Towards a proof theory for heterogeneous logic combining sentences and diagrams

Ryo Takemura

Nihon University, Japan.
takemura.ryo@nihon-u.ac.jp

February 12, 2018

Abstract

We attempt to develop a proof theory for heterogeneous logic combining first-order formulas and diagrams. In proof theory, normal proofs and normalization play a central role, which makes it possible to analyze and characterize the structure of proofs in a given system. Based on the difference between linguistic reasoning and diagrammatic reasoning, we investigate a proof theory particular to diagrammatic reasoning, as distinguished from linguistic reasoning.

Contents

1 Introduction

Heterogeneous reasoning combining various graphical/diagrammatic and sentential/linguistic representations has been an important subject in the study of diagrammatic reasoning, and several heterogeneous systems have been investigated so far. Blocks world systems [2, 5]; Euler and Venn systems [9, 17, 21]; and correspondence table systems [4, 20] are some examples of such studies. However, proof theory of heterogeneous logic has not yet been developed much.

Existing development in proof theory has taken place by investigating logical proofs based on sentential/linguistic representation. One of the major goals of proof theory is to analyze and characterize the structure of proofs, and thereby investigate effective strategies to construct/search proofs in a system. Thus, proof theory offers a basis in theorem proving. In such proof theory, normal proofs and normalization play a central role. Thanks to the normalization theorem, any proof is reduced to a normal form, and we are able to focus on normal proofs for their analysis and characterization. Therefore, Gentzen [8] called the normalization theorem as Hauptsatz (Main Theorem) of proof theory.

If we translate diagrams into formulas of first-order logic (FOL), we can apply the usual proof-theoretic techniques to heterogeneous/diagrammatic logic in a straightforward manner. Based on this idea, the author in [18] investigated a class of Euler diagrammatic proofs called "N-normal diagrammatic proofs" that has a one-to-one correspondence with normal proofs in FOL. Although N-normal diagrammatic proofs have the structure of linguistic FOL proofs, they do not reflect characteristics of diagrammatic proofs. Thus, by using such an N-normal form, it is difficult to characterize the structure of diagrammatic proofs in a general sense.

There is a major difference between linguistic reasoning and diagrammatic reasoning with respect to their methods and strategies. Linguistic reasoning, as characterized by the normalization theorem (cf. [14, 13]) for FOL, consists of (1) decomposition of given premises, and (2) construction of a conclusion by combining the decomposed formulas. In contrast, diagrammatic reasoning consists of (1) construction of a (maximal) diagram by unifying pieces of information contained in given premises, and (2) extraction of a conclusion from the unified diagram (see, for example, [16]).

In light of such a distinction, we investigate proof theory particular to diagrammatic reasoning, as distinguished from linguistic reasoning. We study heterogeneous logic combining first-order formulas and diagrams within the framework of natural deduction. We investigate abstract properties of heterogeneous proofs independent of particular systems. Such properties are shared by various concrete systems such as Euler and Venn systems (e.g., [9, 15, 10, 12]); blocks world systems ([2, 5]); and correspondence table systems (e.g., [4, 20]). In Section 2, we describe our abstract syntax of heterogeneous logic. In Section 3, we introduce our inference rules. Rules for formulas are the usual natural deduction rules for FOL. We investigate, among various inference rules, heterogeneous rules Apply and Observe (cf. [2, 5, 9]), as well as diagrammatic rules Unification and Deletion (cf. [15, 10, 12]) exclusively, since these rules are considered to be the most basic rules and are shared by various heterogeneous systems. In Section 4, we investigate a normalization theorem in our heterogeneous system, and provide a characterization for the structure of our heterogeneous proofs.

2 Syntax of heterogeneous logic

We introduce syntax of heterogeneous logic abstractly. While concrete syntax is defined in each system, here we extract common items to be specified in each system.

The syntax of heterogeneous logic is defined by specifying the following *formulas, diagrams, diagrammatic objects*, and *diagrammatic formulas*:

Formulas: denoted by $\varphi, \psi, \sigma, \varphi_1, \varphi_2, \ldots$. Formulas of FOL (first-order logic) are defined inductively as usual:

$$\varphi ::= A(t_1, \ldots, t_n) \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi \mid \neg\varphi \mid \bot \mid \forall x\varphi \mid \exists x\varphi$$

where $A(t_1, \ldots, t_n)$ is an atomic formula consisting of a predicate A and terms $t_1, \ldots, t_n$. When A is a unary predicate, we usually omit parentheses and write an atomic formula such as At.

Diagrams: denoted by $\mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{D}_1, \mathcal{D}_2, \ldots$. The answer to what qualifies as a concrete diagram depends on each system. Cf. Example 2.2 below.

Diagrammatic objects: denoted by $o, o_1, o_2, \ldots$. They are components of diagrams, and the answer to what qualifies as a concrete diagrammatic object depends on each system. By $ob(\mathcal{D})$, we denote the set of diagrammatic objects that appear on a diagram $\mathcal{D}$.

For example, diagrammatic objects are named circles and points, linking between points, shading in Euler and Venn systems; and blocks such as cubes and tetrahedron in blocks world systems; $\bigcirc$ and $\times$ in correspondence table systems. (Diagrammatic objects can be added to or deleted from a diagram by inference rules.)

Diagrammatic formulas: denoted by $\varphi^d, \psi^d, \sigma^d, \varphi_1^d, \varphi_2^d \ldots$. They describe pieces of basic information contained in diagrams. The answer to what kind of formulas qualify as diagrammatic formulas is specified in each system.

For example, in a Venn diagrammatic system, $\neg\exists x(A_1x \wedge \cdots \wedge A_nx \wedge \neg B_1x \wedge \cdots \wedge \neg B_mx)$ is a diagrammatic formula describing that the region inside circles $A_1, \ldots, A_n$ and outside $B_1, \ldots, B_m$ is shaded. See Example 2.2 for an example of an Euler system. In a blocks world system, atomic formulas of hyperproof such as $small(t), cube(t), leftof(t_1, t_2)$, and so on are diagrammatic formulas. In a correspondence table system of [20], $r(c)$ and $\neg r(c)$ are diagrammatic formulas, which describe that there exists $\bigcirc$ in the cell (r, c) and $\times$ in (r, c), respectively.

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. We further identify a relation holding on a diagram with a diagrammatic formula that describes the relation.

For a diagram $\mathcal{D}$, $type(\mathcal{D})$ is the set $\{\varphi_1^d, \varphi_2^d, \ldots, \varphi_n^d\}$ of diagrammatic formulas, such that the relation φ_i^d holds on $\mathcal{D}$ if and only if $\varphi_i^d \in type(\mathcal{D})$. We identify $type(\mathcal{D})$ with the conjunctive formula $\varphi_1^d \wedge \varphi_2^d \wedge \cdots \wedge \varphi_n^d$. The *type* of a diagram is the symbolic specification of the diagram.

Based on the above specification, we elucidate our **postulates** in this article.

Postulate 2.1

1. We presume the set of diagrammatic formulas to be a subset of FOL formulas. Thus, every piece of basic information contained in the diagrams is described by a formula of FOL.
2. We regard a diagram, in view of FOL, as the conjunction of diagrammatic formulas comprising the diagram. Thus, we do not consider linking between diagrams in this article, since linking between diagrams makes its type disjunctive. (Although we allow linking between points.)
3. We presume $type(\mathcal{D})$ is *deductively closed* with respect to diagrammatic formulas. That is, when $type(\mathcal{D}) = \{\varphi_1^d, \varphi_2^d, \ldots, \varphi_n^d\}$, if $\varphi_1^d \wedge \varphi_2^d \wedge \cdots \wedge \varphi_n^d$ implies a diagrammatic formula ψ^d such that $ob(\psi^d) \subseteq ob(\mathcal{D})$, then $\psi^d \in type(\mathcal{D})$. The above "implies" is considered as an appropriate semantic consequence or syntactic consequence in FOL. (The deductive closedness of diagrams corresponds to the notion of *free ride* of Shimojima [16].)

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

Although we mainly illustrate only one concrete Euler diagrammatic system of [12] below because of space limitation, our definition is valid for other Euler and Venn systems (e.g., [9, 15, 10]); blocks world systems ([2, 5]); correspondence table systems (e.g., [4, 20]), and so on. [1]

Example 2.2 (EUL-diagrams) An Euler diagram of [12], called an EUL-diagram, is defined as a plane with named circles and points. Each EUL-diagram is specified by inclusion and exclusion relations maintained between circles and points on the diagram. EUL-diagrams can express neither disjunctive information with respect to the location of a point, nor information of contradiction.

Diagrams. An EUL-diagram is a plane with a finite number of (named) simple closed curves (simply called (named) circles and denoted by $A, B, C, \dots$), constant points (denoted by $a, b, c, \dots$), and existential points (denoted by $x, y, z, \dots$). Constant points and existential points are collectively called (named) points, and are denoted by $t, s, t_1, t_2, \dots$.

