array}{c}\vdots\, u_1\\ \mathcal{D}\end{array} \quad \begin{array}{c}\vdots\, t_1\\ \alpha\end{array} \quad
\dfrac{
\begin{array}{c}\vdots\, u_2\\ \mathcal{E}\end{array} \quad \begin{array}{c}\vdots\, t_2\\ \alpha'\end{array} \quad
\dfrac{
\begin{array}{c}[\mathcal{D}+\alpha]^x\\ \vdots\, s_1\\ \beta\end{array}
\quad
\begin{array}{c}[\mathcal{E}+\alpha']^y\\ \vdots\, s_2\\ \gamma\end{array}
\quad
\begin{array}{c}[\beta+\gamma]^z\\ \vdots\, s_3\\ \delta\end{array}
}{\delta}\,(+_z)
}{\delta}\,(+_y)
}{\delta}\,(+_x)
$$

An application of permutation reduction always reduces the number of applications of + in a major premise in a given proof. Therefore, by successively applying permutation reduction, we finally obtain, in finite steps, a proof without applications of + in any major premise.

Permutation reduction may provide new $\pm$-redexes. Let us consider the proofs in Remark 4.7, where the second proof is obtained by an application of permutation reduction to the first. Although the first proof does not contain $\pm$-redex, the second proof contains a $\pm$-redex. Thus, we need to apply permutation reduction, as well as simplification reduction, in advance of $\pm$-reduction.

Observe that in Definition 4.4, after simplification and permutation reductions, there are no applications of +-rule in $u, \vec{t_1}, \vec{t_2}$ both in the proofs before and after application of $\pm$-reduction. Thus, $\pm$-reduction does not provide any permutation redex. Therefore, by applying simplification and permutation reductions followed by $\pm$-reduction, we obtain intended normal proofs. In what follows, by normal proofs, we mean proofs which do not contain any $\pm$-redex, simplification redex, nor permutation redex.

Let us now investigate the structure of normal proofs. Following [30], we extend the notion of tracks to our HLe. For simplicity, we define it only for normal proofs.

Definition 4.12 A **track** of a normal proof t in HLe is a sequence of formulas and diagrams $\alpha_1, \ldots, \alpha_n$ such that

1. α_1 is an open or closed assumption not closed by +-rule.
2. α_n is either an app-formula or the conclusion of t.
3. α_i $(1 \leq i \leq n-1)$ is not an app-formula, and either:

(a) α_i is not the major premise of $+$-rule and α_{i+1} is directly below α_i; or

(b) α_i is the major premise of $+$-rule and α_{i+1} is an auxiliary assumption of the $+$-rule.

In the same way as Lemma 3.12, Proposition 3.13, and Corollary 3.14 for basic HL, we obtain the following lemma, proposition, and the maximal diagram property for HLe.

Lemma 4.13 (Track) *In every track in a normal proof in* HLe*, no application of* $-$*-rule can precede an application of* $+$*-rule.*

Proposition 4.14 (Structure of normal form) *In* HLe*, every track in a normal proof of* α *has the form (1)* $\varphi_1, \dots, \varphi_k$ *(when the track starts from a formula), or (2)* $\mathcal{D}_1, \dots, \mathcal{D}_l$*,* $\mathcal{E}_1, \dots, \mathcal{E}_m, \psi_1, \dots, \psi_n, \alpha, \dots, \alpha$ *(when the track starts from a diagram), where*

- $\varphi_2, \dots, \varphi_k$ *and* $\psi_2, \dots, \psi_n$ *are obtained by applications of FOL-rules;*
- $\mathcal{D}_1, \dots, \mathcal{D}_{l-1}$ *are major premises of applications of* $+$*-rule;*
- $\mathcal{D}_l$ *is one of the auxiliary assumptions of the last application of* $+$*-rule;*
- $\mathcal{E}_1, \dots, \mathcal{E}_m, \psi_1$ *are obtained by applying* $-$*-rule;*
- $\alpha, \dots, \alpha$ *are conclusions of applications of* $+$*-rule.*

