Generalized Syllogistic Inference System based on Inclusion and Exclusion Relations (Extended Abstract)

Koji Mineshima, Mitsuhiro Okada, and Ryo Takemura

Department of Philosophy, Keio University.
{minesima,mitsu,takemura}@abelard.flet.keio.ac.jp

1 Introduction

Entailment relations are of central importance in the enterprise of natural language semantics. In modern logic, entailment relations are characterized from two viewpoints: the model-theoretic and proof-theoretic ones. By contrast, most approaches to formalizing entailment relations in natural languages have been solely based on model-theoretic conceptions. Thus, the notion of validity is only characterized in model-theoretic terms, and only few attempts have been made at the relevant proof-theoretic notions such as provability and proof as applied to natural language inferences. The main aim of our study is to offer a simple inference system for *syllogistic* fragment of natural language, thereby making a connection between proof theory and natural language semantics.

Indeed, there are several prior attempts in this direction.[1] One is the study of syllogisms from modern logical viewpoints, which was started by Łukasiewicz [7] and subsequently given their natural deduction formulations by Corcoran [2] and Smiley [23], among others. In connection with this, the program of natural logics, including the so-called calculus of monotonicity proposed in van Benthem [24] and Sanchez [20], has been developed by some linguists and logicians.[2] More recently, there are important developments in the logic and AI literature which aim to extend the traditional syllogistic system to cover more expressive fragments of natural language (cf. Nishihara, Morita, and Iwata [17], Nishihara and Morita [16], Moss [13, 14, 15], and Pratt-Hartmann and Moss [18]). The central aim of these studies is to characterize natural language inferences at the forms as close as possible to surface forms.

Our approach agrees with these studies in that a proof system plays a role in characterizing entailment relations in natural languages. However, an essential difference is that we *decompose* syllogistic inferences and categorical sentences constituting them in terms of more primitive relations (i.e., inclusion and exclusion relations), whereas most approaches in modern logical studies of syllogisms and natural logic take surface forms as primitive and do

[1]As is well known, there is an influential approach applying proof theory to natural language *syntax*, which started with Lambek calculus and has been more fully explored in the recent study of categorial grammar such as Type-Logical Grammar. But our interest here is in applying proof theory to natural language *semantics*; the goal is to represent and analyze semantic notions relevant to natural language inferences such as entailments, rather than syntactic structures of sentences, in proof-theoretic terms.

[2]For recent overviews, see van Eijck [26] and van Benthem [25]. Natural logic has also been applied to the study of textual inference in the field of natural language processing (cf. MacCartney and Manning [8]).

not attempt to reduce them into more primitive forms. We call our inference system "Generalized Syllogistic inference system," which is abbreviated as GS. The inference system GS is formulated in Gentzen's natural deduction style. We prove a normalization theorem, and based on it, we show that the proofs in GS correspond to the chains of categorical syllogisms.

Another important difference between previous approaches and ours is that our system is closely related to an inference system for Euler diagrams. It is well known that Euler diagrams can be used to represent not only Aristotelian categorical sentences but also syllogistic reasoning based on them. Such diagrammatic inferences have been studied in the field of *diagrammatic logic*, which was initiated by philosophers and logicians in the 1990's.[3] However, the exact formulation of inference rules directly operating on Euler diagrams was not clear until recently.[4] Mineshima, Okada, Sato and Takemura [9] and Mineshima, Okada, and Takemura [11] are the first attempts to provide a complete inference system for Euler diagrams that are defined in terms of topological relations between circles.[5] The system is called "Generalized Diagrammatic Syllogistic inference system," which is abbreviated as GDS.

It should be noted that there is a certain similarity between natural language inferences and diagrammatic inferences: the complexities of the representations involved, that is, natural language sentences and diagrams, are limited, as compared to formulas of usual logical languages. Thus, it is often observed that in natural languages, there are certain restrictions on nested occurrences of operators, such as negation, implication, and modality. Similarly, it is well known that usual logic diagrams such as Euler and Venn diagrams do not have devices corresponding to nested logical operators, unless the systems are enriched with additional conventions. By restricting nested structures, the resulting systems can be more efficient and human-oriented tools in actual communication and reasoning.[6]

In view of the connection between natural language inferences and diagrammatic inferences, it is interesting to develop an underlying proof system for these two types of inferences. Indeed, the inference system GS is intended to serve as such a system. The two primitive relations in GS, namely, inclusion (symbolized as $\sqsubset$) and exclusion (symbolized as $\vdash\!\dashv$), are the ones used to define Euler diagrams in our diagrammatic representation system [9, 11].[7] The inference system GS provides a bridge between linguistic syllogistic inferences and Euler-style diagrammatic inferences, and more generally, between natural logic and diagrammatic logic. This opens an interesting possibility to connect studies of natural language inferences with studies of visual or diagrammatic inferences, thereby making it possible to compare these two kinds of inferences in a unified and rigorous logical perspective.