Diagrammatic objects are named circles and points.

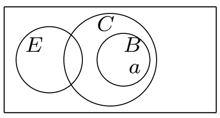
Diagrammatic formulas. An EUL-diagram is specified in terms of the following topological relations between diagrammatic objects.

$A \sqsubset B$	"the interior of A is *inside of* the interior of B,"
$A \vdash\!\dashv B$	"the interior of A is *outside of* the interior of B,"
$A \bowtie B$	"there is at least one *crossing* point between A and B,"
$t \sqsubset A$	"t is *inside of* the interior of A,"
$t \vdash\!\dashv A$	"t is *outside of* the interior of A,"

These relations are expressed by the following diagrammatic formulas:

- $\forall x(Ax \to Bx)$ for $A \sqsubset B$ (A is inside of B);
- $\forall x(Ax \to \neg Bx)$ for $A \vdash\!\dashv B$ (A is outside of B);
- $\forall x(Ax \to Ax) \land \forall x(Bx \to Bx)$ for $A \bowtie B$ (there is at least one crossing point between A and B);
- At for $t \sqsubset A$ (t is inside of A);
- $\neg At$ for $t \vdash\!\dashv A$ (t is outside of A).

For example, the following diagram consists of EUL-relations $B \sqsubset C, B \vdash\!\dashv E, C \bowtie E, a \sqsubset B, a \sqsubset C, a \vdash\!\dashv E$, and hence, its type is $\{\forall x(Bx \to Cx), \forall x(Bx \to \neg Ex), \forall x(Cx \to Cx) \land \forall x(Ex \to Ex), Ba, Ca, \neg Ea\}$.



[1] Strictly speaking, the correspondence table system of [19, 20] is formalized in the framework of sorted logic, and hence, we need to extend slightly the FOL framework in this article to include the system.

3 Inference rules of heterogeneous logic

We first review the usual inference rules of natural deduction for FOL in Section 3.1. Then, in Section 3.2, we introduce our heterogeneous inference rules Apply and Observe, as well as purely diagrammatic inference rules Unification and Deletion, which are shared in typical heterogeneous systems.

3.1 Natural deduction rules for FOL

A proof in natural deduction is structured as a tree consisting of formulas as its nodes and the following inference rules as its edges. The top formulas of the tree are the assumptions, and the other formulas of the tree are ones follow from the formulas immediately above, using one of the rules. A formula A in the tree is said to depend on the assumptions standing above A that have not been closed by some inference preceding A. In the following rules, a formula written within square brackets indicates that, the assumptions of this form occurring above the premises are closed at the inference. See [7, 14] for a detailed introduction to natural deduction.

Definition 3.1 (Rules for FOL) The natural deduction rules for FOL consist of the following dual pairs, each pair consisting of an *introduction (I)* and an *elimination (E)* rules, for each connective $\wedge, \vee, \to, \neg, \forall, \exists$, as well as $\bot E$ and RAA:

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

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

$$\dfrac{\begin{matrix}\vdots\\ \varphi(x)\end{matrix}}{\forall x\varphi(x)}\forall I \quad \dfrac{\begin{matrix}\vdots\\ \forall x\varphi(x)\end{matrix}}{\varphi(t)}\forall E \quad \dfrac{\begin{matrix}\vdots\\ \varphi(t)\end{matrix}}{\exists x\varphi(x)}\exists I \quad \dfrac{\begin{matrix} & [\varphi(x)]^n\\ \vdots & \vdots\\ \exists x\varphi(x) & \psi\end{matrix}}{\psi}\exists E, n \quad \dfrac{\begin{matrix}\vdots\\ \bot\end{matrix}}{\varphi}\bot E \quad \dfrac{\begin{matrix}[\neg\varphi]^n\\ \vdots\\ \bot\end{matrix}}{\varphi}RAA, n$$

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

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, and the set without RAA nor $\bot E$ provides that of minimal logic.

3.2 Heterogeneous rules

Because we regard each diagram as a conjunction of diagrammatic formulas, a diagrammatic inference rule, if it is sound, can be considered, through a translation of diagrams into FOL formulas, as a certain combination of the usual natural deduction rules. However, in contrast to a linguistic FOL rule, whose conclusion is always well-defined given well-defined premises, a diagrammatic rule's conclusion may not be defined even if the premises are well-defined, because of the expressive limitations of diagrams.

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 3.2 Heterogeneous rules of Apply (app) and Observe (obs), and diagrammatic rules of Unification (uni) and Deletion (del) have the following forms:

$$\frac{\begin{matrix}\vdots & \vdots \\ \mathcal{D} & \varphi^d\end{matrix}}{\mathcal{D}+\varphi^d}\ \mathsf{app} \qquad \frac{\begin{matrix}\vdots \\ \mathcal{D}\end{matrix}}{\psi^d}\ \mathsf{obs} \qquad \frac{\begin{matrix}\vdots & \vdots \\ \mathcal{D} & \mathcal{E}\end{matrix}}{\mathcal{D}+\mathcal{E}}\ \mathsf{uni} \qquad \frac{\begin{matrix}\vdots \\ \mathcal{D}\end{matrix}}{\mathcal{D}-\{o_1,\ldots,o_n\}}\ \mathsf{del}$$

where $\psi^d \in type(\mathcal{D})$ in obs, and $o_1, \ldots, o_n \in ob(\mathcal{D})$ in del.
These rules are applicable when every $\mathcal{D}+\varphi^d, \mathcal{D}+\mathcal{E}, \mathcal{D}-\{o_1,\ldots,o_n\}$ is a well-defined diagram.

- In del, $\mathcal{D} - \{o_1, \ldots, o_n\}$ is the diagram obtained by deleting diagrammatic objects $o_1, \ldots, o_n$ from $\mathcal{D}$.

  From a viewpoint of the type of diagrams, all of the diagrammatic formulas associated with the deleted objects $o_1, \ldots, o_n$ are eliminated.
- In app, $\mathcal{D} + \varphi^d$ is the diagram that extends from $\mathcal{D}$ by adding the information of φ^d (cf. [2, 5]).
- In uni, $\mathcal{D} + \mathcal{E}$ is the unified diagram of $\mathcal{D}$ and $\mathcal{E}$ (cf. [10, 12]).
- In obs, as well as app, an observed formula (resp. applied formula) is restricted to be a diagrammatic formula, that is the consequence of our postulate regarding a diagram as a conjunction of diagrammatic formulas.
- Depending on the specific definition of diagrams in each system, $\mathcal{D} + \varphi^d$ and $\mathcal{D} + \mathcal{E}$ are not always defined. There may be various constraints on app and uni in order to avoid the case where its conclusion is undefined. Two of the major constraints are that for *indeterminacy* and for *contradiction* as seen in [12]. In this article, we presume that app as well as uni are applicable when $\mathcal{D} + \varphi^d$ (resp. $\mathcal{D} + \mathcal{E}$) is defined as a single diagram. This allows us to exclude the case where several distinguishable diagrams or linking of them is needed (as [15, 10]) for representing $\mathcal{D} + \varphi^d$ (resp. $\mathcal{D} + \mathcal{E}$). We also do not take the rule of Cases Exhaustive [2, 5] into consideration in this article. (There may be other types of constraints mainly from a cognitive viewpoint. For example, in the Euler system of [12], uni is restricted for application to particular forms of diagrams.) Examples of the systems where we always apply app and uni without any constraint are a Venn system without points, and an Euler system with shading and without points.

Example 3.3 (EUL-diagrams) uni and app in an Euler diagrammatic system of [12] has two constraints. One is the *constraint for determinacy*, which blocks disjunctive ambiguity with respect to the location of points, and the other is the *constraint for consistency*, which blocks representing inconsistent information in a single diagram. uni and app can only be applied when these constraints are satisfied. The following is one of the app-rules. (Although the system of [12] has only diagrammatic rules of uni and del, by regarding atomic diagrams as corresponding FOL formulas, it is easily described as heterogeneous rules. See [12] for the set of complete inference rules.)

Premises: A formula of the form $\forall x(Ax \to \neg Bx)$; and a diagram $\mathcal{D}$ such that B is a circle of $\mathcal{D}$ (but A is not).

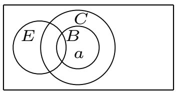
Constraint for determinacy: $t \sqsubset B$ holds for all point t of $\mathcal{D}$.

Operation: Add the circle A to $\mathcal{D}$ (with preservation of all relations on $\mathcal{D}$) so that the following conditions are satisfied on $\mathcal{D} + A \vDash B$: (1) $A \vDash B$ holds; (2) $A \bowtie X$ holds for all circles X ($\not\equiv B$) such that $B \sqsubset X$ or $B \vDash X$ or $B \bowtie X$ holds on $\mathcal{D}$.