Thus, every normal proof has the following form:

$$\dfrac{\mathcal{D}_1 \quad \begin{matrix}\varphi_1\\ \vdots\\ \varphi_k\end{matrix} \quad \dfrac{[\mathcal{D}_2] \quad \begin{matrix}\vdots\\ \beta_2\end{matrix} \quad \begin{matrix} \dfrac{[\mathcal{D}_{l-1}] \quad \begin{matrix}\vdots\\ \beta_{l-1}\end{matrix} \quad \begin{matrix}\dfrac{[\mathcal{D}_l]}{\mathcal{E}_1}\,\mathsf{del}/*E \\ \vdots\ \mathsf{del} \\ \dfrac{\mathcal{E}_m}{\psi_1}\,\mathsf{obs} \\ \vdots\ \text{FOL} \\ \alpha\end{matrix}}{\alpha}\,(+) \\ \vdots\ (+) \\ \alpha \end{matrix}}{\alpha}\,(+)}{\alpha}\,(+)$$

Corollary 4.15 (Maximal diagrams in HLe) *Let* t *be a track* $\mathcal{D}_1, [\mathcal{D}_2], \dots, [\mathcal{D}_l], \mathcal{E}_1, \dots, \mathcal{E}_m$*,* $\psi_1, \dots, \psi_n, \alpha, \dots, \alpha$ *in a normal proof. Then there exists a diagram* $\mathcal{D}$ *in* t *such that* $\mathcal{D}_i, \mathcal{E}_j \sqsubseteq \mathcal{D}$ *for* $1 \leq i \leq l, 1 \leq j \leq m$*, or there exists a sequence of* $*$ *in* t*.*

Proof. We show that $[\mathcal{D}_l]$ is the maximal diagram in t. When $\mathcal{D}_i \neq *$ for $1 \leq i \leq l$, let us consider the following form of an application of $+$:

$$\dfrac{\mathcal{D}' \quad \begin{matrix}\vdots\ t_1\\ \alpha\end{matrix} \quad \begin{matrix}[(\mathcal{D}'+\alpha)_1]^x\\ \vdots\ s_1\\ \beta\end{matrix} \quad \cdots \quad \begin{matrix}[(\mathcal{D}'+\alpha)_o]^x\\ \vdots\ s_o\\ \beta\end{matrix}}{\beta}\,(+_x)$$

If there is another application of $+$ in the above s_k, due to the form of normal proofs (cf. Proposition 4.14), $[(\mathcal{D}' + \alpha)_k]^x$ should be a major premise of an application of $+$ in s_k. In general, every major premise of an application of $+$ is a subdiagram of every exhaustive case of the application of $+$. Hence, $[(\mathcal{D}' + \alpha)_k]^x$ should be a subdiagram of another application of $+$, and every exhaustive case of the topmost application of $+$, i.e., $[\mathcal{D}_l]$ in the above sequence, is the maximal diagram in t. ∎

5 Concluding remarks

Let us summarize our results and discuss relationships and certain dualities between sentential reasoning and diagrammatic reasoning.

The basis of our study is the Postulate 2.1. From the perspective of FOL, we regard our diagram as the deductively closed conjunction of diagrammatic formulas. Based on this observation, we have defined inference rules of HL. Natural deduction inference rules of FOL consist of pairs of introduction and elimination rules for every connective. Inference rules of HL consist of +-rule (app and uni) and −-rule (obs, del, and $*E$). Here, +-rule (i.e., construction of a diagram) can be considered to be a generalization of the conjunction introduction rule of FOL, and −-rule (extraction from a diagram) can be considered to be a generalization of the conjunction elimination rule. (However, this correspondence is more subtle in our extended HLe, where +-rule has the form of the general elimination rule.)

By examining what is a *detour* in heterogeneous proofs given the Postulate 2.1, we have defined our ±-reduction rule for heterogeneous proofs. In the usual natural deduction for FOL, reduction rules are defined for every *introduction-elimination* pair of connectives. In contrast, our ±-reduction rule for heterogeneous proofs is defined for the *elimination-introduction* pair by regarding −-rule as an elimination rule and +-rule as an introduction rule.