In order to investigate the expressive power of the underlying inference system GS, we will show that it corresponds to a *propositional* fragment of minimal logic, where by minimal logic we mean intuitionistic logic without $\perp$ (absurdity) rule. More specifically, GS corresponds to the disjunction-free (i.e. $\wedge$, $\rightarrow$, $\neg$) fragment of minimal logic. The basic idea is to translate GS formula $A \sqsubset B$, which says that A is included in B, as $A \rightarrow B$ in propositional logic, and GS formula $A \vdash\!\dashv B$, which says that A and B are disjoint, as $A \rightarrow \neg B$. In this way,

[3]See Barwise and Etchemendy [1], Shin [22], and Hammer [5].

[4]For a discussion on the difficulty of formalizing reasoning with Euler diagrams, see Hammer and Shin [6].

[5]Euler diagrams are usually formalized based on the method developed in the study on Venn diagrams, where diagrams are abstractly defined as a set of regions. We call such an approach a "region-based" approach. By contrast, we developed a "relation-based" approach, where diagrams are abstractly defined as a set of relations. A comparison between the relation-based approach and the traditional region-based approach is given in Mineshima, Okada and Takemura [12].

[6]Cf. Sato, Mineshima, and Takemura [21] for a cognitive psychological study on efficacy of logic diagrams.

[7]The symbols $\sqsubset$ and $\vdash\!\dashv$ were introduced by Gergonne [4] to represent Euler diagrams symbolically.

the inclusion and exclusion relations between terms are treated like two kinds of *implication* between propositions.

The overall picture is summarized as follows.

Natural Language Inferences	Diagrammatic Inferences
Categorical Syllogism	Diagrammatic Inference System (GDS)

Underlying Inference System (GS)
for inclusion ($\sqsubset$) and exclusion ($\vdash\!\dashv$) relations between terms

Natural Deduction System (ML)
for a propositional ($\wedge$, $\to$, $\neg$) fragment of minimal logic

For reasons of space, we will only consider GS's relations to categorical syllogism and natural deduction system of minimal logic. Also for reasons of space, we cannot discuss possible extensions of GS. In [10], we discuss some extensions of our diagrammatic representation system, including extensions with intersections and unions. It is possible to extend GS in a parallel way. (For example, in a system with intersection we can handle categorical sentences with conjunctive terms as discussed in Nishihara and Morita [16].)

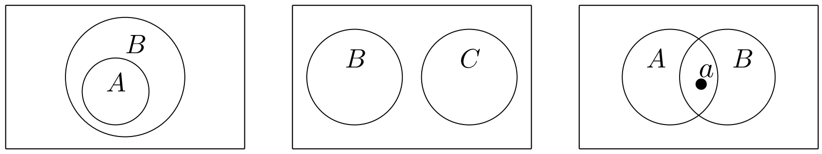
The structure of this paper is as follows. In Section 2, we introduce the Generalized Syllogistic inference system GS, and show that the system is complete with respect to its set-theoretical semantics. We also provide a normalization theorem, which roughly states that every proof of the system has a normal form corresponding to the form of Aristotelian syllogisms. Based on this result, in Section 3, we show the correspondence between GS and categorical syllogism. In Section 4, we show the correspondence between GS and the propositional minimal logic.

2 Generalized Syllogistic inference system GS

Let us start with some examples. Consider the following two syllogisms, which are traditionally called Celarent and Darii.

$$\frac{\textsf{No } B \textsf{ are } C \quad \textsf{All } A \textsf{ are } B}{\textsf{No } A \textsf{ are } C}\ \textsf{Celarent} \qquad \frac{\textsf{All } B \textsf{ are } C \quad \textsf{Some } A \textsf{ are } B}{\textsf{Some } A \textsf{ are } C}\ \textsf{Darii}$$

In our approach, the categorical sentences appeared in the premises and conclusions are analyzed in terms of two primitive relations: inclusion ($\sqsubset$) and exclusion ($\vdash\!\dashv$). To begin with, we translate universal sentence All A are B as $A \sqsubset B$ and No B are C as $B \vdash\!\dashv C$, where $\sqsubset$ and $\vdash\!\dashv$ are semantically interpreted as the subset and disjointness relations, respectively. The notation and terminology come from the logical study of Euler diagrams (see [9, 11]), where these two relations are also used to formalize Euler diagrams. Fig.1 below shows the correspondence between categorical sentences and Euler diagrams. Here diagrams D_1 and D_2 can also be analyzed as $A \sqsubset B$ and $B \vdash\!\dashv C$ at the abstract level, respectively. Existential sentence Some A are B has a diagrammatic representation D_3 in Fig.1. This diagram suggests that the existential statement is decomposed into two primitive assertions, $a \sqsubset A$ and $a \sqsubset B$,



D_1: All A are B $\quad$ D_2: No B are C $\quad$ D_3: Some A are B

Fig. 1 Correspondences between categorical sentences and Euler diagrams.