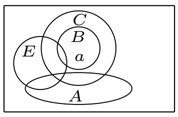
The set of relations $\mathsf{rel}(\mathcal{D} + A \vDash B)$ of the applied diagram $\mathcal{D} + A \vDash B$ is specified as follows:

$$\mathsf{rel}(\mathcal{D}) \cup \{A \vDash B\} \cup \{A \bowtie X \mid B \sqsubset X \text{ or } B \vDash X \text{ or } B \bowtie X \in \mathsf{rel}(\mathcal{D}), X \not\equiv B\}$$
$$\cup \{X \vDash A \mid X \sqsubset B \in \mathsf{rel}(\mathcal{D})\} \cup \{p \vDash A \mid t \text{ is a point of } \mathcal{D}\}$$

The following is an example of an application of app and obs in this Euler diagrammatic system.



$\forall x(Ax \to \neg Bx)$ app



obs

$\neg Aa$

A **heterogeneous proof**, denoted by $\pi, \pi_1, \pi_2, \ldots$, is defined inductively 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_i is a formula or a diagram.

4 Normalization of heterogeneous proofs

We review the notions of *detour*, *reduction*, and *normal proof* in the usual natural deduction for FOL (cf. [13, 14]) in Section 4.1. Then, we discuss their counterparts for our heterogeneous system in Section 4.2. In Section 4.3, we prove our normalization theorem of heterogeneous proofs. Based on the theorem, we investigate a characterization of the structure of heterogeneous proofs in Section 4.4.

4.1 Normal proof and normalization in FOL

We here recall some basic facts and terminology concerning proof normalization. Cf. [13, 14]. See also [7]. (Notions and properties related to normalization are more neatly formulated in Gentzen's *sequent calculus*, which is a refined system of natural deduction (cf. [8, 6]).)

In general, a natural deduction proof may contain some redundant steps and formulas called **maximal formulas**, i.e., formulas that stands at the same time 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, because without them we already have a proof π_1 of φ_1 as illustrated on the right.

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\pi_1 & \vdots\,\pi_2\\ \varphi_1 & \varphi_2\end{matrix}}{\varphi_1 \wedge \varphi_2}\wedge I}{\varphi_1}\wedge E \qquad\qquad \begin{matrix}\vdots\,\pi_1\\ \varphi_1\end{matrix}$$

A maximal formula along with its related pair of applications of an introduction and an elimination rule are together 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**, and it is defined for every pair of the dual introduction and elimination rules. In addition to the above $\wedge$-reduction rule, the reduction rules for $\to, \vee, \forall, \exists$ are defined as follows, where the part in a given proof on the left is rewritten into a form on the right:

- $\to$-reduction

$$\dfrac{\dfrac{\begin{matrix}[\varphi]\\ \vdots\,\pi_1\\ \psi\end{matrix}}{\varphi \to \psi}\to I \qquad \begin{matrix}\vdots\,\pi_2\\ \varphi\end{matrix}}{\psi}\to E \quad \rhd \quad \begin{matrix}\vdots\,\pi_2\\ \varphi\\ \vdots\,\pi_1\\ \psi\end{matrix}$$

- $\vee$-reduction

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\pi\\ \varphi_i\end{matrix}}{\varphi_1 \vee \varphi_2}\vee I \qquad \begin{matrix}[\varphi_1]\\ \vdots\,\pi_1\\ \sigma\end{matrix} \qquad \begin{matrix}[\varphi_2]\\ \vdots\,\pi_2\\ \sigma\end{matrix}}{\sigma}\vee E \quad \rhd \quad \begin{matrix}\vdots\,\pi\\ \varphi_i\\ \vdots\,\pi_i\\ \sigma\end{matrix} \quad {\scriptstyle (i=1,2)}$$

- $\forall$-reduction

$$\dfrac{\dfrac{\begin{matrix}\vdots\,\pi(x)\\ \varphi(x)\end{matrix}}{\forall x \varphi(x)}\forall I}{\varphi(t)}\forall E \quad \rhd \quad \begin{matrix}\vdots\,\pi(t)\\ \varphi(t)\end{matrix}$$

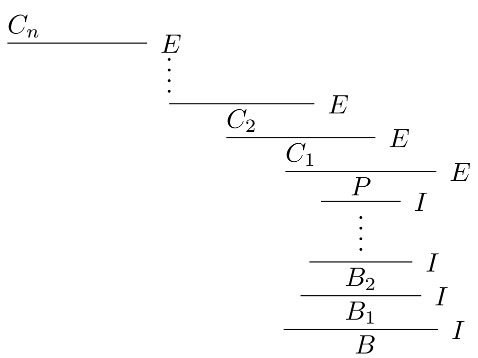
- $\exists$-reduction

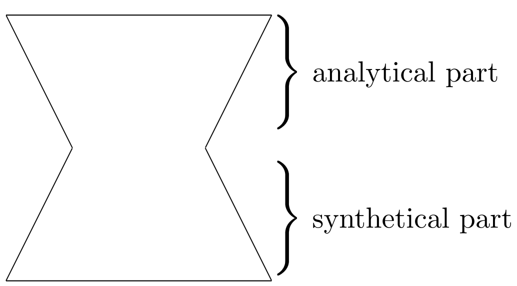
$$\dfrac{\dfrac{\begin{matrix}\vdots\,\pi_1\\ \varphi(t)\end{matrix}}{\exists x \varphi(x)}\exists I \qquad \begin{matrix}[\varphi(x)]\\ \vdots\,\pi_2(x)\\ \sigma\end{matrix}}{\sigma}\exists E \quad \rhd \quad \begin{matrix}\vdots\,\pi_1\\ \varphi(t)\\ \vdots\,\pi_2(t)\\ \sigma\end{matrix}$$

These are the main reduction rules, and see [7, 14] for other technical rules.

A natural deduction proof is said to be in **normal form** when it does not contain any redex, i.e., detour. Then, the following **normalization theorem** holds for the natural deduction system: *If φ is provable from a set of formulas Γ, then there is a normal proof of φ from Γ.*

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 [13, 14] shows that each normal proof consists of two parts: An analytical part in which premises are decomposed into their components by using elimination rules; A synthetic part in which the final components obtained in the analytical part are put together to construct the conclusion by using introduction rules.



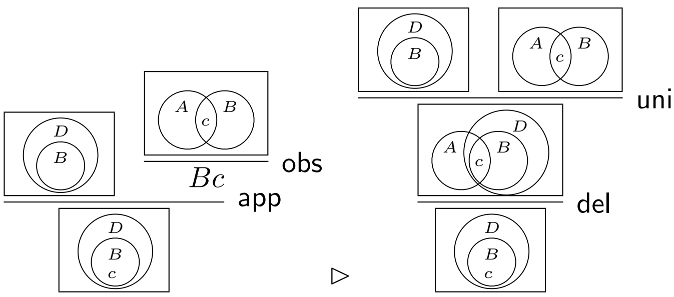


Conceptual form of normal proofs

4.2 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 natural deduction straightforwardly through a translation of diagrams into formulas of FOL. However, as discussed in [18], the resulting "normal" proof does not reflect the characteristics of diagrammatic inference, and it may be called the "linguistic normal" proof. Thus, in this article, we investigate 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 a detour inherent in our diagrammatic inference is. Diagrammatic inference can be characterized by constructing a (maximal) diagram and extracting a conclusion from the diagram. In such a diagrammatic inference, the use of redundant subconcepts (diagrams/formulas) may be considered as a detour. This detour is part of a proof where one infers by deducing subconcepts, even though one can infer directly 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.



$$\dfrac{\begin{matrix}\vdots \\ \mathcal{E}\end{matrix} \quad \begin{matrix}\vdots \\ \dfrac{\mathcal{D}}{\varphi}\,\text{obs} \\ \vdots \\ \psi\end{matrix}}{\mathcal{E}+\psi}\,\text{app} \quad \rhd \quad \dfrac{\dfrac{\begin{matrix}\vdots \\ \mathcal{E}\end{matrix} \quad \begin{matrix}\vdots \\ \mathcal{D}\end{matrix}}{\mathcal{E}+\mathcal{D}}\,\text{uni}}{\mathcal{E}+\psi}\,\text{del}$$

This kind of a detour and its reduction can be formulated by the above rule. When ψ depends only on φ (i.e., $\varphi \vdash \psi$), φ and ψ as well as 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 under the following conditions: (1) $\mathcal{E}+\mathcal{D}$ is defined as a legal diagram; (2) ψ depends only on φ ($\varphi \vdash \psi$). In other words, when we focus and cut out the part from φ to ψ, it stands as a legal proof independent of the other part of the given proof. Otherwise, we cannot deduce ψ from $\mathcal{D}$, as well as $\mathcal{E}+\psi$ from $\mathcal{E}+\mathcal{D}$ after the reduction.

Remark 4.1 Note that the notion of "detour" is conceptual, and it is not necessarily related to the length of proofs. It is known, in symbolic logic proof theory, 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.

Remark 4.2 The following proof on the left cannot be reduced, since $\neg\sigma$ depends not only φ but also on $[\sigma]$, that is, the part from φ to $\neg\sigma$ is not a legal proof in its own right. On the other hand, the proof on the right can be reduced.