In terms of the reduction rule, we have proved the normalization theorems for basic HL (Theorem 3.10) and for extended HLe (Theorem 4.8). In FOL, one of the most important consequences obtained from the normalization theorem is the subformula property. We have investigated a diagrammatic counterpart of the subformula property, and we have shown the maximal diagram property for HL (Corollary 3.14) and for HLe (Corollary 4.15). Thus, in contrast to the reduction in FOL, which reduces a maximal formula, the reduction in heterogeneous logic constructs a maximal diagram.

Based on the normalization theorems, we have further investigated the structure of normal proofs in HL (Proposition 3.13) and in HLe (Proposition 4.14). In FOL, every normal proof consists of, from the top-down, applications of elimination rules followed by applications of introduction rules. Thus, sentential reasoning, as characterized by [19, 18], consists of (1) decomposition of given assumptions, and (2) construction of a conclusion by combining the decomposed formulas. In contrast, every normal proof in HL consists of applications of +-rule (the introduction rule) followed by applications of −-rule (the elimination rule). Thus, diagrammatic reasoning consists of (1) construction of a maximal diagram by unifying pieces of information contained in given assumptions, and (2) extraction of a conclusion from the unified diagram. This observation is also supported by various cognitive studies on diagrammatic reasoning, for example, [1, 24].

Gentzen [8] remarked that an introduction rule gives the meaning of the logical connective in question, and this meaning justifies the corresponding elimination rule. This idea is developed by Prawitz to the notion of the proof-theoretic validity [19], and further developed to proof-theoretic semantics, cf. [22, 15]. In our diagrammatic reasoning, +-rule is a geometric operation such as adding diagrammatic objects to a part of the given diagram, and it does not commit in the meaning of diagrams directly. On the other hand, −-rule is an operation to extract meaning from the given diagram. Thus, −-rule, i.e., the elimination rule, seems to be principal in terms of proof-theoretic semantics. This is also in contrast to sentential reasoning, where introduction rules are usually considered to be principal.

In this way, there are certain dualities between sentential reasoning and diagrammatic reasoning. Their relation has also been extensively discussed in the literature on cognitive science, in particular, as to which of rule-based sentential reasoning and model-based reasoning is fundamental to, or is close to, our actual reasoning. While rule-based sentential reasoning is considered to be conducted by constructing a kind of natural deduction proofs for FOL (e.g., [20]), model-based reasoning is considered to be conducted by constructing a mental model (e.g., [11]), which was derived from the notion of models in logic. A mental model is a restricted finite model in logic, and it is constructed as long as it is conceivable for a reasoner depending on his/her state of knowledge. Such a mental model can be considered to be represented by a diagram, and hence, reasoning with mental models can be considered as reasoning with certain diagrams. Thus, our investigation on heterogeneous logic combining sentences and diagrams may shed light on the discussion from the perspective of proof theory.

Based on the above findings on the duality between sentential and diagrammatic reasoning, we need further philosophical investigation into this relationship.

References

[1] Gerard Allwein and Jon Barwise, eds., *Logical reasoning with diagrams*, Oxford Studies In Logic And Computation Series, 1996.

[2] Dave Barker-Plummer, Jon Barwise, and John Etchemendy, *Logical Reasoning With Diagrams & Sentences: Using Hyperproof*, CSLI Publications, 2017.

[3] Dave Barker-Plummer and John Etchemendy, A computational architecture for heterogeneous reasoning, *Journal of Experimental & Theoretical Artificial Intelligence*, 19(3), 195-225, 2007.

[4] Dave Barker-Plummer and Nik Swoboda, Reasoning with coincidence grids–A sequent-based logic and an analysis of complexity, *Journal of Visual Languages and Computing*, 22 (1), 56-65, 2011.

[5] Jon Barwise and John Etchemendy, *Hyperproof: For Macintosh*, The Center for the Study of Language and Information Publications, 1995.

[6] Dirk van Dalen, *Logic and Structure*, Springer, 4th ed. 2004.

[7] Roy Dyckhoff, Implementing a simple proof assistant, *Proceedings of the Workshop on Programming for Logic Teaching*, Leeds Centre for Theoretical Computer Science Proceedings 23.88, 49-59, 1988.

[8] Gerhard Gentzen, Untersuchungen über das logische Schließen, *Mathematische Zeitschrift*, 39, 176-210, 405-431, 1934. English Translation: Investigations into logical deduction, in M. E. Szabo, ed., *The collected Papers of Gerhard Gentzen*, 1969.