where a is a constant denoting a witness of the existential sentence. Semantically, constants are interpreted as a singleton set, which enables us to treat $\sqsubset$ as the subset relation in a uniform way. Given these translations, Celarent and Darii are simulated as follows.

$$\dfrac{\begin{array}{c}\text{No } B \text{ are } C\\ B \vdash\!\!\dashv C\end{array} \qquad \begin{array}{c}\text{All } A \text{ are } B\\ A \sqsubset B\end{array}}{\begin{array}{c}A \vdash\!\!\dashv C\\ \text{No } A \text{ are } C\end{array}}\,(\vdash\!\!\dashv)$$

$$\dfrac{\dfrac{\begin{array}{c}\text{Some } A \text{ are } B\\ a \sqsubset A,\, a \sqsubset B\end{array}}{a \sqsubset A}\,(-) \qquad \dfrac{\dfrac{\begin{array}{c}\text{Some } A \text{ are } B\\ a \sqsubset A,\, a \sqsubset B\end{array}}{a \sqsubset B}\,(-) \qquad \begin{array}{c}\text{All } B \text{ are } C\\ B \sqsubset C\end{array}}{a \sqsubset C}\,(\sqsubset)}{\begin{array}{c}a \sqsubset A,\, a \sqsubset C\\ \text{Some } A \text{ are } C\end{array}}\,(+)$$

Fig. 2 Celarent (left) and Darii (right) in our inference system.

The crucial inference rules here are the transitivities of $\sqsubset$ and $\vdash\!\!\dashv$, which are labeled here as $(\sqsubset)$ and $(\vdash\!\!\dashv)$, respectively.

Now let us introduce the syntax and semantics of GS. We first present the language of GS.

Definition 2.1 (Language)
Terms of GS are *general terms*, denoted by $A, B, C, \ldots$, and *singular terms*, denoted by $a, b, c, \ldots$. They are collectively denoted by $s, t, u, \ldots$.
Relations between terms are $\sqsubset$ (inclusion relation) and $\vdash\!\!\dashv$ (exclusion relation).
Atoms of GS are of the form $s \sqsubset t$ or $s \vdash\!\!\dashv t$ and denoted by $P, Q, \ldots$.
A *formula* of GS is a non-empty set $\{P_1, \ldots, P_n\}$ of atoms and denoted by $\mathcal{P}, \mathcal{Q}, \ldots$.

We assume that the relation $\vdash\!\!\dashv$ is symmetric, and syntactically identifies $s \vdash\!\!\dashv t$ with $t \vdash\!\!\dashv s$.

Our definition of formulas deviates from the standard one in that we call a set of atoms $\{P_1, \ldots, P_n\}$ a formula, but this makes comparisons of GS with categorical syllogisms and Euler diagrams simpler. A formula $\mathcal{P} = \{P_1, \ldots, P_n\}$ can be considered as a conjunction $P_1 \wedge \cdots \wedge P_n$ of atoms (cf. Definition 2.3 of semantics below). For simplicity, we sometimes omit braces $\{$ and $\}$ of a set of atoms.

Next we introduce a set-theoretical semantics of GS.

Definition 2.2 A **structure** M is a pair (U, I), where U is a non-empty set (the domain of M), and I is an interpretation function which assigns to each term s a non-empty subset of U. In particular, $I(a)$ is a singleton for any singular term a, and $I(a) \neq I(b)$ for any distinct singular terms a and b.

Definition 2.3 Let $\mathcal{P}$ be a formulas of GS. $M = (U, I)$ is a **model** of $\mathcal{P}$, written as $M \models \mathcal{P}$, if the following **truth-conditions** (1) and (2) hold: for all terms s, t of $\mathcal{P}$,

$$(1)\ \ I(s) \subseteq I(t)\ \text{ if } s \sqsubset t \in \mathcal{P}, \quad \text{and} \quad (2)\ \ I(s) \cap I(t) = \emptyset\ \text{ if } s \vdash\!\!\dashv t \in \mathcal{P}.$$

Note that when s is an singular term a, we have $I(a) = \{e\}$ for some $e \in U$, and $I(a) \subseteq I(t)$ in (1) above is equivalent to $e \in I(t)$. Similarly, $I(a) \cap I(t) = \emptyset$ in (2) is equivalent to $e \notin I(t)$.

$\mathcal{Q}$ is a **semantically valid consequence** of $\mathcal{P}_1, \ldots, \mathcal{P}_n$ written as $\mathcal{P}_1, \ldots, \mathcal{P}_n \models \mathcal{Q}$, when the following holds: for any model M, if $M \models \mathcal{P}_i$ for all $1 \leq i \leq n$, then $M \models \mathcal{Q}$.

The syntax of GS is defined in the style of Gentzen's natural deduction.