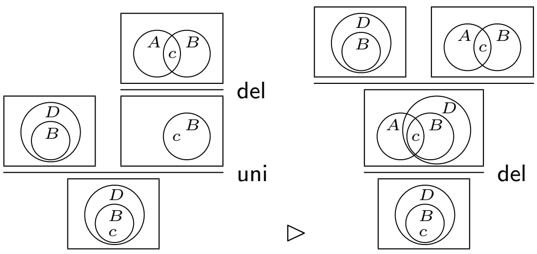
$$\dfrac{\mathcal{E} \qquad \dfrac{\begin{matrix}[\sigma]^1\\ \vdots\\ \dfrac{\mathcal{D}}{\varphi}\,\mathsf{obs}\\ \vdots\\ \bot\end{matrix}}{\neg\sigma}\,1}{\mathcal{E}+\neg\sigma}\,\mathsf{app} \qquad\qquad \dfrac{\begin{matrix}\vdots\\ \mathcal{E}\end{matrix} \qquad \begin{matrix}\vdots\\ \dfrac{\mathcal{D}}{\varphi}\,\mathsf{obs}\end{matrix} \quad \dfrac{\begin{matrix}[\sigma]^1\\ \vdots\\ \bot\end{matrix}}{\neg\sigma}\,1}{\mathcal{E}+\neg\sigma}\,\mathsf{app} \quad \rhd \quad \dfrac{\dfrac{\begin{matrix}\vdots\\ \mathcal{E}\end{matrix} \quad \begin{matrix}\vdots\\ \mathcal{D}\end{matrix}}{\mathcal{E}+\mathcal{D}}\,\mathsf{uni}}{\mathcal{E}+\neg\sigma}\,\mathsf{del}$$

The above obs-app-reduction is generalized as follows: When (1) $\mathcal{E}+\sum\left(\sum\mathcal{D}_n+\sigma^d_l\right)$ is defined, where $\sum\mathcal{D}_n$ is the unified diagram $\mathcal{D}_1+\mathcal{D}_2+\cdots+\mathcal{D}_n$; and (2) $\varphi_1,\ldots,\varphi_n,\sigma^d_1,\ldots,\sigma^d_l \vdash \psi$, the following part in a proof on the left is reduced to the part on the right:

$$\dfrac{\begin{matrix}\vdots\\ \mathcal{E}\end{matrix} \qquad \begin{matrix}\begin{matrix}\vdots\\ \dfrac{\mathcal{D}_1}{\varphi_1}\end{matrix}\,\mathsf{obs}\ldots\begin{matrix}\vdots\\ \dfrac{\mathcal{D}_n}{\varphi_n}\end{matrix}\,\mathsf{obs} \quad \begin{matrix}\vdots\\ \sigma^d_1\end{matrix}\cdots\begin{matrix}\vdots\\ \sigma^d_l\end{matrix}\\ \vdots\\ \psi\end{matrix}}{\mathcal{E}+\psi}\,\mathsf{app} \qquad \rhd \qquad \dfrac{\dfrac{\begin{matrix}\vdots\\ \mathcal{E}\end{matrix} \quad \dfrac{\dfrac{\begin{matrix}\vdots\\ \mathcal{D}_1\end{matrix}\cdots\begin{matrix}\vdots\\ \mathcal{D}_n\end{matrix}}{\sum\mathcal{D}_n}\,\mathsf{uni} \quad \begin{matrix}\vdots\\ \sigma^d_1\end{matrix}\cdots\begin{matrix}\vdots\\ \sigma^d_l\end{matrix}}{\sum\left(\sum\mathcal{D}_n+\sigma^d_l\right)}\,\mathsf{app}}{\mathcal{E}+\sum\left(\sum\mathcal{D}_n+\sigma^d_l\right)}\,\mathsf{uni}}{\mathcal{E}+\psi}\,\mathsf{del}$$

In the above reduced proof on the right, the double line ═══ uni means repeated applications of uni to $\mathcal{D}_1,\ldots,\mathcal{D}_n$; similarly for ═══ app. Each σ^d_i is a diagrammatic formula independent of obs. Since ψ may depend not only on diagrams $\mathcal{D}_1,\ldots,\mathcal{D}_n$ but also on formulas $\sigma^d_1,\ldots,\sigma^d_l$, the similar structure of our detour may occur in such a part. Thus, we generalize our reduction by including σ^d_i to reduce such a part; cf. Fig. 1 of Example 4.5.

Let us further consider the dual pair del and uni of diagrammatic rules. The following proof on the left consisting only of diagrams, is considered to contain a similar detour as the obs-app pair. Without deducing the diagram of Bc by deleting circle A, we may directly unify the given two premise diagrams as illustrated on the right.



$$\dfrac{\begin{matrix}\vdots\\ \mathcal{E}\end{matrix} \qquad \begin{matrix}\vdots\\ \dfrac{\mathcal{F}}{\mathcal{F}'}\,\mathsf{del}\\ \vdots\\ \mathcal{D}\end{matrix}}{\mathcal{E}+\mathcal{D}}\,\mathsf{uni} \quad \rhd \quad \dfrac{\dfrac{\begin{matrix}\vdots\\ \mathcal{E}\end{matrix} \quad \begin{matrix}\vdots\\ \mathcal{F}\end{matrix}}{\mathcal{E}+\mathcal{F}}\,\mathsf{uni}}{\mathcal{E}+\mathcal{D}}\,\mathsf{del}$$

Thus, we can also formulate a reduction rule for a del-uni pair, where, as in the obs-app-reduction, we assume that (1) $\mathcal{E}+\mathcal{F}$ is defined; and (2) $\mathcal{D}$ depends only on $\mathcal{F}'$, i.e., $\mathcal{F}'\vdash\mathcal{D}$.

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.

$$\dfrac{\begin{matrix}\vdots\\ \dfrac{\mathcal{F}}{\mathcal{F}'}\,\mathsf{del}\\ \vdots\\ \mathcal{E}\end{matrix} \qquad \begin{matrix}\vdots\\ \psi\end{matrix}}{\mathcal{E}+\psi}\,\mathsf{app} \quad \rhd \quad \dfrac{\dfrac{\begin{matrix}\vdots\\ \mathcal{F}\end{matrix} \quad \begin{matrix}\vdots\\ \psi\end{matrix}}{\mathcal{F}+\psi}\,\mathsf{app}}{\mathcal{E}+\psi}\,\mathsf{del}$$

Thus, for removing all of the detours of a same kind in a proof, we generalize our reduction to obs/del-app-reduction and del/obs-uni-reduction as follows.

Definition 4.3 obs/del-app-reduction and del/obs-uni-reduction are defined as follows.

- obs/del-app-**reduction (1)**
When $\overrightarrow{\varphi_n}, \overrightarrow{\mathcal{F}'_m}, \overrightarrow{\sigma_l^d} \vdash \psi$ and $\mathcal{E} + \sum\left(\sum\left(\sum D_n + F_m\right) + \sigma_l^d\right)$ is defined in the following part in a proof:

$$\dfrac{\begin{array}{c}\vdots\;\pi'\\ \mathcal{E}\end{array}\qquad \begin{array}{c}\dfrac{\begin{array}{c}\vdots\;\pi_1^1\\ \mathcal{D}_1\end{array}}{\varphi_1}\,\mathsf{obs}\cdots\dfrac{\begin{array}{c}\vdots\;\pi_n^1\\ \mathcal{D}_n\end{array}}{\varphi_n}\,\mathsf{obs}\cdots\dfrac{\begin{array}{c}\vdots\;\pi_1^2\\ \mathcal{F}_1\end{array}}{\mathcal{F}'_1}\,\mathsf{del}\cdots\dfrac{\begin{array}{c}\vdots\;\pi_m^2\\ \mathcal{F}_m\end{array}}{\mathcal{F}'_m}\,\mathsf{del}\cdots\begin{array}{c}\vdots\;\pi_1^3\\ \sigma_1^d\end{array}\cdots\begin{array}{c}\vdots\;\pi_l^3\\ \sigma_l^d\end{array}\\ \vdots\;\pi''\\ \psi\end{array}}{\mathcal{E}+\psi}\,\mathsf{app}$$

where $n \neq 0$ or $m \neq 0$, it is reduced to:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\;\pi'\\ \mathcal{E}\end{array}\qquad \dfrac{\dfrac{\dfrac{\begin{array}{c}\vdots\;\pi_1^1\\ \mathcal{D}_1\end{array}\cdots\begin{array}{c}\vdots\;\pi_n^1\\ \mathcal{D}_n\end{array}}{\sum \mathcal{D}_n}\,\mathsf{uni}\quad \begin{array}{c}\vdots\;\pi_1^2\\ \mathcal{F}_1\end{array}\cdots\begin{array}{c}\vdots\;\pi_m^2\\ \mathcal{F}_m\end{array}}{\sum\left(\sum \mathcal{D}_n + \mathcal{F}_m\right)}\,\mathsf{uni}\quad \begin{array}{c}\vdots\;\pi_1^3\\ \sigma_1^d\end{array}\cdots\begin{array}{c}\vdots\;\pi_l^3\\ \sigma_l^d\end{array}}{\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma_l^d\right)}\,\mathsf{app}}{\mathcal{E} + \sum\left(\sum\left(\sum D_n + F_m\right) + \sigma_l^d\right)}\,\mathsf{uni}}{\mathcal{E}+\psi}\,\mathsf{del}$$