[9] Eric Hammer, Reasoning with Sentences and Diagrams, *Notre Dame Journal of Formal Logic*, Volume 35, Number 1, 73-87, 1994.

[10] John Howse, Gem Stapleton, and John Taylor, Spider Diagrams, *LMS Journal of Computation and Mathematics*, Volume 8, 145-194, London Mathematical Society, 2005.

[11] Philip N Johnson-Laird, *Mental Models: Toward a Cognitive Science of Language, Inference and Consciousness*, Harvard University Press, 1983.

[12] López-Escobar, Standardizing the N systems of Gentzen, *Models, algebras, and proofs*, Caicedo, X. and Montenegro, C., editors, Lecture Notes in Pure and Applied Mathematics, Vol. 203, 411-434, 1999.

[13] Koji Mineshima, Mitsuhiro Okada, and Ryo Takemura, A Diagrammatic Inference System with Euler Circles, *Journal of Logic, Language and Information*, 21, 3, 365-391, 2012.

[14] Koji Mineshima, Mitsuhiro Okada, and Ryo Takemura, Two Types of Diagrammatic Inference Systems: Natural Deduction Style and Resolution Style, *Diagrammatic Representation and Inference: 6th International Conference, Diagrams 2010*, Lecture Notes In Artificial Intelligence, Springer, 99-114, 2010.

[15] Thomas Piecha and Peter Schroeder-Heister (Eds.), *Advances in Proof-Theoretic Semantics*, Springer, 2016.

[16] Sara Negri and Jan von Plato, *Structural Proof Theory*, Cambridge University Press, 2001.

[17] Jan von Plato, Natural deduction with general elimination rules, *Archive for Mathematical Logic*, Volume 40, Issue 7, 541-567, 2001.

[18] Dag Prawitz, *Natural Deduction*, Almqvist & Wiksell, 1965 (Dover, 2006).

[19] Dag Prawitz, Ideas and results in proof theory, *Proceedings 2nd Scandinavian Logic Symposium*, 237-309, 1971.

[20] Lance J. Rips, *The psychology of proof: deductive reasoning in human thinking*, MIT Press Cambridge, MA, USA, 1994.

[21] Peter Schroeder-Heister, A Natural Extension of Natural Deduction, *Journal of Symbolic Logic*, Volume 49, Issue 4, 1284-1300, 1984.

[22] Peter Schroeder-Heister, Validity concepts in proof-theoretic semantics, *Synthese*, 148(3), 525-571, 2006.

[23] Sun-Joo Shin, *The Logical Status of Diagrams*, Cambridge University Press, 1994.

[24] Atsushi Shimojima, *Semantic Properties of Diagrams and Their Cognitive Potentials*, CSLI Publications, Stanford, CA, 2015.

[25] Nik Swoboda and Gerard Allwein, Heterogeneous Reasoning with Euler/Venn Diagrams Containing Named Constants and FOL, *Electronic Notes in Theoretical Computer Science*, 134, 153-187, 2005.

[26] Ryo Takemura, Proof theory for reasoning with Euler diagrams: a logic translation and normalization, *Studia Logica*, 101, 1, 157-191, 2013.

[27] Ryo Takemura, Towards a proof theory for heterogeneous logic combining sentences and diagrams, *Diagrammatic Representation and Inference*, Lecture Notes in Computer Science, Springer, 607-623, 2018.

[28] Ryo Takemura, Atsushi Shimojima, Yasuhiro Katagiri, Logical investigation of reasoning with tables, *Diagrammatic Representation and Inference*, Lecture Notes in Computer Science, Springer, Volume 8578, 261-276, 2014.

[29] Neil Tennant, *Autologic*, Edinburgh University Press, 1992,

[30] Anne S. Troelstra and Helmut Schwichtenberg, *Basic Proof Theory*, 2nd edition, Cambridge University Press, 2000.

[31] Matej Urbas and Mateja Jamnik, Heterogeneous Proofs: Spider Diagrams Meet Higher-Order Provers, *Interactive Theorem Proving*, LNCS, Volume 6898, 376-382, 2011.