Definition 2.4 (Inference rules) Inference rules of GS are defined as follows:

$$\frac{}{s \sqsubset s}\ ax \qquad \frac{s \sqsubset t \quad t \sqsubset u}{s \sqsubset u}\ (\sqsubset) \qquad \frac{s \sqsubset t \quad t \vDash\!\!\dashv u}{s \vDash\!\!\dashv u}\ (\vDash\!\!\dashv) \qquad \frac{\mathcal{P} \quad \mathcal{Q}}{\mathcal{P} \cup \mathcal{Q}}\ (+) \qquad \frac{\mathcal{P}}{\mathcal{P}'}\ (-) \text{ for } \mathcal{P}' \subsetneq \mathcal{P}$$

For $(+)$ rule, we assume that $\mathcal{P}$ and $\mathcal{Q}$ are different sets, i.e., $\mathcal{P} \neq \mathcal{Q}$.

Since a formulas $\mathcal{P} = \{P_1, \ldots, P_n\}$ means the conjunction $P_1 \wedge \cdots \wedge P_n$, $(+)$ and $(-)$ rules can be considered as generalizations of usual $\wedge$-introduction and $\wedge$-elimination rules of Gentzen's natural deduction.

Proofs of GS are defined inductively in the usual way.

Definition 2.5 (GS-proof) *Proofs* of GS are defined inductively as follows:

1. A formula $\mathcal{P}$ is a proof from the premise $\mathcal{P}$ to the conclusion $\mathcal{P}$.
2. Let π_1 be a proof of $\mathcal{Q}$ from $\mathcal{P}_1, \ldots, \mathcal{P}_n$, and π_2 be a proof of $\mathcal{S}$ from $\mathcal{R}_1, \ldots, \mathcal{R}_m$, respectively. If $\mathcal{T}$ is obtained by an application of $(\sqsubset), (\vDash\!\!\dashv), (+)$ rules to $\mathcal{Q}$ and $\mathcal{S}$, then the following (i) is a proof of $\mathcal{T}$ from $\mathcal{P}_1, \ldots, \mathcal{P}_n, \mathcal{R}_1, \ldots, \mathcal{R}_m$.
3. Let π_1 be a proof of $\mathcal{Q}$ from $\mathcal{P}_1, \ldots, \mathcal{P}_n$. If $\mathcal{T}$ is obtained by applying $(-)$ rule to $\mathcal{Q}$, then the following (ii) is a proof of $\mathcal{T}$ from $\mathcal{P}_1, \ldots, \mathcal{P}_n$.

$$\text{(i)}\quad \dfrac{\begin{matrix}\mathcal{P}_1 \cdots \mathcal{P}_n \\ \vdots\ \pi_1 \\ \mathcal{Q}\end{matrix} \qquad \begin{matrix}\mathcal{R}_1 \cdots \mathcal{R}_m \\ \vdots\ \pi_2 \\ \mathcal{S}\end{matrix}}{\mathcal{T}} \qquad\qquad \text{(ii)}\quad \dfrac{\begin{matrix}\mathcal{P}_1 \cdots \mathcal{P}_n \\ \vdots\ \pi_1 \\ \mathcal{Q}\end{matrix}}{\mathcal{T}}$$

Here $\begin{matrix}\vdots\ \pi \\ \mathcal{Q}\end{matrix}$ means a proof π having $\mathcal{Q}$ as its conclusion. The *length* of a proof is defined as the number of applications of inference rules.

Provability relation of GS is defined as follows:

Definition 2.6 (Provability) A formula $\mathcal{Q}$ is *provable* from $\mathcal{P}_1, \ldots, \mathcal{P}_n$, written as $\mathcal{P}_1, \ldots, \mathcal{P}_n \vdash \mathcal{Q}$, if there is a proof of $\mathcal{Q}$ in GS from $\mathcal{P}_1, \ldots, \mathcal{P}_m$ for $1 \leq m \leq n$.

Given that $(+)$ and $(-)$ rules are generalizations of $\wedge$-introduction and $\wedge$-elimination rules of Gentzen's natural deduction, it is possible to give a rewriting procedure of a GS-proof into an appropriate form of a *normal* GS-*proof*, in a similar way as natural deduction.

Definition 2.7 (Redex and normal GS-proof) A *redex* is a part of a GS-proof which has the form on the left below, and it is rewritten to the form on the right:

$$\dfrac{\dfrac{\mathcal{P}_1 \quad \mathcal{P}_2}{\mathcal{P}_1 \cup \mathcal{P}_2}\,(+)}{\mathcal{Q}}\,(-) \quad \rhd \quad \begin{cases} \dfrac{\mathcal{P}_i}{\mathcal{Q}}\,(-) & \text{when } \mathcal{Q} \subseteq \mathcal{P}_i \text{ for } i = 1 \text{ or } 2 \\[2em] \dfrac{\dfrac{\mathcal{P}_1}{\mathcal{Q}_1}\,(-) \quad \dfrac{\mathcal{P}_2}{\mathcal{Q}_2}\,(-)}{\mathcal{Q}}\,(+) & \begin{array}{c}\text{when } \mathcal{Q} = \mathcal{Q}_1 \cup \mathcal{Q}_2 \text{ such that} \\ \mathcal{Q}_1 \subseteq \mathcal{P}_1 \text{ and } \mathcal{Q}_2 \subseteq \mathcal{P}_2\end{array} \end{cases}$$

Particularly when $\mathcal{Q} = \mathcal{P}_i$ (resp. $\mathcal{Q}_i = \mathcal{P}_i$), the $(-)$ rule on the right above is not applied. A GS-proof π is in *normal form* when it contains no redex.