- obs/del-app-**reduction (2)**
When $\overrightarrow{\varphi_n}, \overrightarrow{\mathcal{F}'_m}, \overrightarrow{\sigma_l^d} \vdash \mathcal{E}$ and $\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma_l^d\right) + \psi$ is defined in the following part in a proof:

$$\dfrac{\begin{array}{c}\dfrac{\begin{array}{c}\vdots\;\pi_1^2\\ \mathcal{F}_1\end{array}}{\mathcal{F}'_1}\,\mathsf{del}\cdots\dfrac{\begin{array}{c}\vdots\;\pi_m^2\\ \mathcal{F}_m\end{array}}{\mathcal{F}'_m}\,\mathsf{del}\cdots\dfrac{\begin{array}{c}\vdots\;\pi_1^1\\ \mathcal{D}_1\end{array}}{\varphi_1}\,\mathsf{obs}\cdots\dfrac{\begin{array}{c}\vdots\;\pi_n^1\\ \mathcal{D}_n\end{array}}{\varphi_n}\,\mathsf{obs}\cdots\begin{array}{c}\vdots\;\pi_1^3\\ \sigma_1^d\end{array}\cdots\begin{array}{c}\vdots\;\pi_l^3\\ \sigma_l^d\end{array}\\ \vdots\;\pi'\\ \mathcal{E}\end{array}\qquad\begin{array}{c}\vdots\;\pi''\\ \psi\end{array}}{\mathcal{E}+\psi}\,\mathsf{app}$$

where $m \neq 0$ or $n \neq 0$, it is reduced to:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\vdots\;\pi_1^1\\ \mathcal{D}_1\end{array}\cdots\begin{array}{c}\vdots\;\pi_n^1\\ \mathcal{D}_n\end{array}}{\sum \mathcal{D}_n}\,\mathsf{uni}\quad \begin{array}{c}\vdots\;\pi_1^2\\ \mathcal{F}_1\end{array}\cdots\begin{array}{c}\vdots\;\pi_m^2\\ \mathcal{F}_m\end{array}}{\sum\left(\sum \mathcal{D}_n + \mathcal{F}_m\right)}\,\mathsf{uni}\quad \begin{array}{c}\vdots\;\pi_1^3\\ \sigma_1^d\end{array}\cdots\begin{array}{c}\vdots\;\pi_l^3\\ \sigma_l^d\end{array}}{\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma_l^d\right)}\,\mathsf{app}\quad \begin{array}{c}\vdots\;\pi''\\ \psi\end{array}}{\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma_l^d\right) + \psi}\,\mathsf{app}}{\mathcal{E}+\psi}\,\mathsf{del}$$

- del/obs-uni-**reduction (1)**
When $\overrightarrow{\varphi_n}, \overrightarrow{\mathcal{F}'_m}, \overrightarrow{\sigma^d_l} \vdash \mathcal{D}$ and $\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma^d_l\right) + \mathcal{E}$ is defined in the following part in a proof:

$$
\dfrac{\dfrac{\dfrac{\vdots\,\pi^1_1}{\mathcal{D}_1}}{\varphi_1}\,\mathsf{obs}\ldots\dfrac{\dfrac{\vdots\,\pi^1_n}{\mathcal{D}_n}}{\varphi_n}\,\mathsf{obs}\ldots\dfrac{\dfrac{\vdots\,\pi^2_1}{\mathcal{F}_1}}{\mathcal{F}'_1}\,\mathsf{del}\ldots\dfrac{\dfrac{\vdots\,\pi^2_m}{\mathcal{F}_m}}{\mathcal{F}'_m}\,\mathsf{del}\ldots\overset{\vdots\,\pi^3_1}{\sigma^d_1}\ \cdots\overset{\vdots\,\pi^3_l}{\sigma^d_l} \atop \vdots\,\pi' \atop \mathcal{D} \qquad\qquad \overset{\vdots\,\pi''}{\mathcal{E}}}{\mathcal{D}+\mathcal{E}}\ \mathsf{uni}
$$

where $m \neq 0$ or $n \neq 0$, it is reduced to:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{\vdots\,\pi^1_1}{\mathcal{D}_1}\ \cdots\overset{\vdots\,\pi^1_n}{\mathcal{D}_n}}{\sum \mathcal{D}_n}\,\mathsf{uni}\quad \overset{\vdots\,\pi^2_1}{\mathcal{F}_1}\ \cdots\overset{\vdots\,\pi^2_m}{\mathcal{F}_m}}{\sum\left(\sum \mathcal{D}_n + \mathcal{F}_m\right)}\,\mathsf{uni}\quad \overset{\vdots\,\pi^3_1}{\sigma^d_1}\ \cdots\overset{\vdots\,\pi^3_l}{\sigma^d_l}}{\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma^d_l\right)}\,\mathsf{app}\quad \overset{\vdots\,\pi''}{\mathcal{E}}}{\sum\left(\sum\left(\sum D_n + F_m\right) + \sigma^d_l\right) + \mathcal{E}}\,\mathsf{uni}}{\mathcal{D}+\mathcal{E}}\,\mathsf{del}
$$

- del/obs-uni-**reduction (2)** is defined similarly for π''.

A **redex** is a tuple of applications of rules and diagrammatic formulas (obs, ..., obs, del, ..., del, $\sigma^d_1, \ldots, \sigma^d_l$; app) or (del, ..., del, obs, ..., obs, $\sigma^d_1, \ldots, \sigma^d_l$; uni) to which, a reduction rule can be applied. [2]A heterogeneous proof is said to be in **normal form** when it does not contain any redex.

Remark 4.4 As shown in the following proof of our normalization theorem, we may restrict our del/obs-uni-reduction to del-uni-reduction without including obs.

Example 4.5 (obs-app-reduction) By reducing the obs-app pair of the following proof on the left in Fig. 1, we obtain the normal proof on the right.

Remark 4.6
In a given proof, our redex is not uniquely determined, and there are several redexes with respect to the same app/uni. For example, in the following form of a proof, each of del and obs rules induces a redex with respect to the same app.

$$
\dfrac{\begin{matrix} \vdots \\ \mathcal{E} \end{matrix} \quad \begin{matrix} \dfrac{\dfrac{\mathcal{F}}{\mathcal{F}'}\,\mathsf{del}}{\varphi}\,\mathsf{obs} \\ \vdots \\ \psi \end{matrix}}{\mathcal{E}+\psi}\,\mathsf{app}
$$

Remark 4.7 In the usual natural deduction for FOL, reduction rules are defined for an introduction and an elimination pairs of connectives. On the other hand, our reduction rules for heterogeneous proofs may be considered to be defined for some elimination and an introduction pairs by regarding obs and del as elimination rules (extraction of information) and app and uni as introduction rules (addition of information).

[2]Strictly speaking, we need to distinguish each application of inference rules explicitly to define the notion of redex, although it makes our description cumbersome.

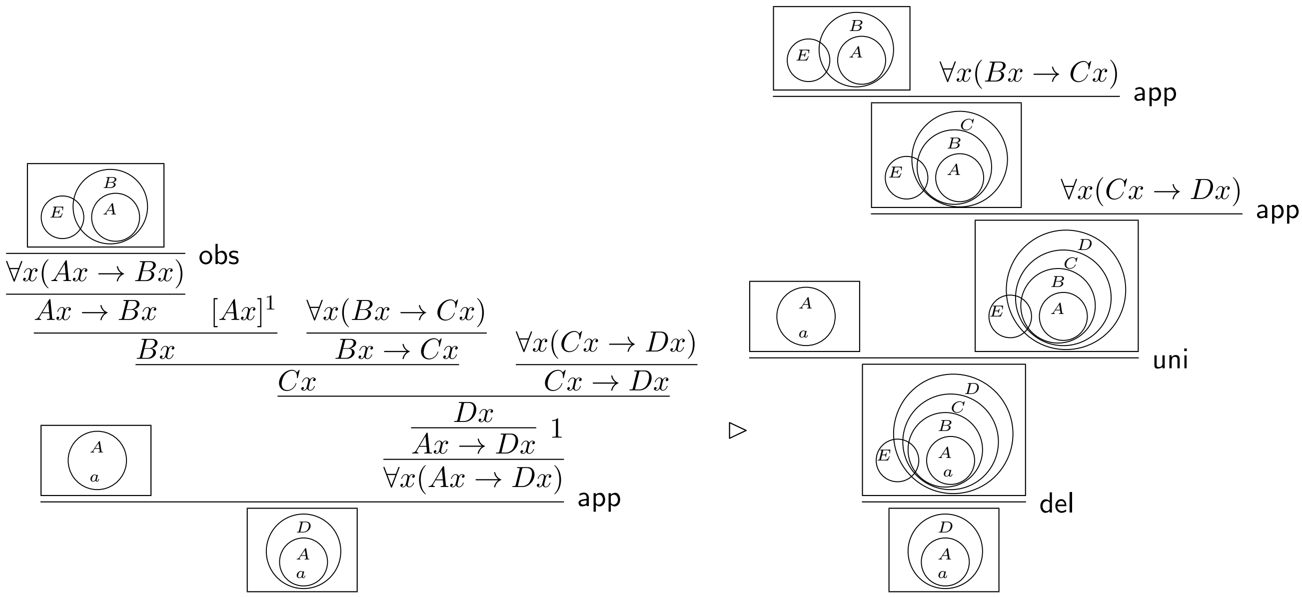


Fig. 1 obs-app-reduction

4.3 Normalization

Based on our reduction rules, we establish our normalization theorem. One of the difficulties is that after an obs/del-app-reduction, new applications of app and uni are provided and they may induce new redexes. To overcome this difficulty, we choose the topmost-leftmost redex in a given proof, and apply our reduction twice in a row.

Theorem 4.8 (Normalization) *Let 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 distinguish linguistic FOL parts and diagrammatic parts in a given proof, and we first reduce the linguistic parts, whose normalization theorem is already established. Let π be a heterogeneous proof whose linguistic parts are already reduced to normal form. For every application of app or uni in π, we define its degree $deg(\mathsf{app})$ or $deg(\mathsf{uni})$ as the number of applications of obs and del that form redexes with respect to the application of app or uni in question. Thus, in any normal form, $deg(\mathsf{app}) = deg(\mathsf{uni}) = 0$ for every application of app and uni in the proof. We choose the topmost-leftmost redex in π, which is the leftmost application of app or uni whose degree is minimal in π. We divide the cases according to the rule: app or uni.

When the rule is app, let the topmost-leftmost redex be the following form:

$$
\dfrac{\dfrac{\dfrac{\overset{\vdots\ \pi_1^1}{\mathcal{E}_1}}{\mathcal{E}'_1}\,\mathsf{del} \cdots \dfrac{\overset{\vdots\ \pi_k^1}{\mathcal{E}_k}}{\mathcal{E}'_k}\,\mathsf{del}}{\overset{\vdots\ \pi'}{\mathcal{E}}} \qquad \dfrac{\dfrac{\overset{\vdots\ \pi_1^2}{\mathcal{D}_1}}{\varphi_1}\,\mathsf{obs} \cdots \dfrac{\overset{\vdots\ \pi_n^2}{\mathcal{D}_n}}{\varphi_n}\,\mathsf{obs} \cdots \dfrac{\overset{\vdots\ \pi_1^3}{\mathcal{F}_1}}{\mathcal{F}'_1}\,\mathsf{del} \cdots \dfrac{\overset{\vdots\ \pi_m^3}{\mathcal{F}_m}}{\mathcal{F}'_m}\,\mathsf{del} \cdots \overset{\vdots\ \pi_1^4}{\sigma_1^d} \cdots \overset{\vdots\ \pi_l^4}{\sigma_l^d}}{\overset{\vdots\ \pi''}{\psi}}}{\mathcal{E} + \psi}\,\mathsf{app}
$$

where every del and obs is the topmost application that forms a redex with respect to the given app, and hence there are no obs nor del in $\overrightarrow{\pi_k^1}, \overrightarrow{\pi_n^2}, \overrightarrow{\pi_m^3}$ that forms a redex with respect to the given app.

Note that there is application of neither obs nor del in $\overrightarrow{\pi^4_l}$. Since, if such an application exists, it has to be one of π^2 or π^3. Furthermore, above $\mathcal{E}$, i.e., in π' and $\overrightarrow{\pi^1_k}$, there is only application of del without obs, or there is application of neither del nor obs (i.e., $k = 0$). This is because, if there is an obs, then there has to be an application of app for inferring the diagram $\mathcal{E}$ and $\mathcal{E}_i$. This contradicts the assumption that the given app forms the topmost-leftmost redex.

By applying del-app-reduction to π', we obtain the following proof:

$$\dfrac{\dfrac{\dfrac{\sum\mathcal{E}_k}{}\!\!\!\!\!\!\!\!\!\!\!\!\;\;\;\;\;\;\;\;\;\;\;\;\begin{array}{c}\vdots\,\pi^1_1 \quad \vdots\,\pi^1_k\\ \mathcal{E}_1 \quad \cdots\mathcal{E}_k\\ \hline\hline \sum\mathcal{E}_k\end{array}\ \textsf{uni} \qquad \begin{array}{c} \begin{array}{c}\vdots\,\pi^2_1\\ \mathcal{D}_1\\ \hline \varphi_1\end{array}\textsf{obs}\cdots\begin{array}{c}\vdots\,\pi^2_n\\ \mathcal{D}_n\\ \hline \varphi_n\end{array}\textsf{obs}\cdots\begin{array}{c}\vdots\,\pi^3_1\\ \mathcal{F}_1\\ \hline \mathcal{F}'_1\end{array}\textsf{del}\cdots\begin{array}{c}\vdots\,\pi^3_m\\ \mathcal{F}_m\\ \hline \mathcal{F}'_m\end{array}\textsf{del}\cdots\begin{array}{c}\vdots\,\pi^4_1\\ \sigma^d_1\end{array}\cdots\begin{array}{c}\vdots\,\pi^4_l\\ \sigma^d_l\end{array} \\ \vdots\,\pi'' \\ \psi\end{array}}{\sum\mathcal{E}_k+\psi}\,\textsf{app}}{\mathcal{E}+\psi}\,\textsf{del}$$

By further applying obs/del-app-reduction to π'', we obtain the following proof:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\pi^1_1 \quad \vdots\,\pi^1_k\\ \mathcal{E}_1 \quad \cdots\mathcal{E}_k\\ \hline\hline \sum\mathcal{E}_k\end{array}\ \textsf{uni} \qquad \begin{array}{c}\begin{array}{c}\begin{array}{c}\vdots\,\pi^2_1 \quad \vdots\,\pi^2_n\\ \mathcal{D}_1 \quad \cdots\mathcal{D}_n\\ \hline\hline \sum\mathcal{D}_n\end{array}\ \textsf{uni} \quad \begin{array}{c}\vdots\,\pi^3_1 \quad \vdots\,\pi^3_m\\ \mathcal{F}_1 \quad \cdots\mathcal{F}_m\end{array}\\ \hline\hline \sum\left(\sum\mathcal{D}_n+\mathcal{F}_m\right)\end{array}\ \textsf{uni} \quad \begin{array}{c}\vdots\,\pi^4_1 \quad \vdots\,\pi^4_l\\ \sigma^d_1 \quad \cdots\sigma^d_l\end{array}\\ \hline\hline \sum\left(\sum\left(\sum\mathcal{D}_n+\mathcal{F}_m\right)+\sigma^d_l\right)\end{array}\ \textsf{app}}{\sum\mathcal{E}_k+\sum\left(\sum\left(\sum\mathcal{D}_n+\mathcal{F}_m\right)+\sigma^d_l\right)}\,\textsf{uni}}{\mathcal{E}+\psi}\,\textsf{del}$$

Although new applications of uni and app are produced, there is neither obs nor del in $\overrightarrow{\pi^1_k}, \overrightarrow{\pi^2_n}, \overrightarrow{\pi^3_m}, \overrightarrow{\pi^4_l}$ that forms a new redex with respect to these uni and app. This is justified because, if there is an obs or del that forms a new redex, then it must have already formed a redex in the original proof, which contradicts the assumption that the deleted obs and del are topmost. Thus, the degrees of these new uni and app are 0.

The case where the topmost-leftmost redex consists of an application of uni, let us assume it is the following form:

$$\dfrac{\begin{array}{c}\begin{array}{c}\vdots\,\pi^1_1\\ \mathcal{D}_1\\ \hline \mathcal{D}'_1\end{array}\textsf{del}\cdots\begin{array}{c}\vdots\,\pi^1_n\\ \mathcal{D}_n\\ \hline \mathcal{D}'_n\end{array}\textsf{del}\\ \vdots\,\pi'\\ \mathcal{D}\end{array} \qquad \begin{array}{c}\begin{array}{c}\vdots\,\pi^2_1\\ \mathcal{E}_1\\ \hline \mathcal{E}'_1\end{array}\textsf{del}\cdots\begin{array}{c}\vdots\,\pi^2_m\\ \mathcal{E}_m\\ \hline \mathcal{E}'_n\end{array}\textsf{del}\\ \vdots\,\pi''\\ \mathcal{E}\end{array}}{\mathcal{D}+\mathcal{E}}\,\textsf{uni}$$

where every del is the topmost application that forms a redex with respect to the given uni, and hence there are no del nor obs in $\overrightarrow{\pi^1_n}, \overrightarrow{\pi^2_m}$ that forms a redex with respect to the given uni.