Note that, as it is in usual natural deduction, the complexity of a proof (the number of formulas appeared in it) is strictly reduced by applying a reduction rule. Thus in a similar way as well-known normalization for natural deduction systems, we obtain the following normalization theorem of GS.

Theorem 2.8 (Normalization of GS) *If $\mathcal{P}_1, \ldots, \mathcal{P}_n \vdash \mathcal{Q}$ in GS, then $\mathcal{P}_1, \ldots, \mathcal{P}_n \vdash \mathcal{Q}$ with a normal GS-proof.*

As a corollary of Theorem 2.8, it is shown that, when a formula of the form $s \sqsubset t$ or $s \vDash t$ is provable, it is provable without using $(-)$ rule or $(+)$ rule.

■

Soundness of GS is easily proved by induction on the length of GS-proof. Completeness of GS is proved in a similar way as [11]. [8]

Theorem 2.9 (Soundness and completeness) *Let $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be GS-formulas such that for some model M, $M \models \mathcal{P}_i$ for all $1 \leq i \leq n$. Then, $\mathcal{P}_1, \ldots, \mathcal{P}_n \vdash \mathcal{Q}$, if and only if, $\mathcal{P}_1, \ldots, \mathcal{P}_n \models \mathcal{Q}$.*

3 GS and categorical syllogisms

In this section, we study a relationship between GS and categorical syllogisms. We show the correspondence between the normal GS-proofs in the syllogistic fragment and the chains of categorical syllogisms.

We consider syllogisms composed of categorical sentences of the forms All A are B, No A are B, Some A are B, and Some A are not B, where we require that A and B are distinct terms. Aristotle and his modern followers showed that all valid syllogisms are reduced to the syllogisms of the first figure, called the *perfect* syllogisms. Instead, for simplicity, here we treat all the valid patterns of syllogisms as primitive inference rules (see Appendix for the list of them), and consider a *chain* of categorical syllogisms (i.e. sorites), defined as follows.[9]

[8]Note that we impose a model existence condition for premises. In place of the semantic restriction, it is possible to extend GS by adding an inference rule corresponding to the absurdity rule of Gentzen's natural deduction system, namely, the rule which permits to infer any formula from a pair of inconsistent formulas.

[9]The validity of some patterns depends on the existential import of the subject term of universal sentences, which licenses to derive Some A are B from All A are B. The status of such an existential import has been much debated in the philosophy and linguistics literature. For simplicity, we exclude these patterns and only consider syllogisms which are valid without existential imports.

Definition 3.1 (Chain of syllogism) A chain of syllogisms of S from premises $S_1, \ldots, S_n$, written by $S_1, \ldots, S_n \vdash S$, is defined inductively as follows:

1. Each syllogism of S with two premises S_1 and S_2 of the following form is a chain of syllogism $S_1, S_2 \vdash S$:
$$\frac{S_1 \quad S_2}{S}$$
2. Let π_1 be a chain of syllogisms $S_1, \ldots, S_i \vdash T$, and π_2 be a chain of syllogisms $S'_1, \ldots, S'_j \vdash T'$. If S is obtained by a syllogism with two premises T and T', then the following is a chain of syllogisms $S_1, \ldots, S_i, S'_1, \ldots, S'_j \vdash S$:
$$\frac{T \quad T'}{S}$$

Definition 3.2 (Syllogistic formula) Categorical sentences are translated into GS-formulas, called *syllogistic formulas*, as follows:

All A are B	is	$\{A \sqsubset B\}$
No A are B	is	$\{A \vdash\!\dashv B\}$
Some A are B	is	$\{c \sqsubset A, c \sqsubset B\}$ for some singular term c
Some A are not B	is	$\{c \sqsubset A, c \vdash\!\dashv B\}$ for some singular term c

We sometimes call a formula of the form $\{A \sqsubset B\}$ or $\{A \vdash\!\dashv B\}$ a *universal* formula, and a formula of the form $\{c \sqsubset A, c \sqsubset B\}$ or $\{c \sqsubset A, c \vdash\!\dashv B\}$ an *existential* formula.

By our convention, two sentences No A are B and No B are A (resp. Some A are B and Some B are A) are interpreted by the same GS-formula. We denote by S° the interpretation of a categorical sentence S. When we translate syllogisms to GS-proofs, we need to assign different singular terms to different existential premises. However, we do not have to care about it when we consider a chain of valid syllogisms, since as is well known, at most one existential sentence can appear in the premises.

Proposition 3.3 (Chains of syllogisms as GS-proofs) *If $S_1, \ldots, S_n \vdash S$ with syllogisms, then $S_1^\circ, \ldots, S_n^\circ \vdash S^\circ$ in* GS.

The valid syllogisms are translated in the following way (see Appendix for the labels of syllogisms used here).

(1) Barbara is translated as $\dfrac{A \sqsubset B \quad B \sqsubset C}{A \sqsubset C}\,(\sqsubset)$

(2) Celarent, Cesare and Camestres are translated as $\dfrac{B \vdash\!\dashv C \quad A \sqsubset B}{A \vdash\!\dashv C}\,(\vdash\!\dashv)$

(3) Darii, Disamis, Datisi, and Dimaris are translated as in the left below.

(4) Ferio, Festino, Ferison, and Fresison are translated as in the right below.

$$\dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset A}\,(-) \quad \dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset B}\,(-) \quad B \sqsubset C}{a \sqsubset C}\,(\sqsubset)}{a \sqsubset A, a \sqsubset C}\,(+) \qquad \dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset A}\,(-) \quad \dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset B}\,(-) \quad B \vdash\!\dashv C}{a \vdash\!\dashv C}\,(\vdash\!\dashv)}{a \sqsubset A, a \vdash\!\dashv C}\,(+)$$

(5) Baroco is translated as in the left below.

(6) Bocardo is translated as in the right below.

$$\dfrac{\dfrac{a \sqsubset A, a \mathrel{\dashv\vdash} B}{a \sqsubset A}(-) \qquad \dfrac{\dfrac{a \sqsubset A, a \mathrel{\dashv\vdash} B}{a \mathrel{\dashv\vdash} B}(-) \quad C \sqsubset B}{a \mathrel{\dashv\vdash} C}(\dashv\vdash)}{a \sqsubset A, a \mathrel{\dashv\vdash} C}(+) \qquad \dfrac{\dfrac{a \sqsubset B, a \mathrel{\dashv\vdash} C}{a \mathrel{\dashv\vdash} C}(-) \qquad \dfrac{\dfrac{a \sqsubset B, a \mathrel{\dashv\vdash} C}{a \sqsubset B}(-) \quad B \sqsubset A}{a \sqsubset A}(\sqsubset)}{a \sqsubset A, a \mathrel{\dashv\vdash} C}(+)$$

Type (1) and (2) correspond to a single rule of GS, while the other four types, which contain an existential sentence in premises, are translated by combinations of some inference rules of GS. Note that type (5) (resp. type (6)) has exactly the same proof structure as type (4) (resp. type (3)). In this sense, the categorical syllogisms are classified in GS into four patterns.

Example 3.4 A chain of syllogisms is simulated by a combination of these patterns. For example, a chain of syllogisms

$$\dfrac{\dfrac{\text{Some } A \text{ are } B \quad \text{All } A \text{ are } C}{\text{Some } C \text{ are } B}\text{Disamis} \quad \text{No } B \text{ are } D}{\text{Some } C \text{ are not } D}\text{Ferio}$$

is simulated in GS as follows.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset A}(-) \quad A \sqsubset C}{a \sqsubset C}(\sqsubset) \quad \dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset B}(-)}{a \sqsubset C, a \sqsubset B}(+)}{a \sqsubset C}(-) \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset A}(-) \quad A \sqsubset C}{a \sqsubset C}(\sqsubset) \quad \dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset B}(-)}{a \sqsubset C, a \sqsubset B}(+)}{a \sqsubset B}(-) \quad B \mathrel{\dashv\vdash} D}{a \mathrel{\dashv\vdash} D}(\dashv\vdash)}{a \sqsubset C, a \mathrel{\dashv\vdash} D}(+)$$

This proof transforms to the following normal proof in GS.

$$\dfrac{\dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset A}(-) \quad A \sqsubset C}{a \sqsubset C}(\sqsubset) \qquad \dfrac{\dfrac{a \sqsubset A, a \sqsubset B}{a \sqsubset B}(-) \quad B \mathrel{\dashv\vdash} D}{a \mathrel{\dashv\vdash} D}(\dashv\vdash)}{a \sqsubset C, a \mathrel{\dashv\vdash} D}(+)$$

Next, we show the converse of the above proposition.

Let S° be a syllogistic formula of GS. We denote by $gtm(S^\circ)$ (resp. $stm(S^\circ)$) a set of general (resp. singular) terms which appear in S°. In order to characterize the syllogistic fragment of GS, we introduce the following notion (cf. [3]):

Definition 3.5 (Cycle) A sequence $S_1^\circ, \ldots, S_n^\circ, S^\circ$ of syllogistic formulas is a *cycle* if

- $gtm(S_1^\circ) = \{A_1, A_2\}, \ldots, gtm(S_i^\circ) = \{A_i, A_{i+1}\}, \ldots, gtm(S^\circ) = \{A_{n+1}, A_1\}$, where $A_i \not\equiv A_j$ for any $1 \leq i, j \leq n+1$ and
- for any S_i° and S_j° with $i \neq j$, if $c \in stm(S_i^\circ)$ and $d \in stm(S_j^\circ)$, then $c \neq d$.