Note that there is no application of obs that forms a redex with respect to the uni. This is because, if there is an obs, then there has to be an application of app to derive the diagram $\mathcal{D}$ or $\mathcal{E}$. This contradicts the assumption that the given uni forms the topmost-leftmost redex.

By applying del-uni-reductin to π', we obtain the following proof:

$$\dfrac{\dfrac{\dfrac{\overset{\vdots\ \pi_1^1}{\mathcal{D}_1}\quad\cdots\overset{\vdots\ \pi_n^1}{\mathcal{D}_n}}{\sum\mathcal{D}_n}\ \mathsf{uni}\qquad \overset{\dfrac{\overset{\vdots\ \pi_1^2}{\mathcal{E}_1}}{\mathcal{E}_1'}\ \mathsf{del}\ \cdots\dfrac{\overset{\vdots\ \pi_m^2}{\mathcal{E}_m}}{\mathcal{E}_n'}\ \mathsf{del}}{\overset{\vdots\ \pi''}{\mathcal{E}}}}{\sum\mathcal{D}_n+\mathcal{E}}\ \mathsf{uni}}{\mathcal{D}+\mathcal{E}}\ \mathsf{del}$$

By further applying del-uni-reduction to π'', we obtain the following proof:

$$\dfrac{\dfrac{\dfrac{\overset{\vdots\ \pi_1^1}{\mathcal{D}_1}\quad\cdots\overset{\vdots\ \pi_n^1}{\mathcal{D}_n}}{\sum\mathcal{D}_n}\ \mathsf{uni}\qquad \dfrac{\overset{\vdots\ \pi_1^2}{\mathcal{E}_1}\quad\cdots\overset{\vdots\ \pi_m^2}{\mathcal{E}_m}}{\sum\mathcal{E}_m}\ \mathsf{uni}}{\sum\mathcal{D}_n+\sum\mathcal{E}_m}\ \mathsf{uni}}{\mathcal{D}+\mathcal{E}}\ \mathsf{del}$$

Although new applications of uni are produced, there is neither obs nor del in $\overrightarrow{\pi_n^1}, \overrightarrow{\pi_m^2}$ that forms a new redex with respect to these uni. This is justified because, if there is an obs or del that forms a new redex, then it must have already formed a redex in the original proof, which contradicts the assumption that the deleted del are topmost. Thus, the degrees of these new uni are 0.

Let the degree $deg(\pi)$ of the given proof π be the sum of all degrees of applications of app and uni in π. Let $\pi^\star$ be the proof obtained by the above topmost-leftmost reduction. Then, we have $deg(\pi^\star) < deg(\pi)$. Therefore, by repeated applications of the topmost-leftmost reduction, we obtain a proof whose degree is 0, i.e., we obtain a normal proof. ■

Example 4.9 (Normalization) By repeatedly reducing the topmost-leftmost redex, we obtain a normal proof as shown in Fig. 2.

Note that in the normal form, the application of uni seems to be redundant, since one of its premises and its conclusion are the same diagrams. There may be this kind of redundancy in our normal form, as we see in the usual normal form in natural deduction for FOL. However, this kind of redundancy is different from our essential detour, which uses redundant subconcepts. Thus, we leave this kind of inessential redundant parts untouched in our normal form.

By applying our reduction, every redundant linguistic part that lies between diagrammatic parts is removed. From the perspective of diagrams reducing certain complexity of linguistic inference, it is ineffective to infer diagrammatically by way of some linguistic parts, and hence our reduction is also verified from this perspective.

4.4 Characterization of normal heterogeneous proofs

Let us investigate how diagrammatic inference and linguistic inference appear, and are related in our heterogeneous proofs. The following proposition holds in a system that does not constrain any inference rules such as Venn and Euler systems without any points.

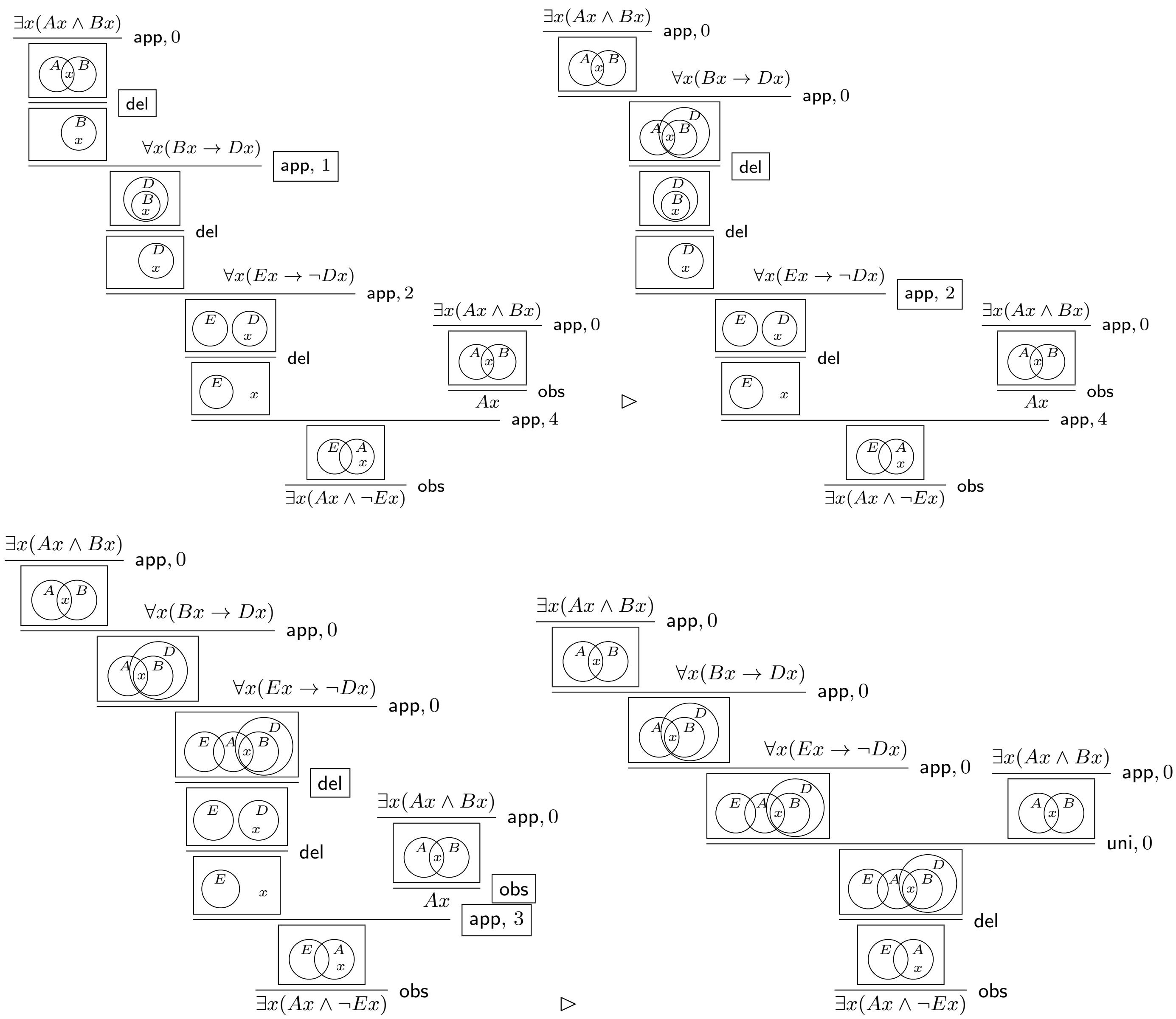
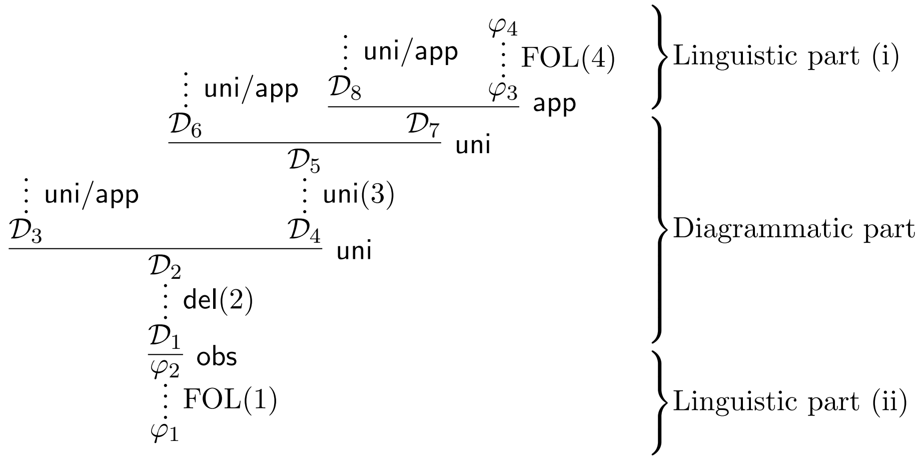


Fig. 2 Normalization. (The number beside an application of app or uni is its degree.)