The following lemmas hold particularly in the syllogistic fragment of GS, which also hold for chains of (linguistic) syllogisms.

Lemma 3.6 *When a sequence $S_1^\circ, \dots, S_n^\circ$ of syllogistic formulas is a cycle, for any $A \in gtm(S_1^\circ \cup \cdots \cup S_n^\circ)$, A appears exactly twice in the sequence.*

Lemma 3.7 *Let π be a normal* GS*-proof of S° from $S_1^\circ, \dots, S_n^\circ$.*

1. *There is a premise S_i° which is existential, if and only if, the conclusion S° is existential.*
2. *There are no distinct singular terms c and d such that c is contained in S_i° and d is contained in S_j° with $i \neq j$.*
3. *All premises S_i° are universal, if and only if, the conclusion S° is universal.*

We show that each normal proof in the syllogistic fragment of GS is translated into a chain of syllogisms.

Proposition 3.8 (GS-proofs as chains of syllogisms) *Let π be a normal* GS*-proof of S° from $S_1^\circ, \dots, S_n^\circ$ such that $S_1^\circ, \dots, S_n^\circ, S^\circ$ is cycle. Then there exists a chain of syllogisms such that $S_1, \dots, S_n \vdash S$.*

Proof (sketch). We divide the following two cases according to whether or not the conclusion S° contains a singular term.
(1) When S° does not contain any singular term, by Lemma 3.7, no premises $S_1^\circ, \dots, S_n^\circ$ contain any singular term. Hence, by Theorem 2.8, only $(\sqsubset)$ and $(\vdash\!\dashv)$ rules appear in π. Since $(\sqsubset)$ rule corresponds to Barbara and $(\vdash\!\dashv)$ rule to Ceralent, Cesare, and Camestres, π is easily translated into a chain of syllogisms.
(2) When S° contains a singular term c, the normal GS-proof π is translated into a chain of syllogisms by using the following derived rules.

$$\frac{c \sqsubset A_1, c\Box A_2 \quad A_1 \sqsubset B}{c \sqsubset B, c\Box A_2}\,(\sqsubset)^+ \qquad \frac{c \sqsubset A_1, c\Box A_2 \quad A_1 \vdash\!\dashv B}{c \vdash\!\dashv B, c\Box A_2}\,(\vdash\!\dashv)^+ \qquad \frac{c \vdash\!\dashv A_1, c\Box A_2 \quad B \sqsubset A_1}{c \vdash\!\dashv B, c\Box A_2}\,(\vdash\!\dashv)^+$$

These rules are slight extensions of $(\sqsubset)$ and $(\vdash\!\dashv)$ rules. For example, the normal proof in Example 3.4 is transformed into the following proof on the left, which corresponds to the original chain of syllogisms on the right.

$$\frac{\dfrac{a \sqsubset A, a \sqsubset B \quad A \sqsubset C}{a \sqsubset C, a \sqsubset B}\,(\sqsubset)^+ \quad B \vdash\!\dashv D}{a \sqsubset C, a \vdash\!\dashv D}\,(\vdash\!\dashv)^+ \qquad \frac{\dfrac{\text{Some } A \text{ are } B \quad \text{All } A \text{ are } C}{\text{Some } C \text{ are } B} \quad \text{No } B \text{ are } D}{\text{Some } C \text{ are not } D}$$

In this way, any normal GS-proof is translated into a chain of syllogisms by using the derived rules above. ■

4 GS and natural deduction

In this section, we show that GS corresponds to the natural deduction system for disjunction-free (i.e., $\wedge$, $\to$, $\neg$) fragment of minimal logic (cf. Prawitz [19]), which is abbreviated as ML. In what follows, for ML-formulas, under the associativity and the commutativity of $\wedge$-connective, we consider $\wedge$-connective as an n-ary connective for an appropriate n. Furthermore, for simplicity, we denote by a sequence (set) $\varphi_1, \dots, \varphi_n$ a conjunction $\varphi_1 \wedge \cdots \wedge \varphi_n$, where we assume all conjuncts are distinct under the idempotency of $\wedge$. We also generalize $\wedge I$ and $\wedge E$ rules of natural deduction for ML to those for the n-ary $\wedge$-connective.

We first give a translation of each GS-formula to a propositional implicational formula.