Proposition 4.10 (Normal form) *In a heterogeneous system, where* app *and* uni *are applicable to any diagram and formula without any constraint, every normal proof has the following form:*



Proof. Starting from a given conclusion, we examine possible rules in a normal proof from the bottom-up. The following number corresponds to that in the above figure. Assume that the given conclusion is a formula φ_1. When the conclusion is a diagram, we start from the following case (2).

(1) φ_1 should be obtained by FOL-rules or obs, since the conclusion of other rules del,uni, and app is a diagram. Thus, we assume without loss of generality, that φ_1 is obtained after application of FOL-rules followed by an application of obs from the bottom-up.

(2) Above the obs of (1), possible rules are del, uni, and app. (FOL-rules and obs are not possible, since their conclusion has to be a formula.) Since there is no del above uni or app in a normal proof, we assume without loss of generality, that del is applied certain times.

(3) Above the del-rule of (2), possible rules are uni or app. Although any of them is possible, we assume without loss of generality, that $\mathcal{D}_2$ is obtained by some applications of uni followed by an application of app.

(4) Above the app of (3), the only possible rule is one of the FOL-rules, since there is no obs (nor del) above app in a normal proof. Therefore, only successive applications of FOL-rules are possible. ■

By Proposition 4.10 above, a normal heterogeneous proof is divided into the following three parts from the top-down, which also indicates a strategy to construct heterogeneous proofs.

Linguistic part (i) By FOL-rules, given premises represented by formulas are decomposed, and apply-formulas are constructed.

Diagrammatic part uni is applied to given premises represented by diagrams, and app is applied to apply-formulas obtained at the linguistic part (i), and a maximal diagram is constructed. Then, by del and obs, diagrammatic formulas are extracted from the maximal diagram.

Linguistic part (ii) By FOL-rules, the conclusion is constructed.

Example 4.11 (Normal proof) Fig. 3 is an example of a typical normal heterogeneous proof.

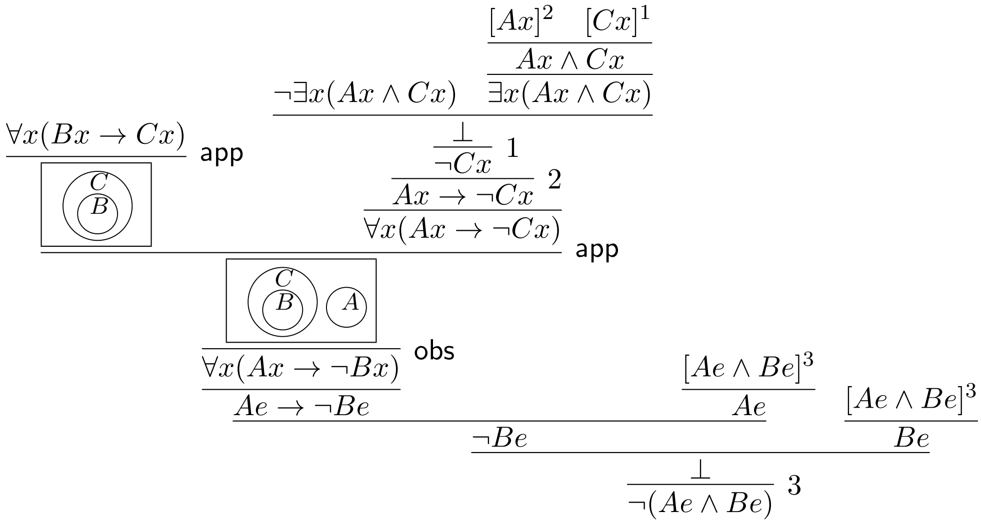


Fig. 3 Normal proof

5 Discussion and future work

By slightly extending the notion of free ride of Shimojima [16], let us call diagrammatic formulas of the conclusion of app and uni, *free rides*, if they do not appear in the given premise diagrams or sentences, but (automatically) appear in its conclusion (cf. [18]). From the perspective of symbolic specification, a diagram is a deductively closed set of diagrammatic formulas. The deductive closedness of diagrams induces the free rides. The larger a diagram is, the more free rides appear in general. Thus, since a maximal diagram is constructed in our normal heterogeneous proof, we may say that a normal heterogeneous proof takes full advantage of the free rides. However, from a cognitive standpoint, a maximal diagram is not necessarily comprehensible or manageable. This is because, the more complex a diagram is, more cognitive cost is required to construct and read the diagram in general. In [11], to make proofs readable by avoiding clutter in diagrams, tactics are introduced to an interactive theorem prover for spider diagrams Speedith [23].

It is often pointed out that there is a trade-off between the expressive power and the cognitive clarity/complexity of diagrams. In general, on top of inherent geometrical constraints of diagrams, if we increase their expressive power by introducing various conventional devices (for example, linking between points as well as between diagrams), then it is appropriate that the cognitive clarity of the diagrams is decreased. Conversely, if we restrict introducing conventional devices, then such diagrams maintain their cognitive clarity in exchange for limited expressive power.

Our characterization of heterogeneous proofs of Proposition 4.10 shows another trade-off between constraints on inference rules, and the complexity of the structure of proofs or of the strategy to construct proofs. As discussed in Section 3.2, there may be various constraints on inference rules such as constraints for indeterminacy and for contradiction of app and uni. Such a constraint is mainly imposed to avoid cognitive complexity or to maintain actual feasibility of the rule. (Although a constraint on inference rules pertains to expressive limitation of diagrams, they are not the same.) The characterization of the structure for heterogeneous proofs of Proposition 4.10 is valid for systems without constraints on inference rules app and uni. Thus, in a system with some constraints, we cannot apply our strategy to construct proofs in a straightforward manner. We need a more complex strategy or worse, a heuristic method to construct proofs in such a system. In general, within a system with various constraints on inference rules, although cognitive clarity and actual feasibility of the rules are maintained due to those constraints, the structure of proofs in such a system becomes complex and an automatic strategy to construct proofs cannot be applied. Conversely, in a system with few constraints on rules, the structure of proofs in such a system is simpler and we are able to apply an automatic strategy to construct proofs. However, the cognitive clarity and actual feasibility of each inference rule are decreased in such a system.

In this article, we restrict our type of a diagram to the conjunction of diagrammatic formulas thereof, and hence, we exclude from our consideration, the linking between diagrams that represents certain kind of disjunctive information. Similarly, we do not investigate the rule of so-called Cases Exhaustive (cf. [2]), which is a generalization of our app, and which handles disjunctive information. For future work, we aim to extend our framework in order to include such diagrams representing disjunctive information.

One of the most important characteristics of the notion of reduction in the usual FOL is that it corresponds to the notion of computation via the Curry-Howard correspondence between formulas and types, as well as proofs and programs (cf. [14]). Our reduction is different from the usual one, and it is interesting to investigate its counterparts in the theory of computation.

References

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

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

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

[4] D. Barker-Plummer and N. 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] J. Barwise and J. Etchemendy, *Hyperproof: For Macintosh*, The Center for the Study of Language and Information Publications, 1995.

[6] S. R. Buss, An Introduction to Proof Theory, in *Handbook Proof Theory*, S. R. Buss ed., Elsevier, Amsterdam, 1998.

[7] D. van Dalen, *Logic and Structure*, Springer, 4th ed. 2004.

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

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

[11] S. Linker, J. Burton, M. Jamnik, Tactical Diagrammatic Reasoning, *Proceedings of the 12th Workshop on User Interfaces for Theorem Provers*, EPTCS 239, 29-42, 2017.

[12] K. Mineshima, M. Okada, and R. Takemura, A Diagrammatic Inference System with Euler Circles, *Journal of Logic, Language and Information*, 21, 3, 365-391, 2012.

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

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

[15] S.-J. Shin, *The Logical Status of Diagrams*, Cambridge University Press, 1994.

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

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

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

[19] R. Takemura, A Heterogeneous Logic with Tables, *Proceedings of International Workshop on Diagrams, Logic and Cognition (DLAC 2013)*, Jim Burton, Lopamudra Choudhury, eds., CEUR Series Vol. 1132, 9-16, 2014.

[20] R. Takemura, A. Shimojima, and Y. Katagiri, A Logical Investigation on Global Reading of Diagrams, *Diagrams 2012*, LNCS, Volume 7352, 330-333, 2012.

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

[22] M. Urbas and M. Jamnik, A Framework for Heterogeneous Reasoning in Formal and Informal Domains, *Diagrams 2014*, LNCS, Volume 8578, 277-292, 2014.

[23] M. Urbas, M. Jamnik, G. Stapleton, Speedith: A Reasoner for Spider Diagrams, *Journal of Logic, Language and Information*, 24 (4), 487-540, 2015.