Definition 4.1 (From GS-formulas to ML-formulas) General terms and singular terms of GS are translated into atoms of ML. Then each atom P of GS is translated into an implicational formula $P^\bullet$ as follows:

$$(s \sqsubset t)^\bullet := s \to t \qquad\qquad (s \vDash t)^\bullet := s \to \neg t$$

A GS-formula $\mathcal{P} = \{P_1, \ldots, P_n\}$ is translated into a conjunction:

$$\mathcal{P}^\bullet := P_1^\bullet, \ldots, P_n^\bullet.$$

Each inference rule of GS is translated as a combination of inference rules of ML.

Definition 4.2 (Translation of GS) Inference rules of GS are translated as follows.

The axiom of the form $\dfrac{}{s \sqsubset s}\ ax$ is translated as $\dfrac{[s]^n}{s \to s} \to I, n$.

($\sqsubset$) rule of the form $\dfrac{s \sqsubset t \quad t \sqsubset u}{s \sqsubset u}$ ($\sqsubset$) is translated as in the left below.

($\vDash$) rule of the form $\dfrac{s \sqsubset t \quad t \vDash u}{s \vDash u}$ ($\vDash$) is translated as in the right below.

$$\dfrac{\dfrac{\dfrac{[s]^n \quad s \to t}{t} \to E \qquad t \to u}{\dfrac{u}{s \to u} \to I, n} \to E}{} \qquad\qquad \dfrac{\dfrac{\dfrac{[s]^n \quad s \to t}{t} \to E \qquad t \to \neg u}{\dfrac{\neg u}{s \to \neg u} \to I, n} \to E}{}$$

(+) rule is the same as n-ary $\wedge I$-rule.

($-$) rule is the same as n-ary $\wedge E$-rule.

Definition 4.2 of the translation of inference rules of GS gives, by induction, a translation of any GS-proof π into a natural deduction proof $\pi^\bullet$. Hence the following theorem is immediate:

Theorem 4.3 (From GS to ML) *Let $\mathcal{P}_1, \ldots, \mathcal{P}_n, \mathcal{Q}$ be GS-formulas. If π is a GS-proof of $\mathcal{Q}$ from $\mathcal{P}_1, \ldots, \mathcal{P}_n$, then $\pi^\bullet$ is a natural deduction proof of $\mathcal{Q}^\bullet$ from $\mathcal{P}_1^\bullet, \ldots, \mathcal{P}_n^\bullet$.*

The converse of Theorem 4.3 is shown in the same way as [12].

Theorem 4.4 (Translation of ML) *Let $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be GS-formulas which have a model. Let $\mathcal{Q}$ be a GS-formula. Any proof of $\mathcal{Q}^\bullet$ from $\mathcal{P}_1^\bullet, \ldots, \mathcal{P}_n^\bullet$ in disjunction-free minimal logic is transformed into a GS-proof of $\mathcal{Q}$ from $\mathcal{P}_1, \ldots, \mathcal{P}_n$.*

References

[1] J. Barwise and J. Etchemendy. Visual information and valid reasoning. In G. Allwein, J. Barwise (eds). *Logical Reasoning with Diagrams*, Oxford University Press, 3–26, 1996.

[2] J. Corcoran. Aristotle's natural deduction system, in *Ancient Logic and its Modern Interpretations*, J. Corcoran, ed., D. Reidel, Dordrecht-Holland, 85–131, 1974.

[3] M. Crabbé. The formal theory of syllogisms, *The Review of Modern Logic*, Volume 9, Number 1-2, 29–52, 2001.

[4] J. D. Gergonne. Essai de dialectique rationelle, *Annales de mathématiques pures et appliqées*, 7, 189–228, 1817.

[5] E. Hammer. *Logic and Visual Information*, CSLI Publications, 1995.

[6] E. Hammer and S.-J. Shin. Euler's visual logic, *History and Philosophy of Logic*, 19, 1-29, 1998.

[7] Jan Łukasiewicz. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*, Second edition, Oxford: Clarendon Press, 1957.

[8] B. MacCartney and C. Manning. Modeling Semantic Containment and Exclusion in Natural Language Inference, *Coling-08*, 2008.

[9] K. Mineshima, M. Okada, Y. Sato, and R. Takemura. Diagrammatic Reasoning System with Euler Circles: Theory and Experiment Design, *Diagrams 08*, LNAI, Vol. 5223, Springer, 188-205, 2008.

[10] K. Mineshima, M. Okada, and R. Takemura. Conservativity for a hierarchy of Euler and Venn reasoning systems, *Proc. of Visual Languages and Logic 2009*, CEUR Series Vol. 510, 37-61, 2009.

[11] K. Mineshima, M. Okada, and R. Takemura. A Diagrammatic Reasoning System with Euler Circles, 2009, submitted to a journal.

[12] K. Mineshima, M. Okada, and R. Takemura. Two Types of Diagrammatic Inference Systems: Natural Deduction Style and Resolution Style, to appear in *Proceedings of Diagrams 2010*, *Lecture notes in computer science*, Springer. An extended version is available at: http://abelard.flet.keio.ac.jp/person/takemura/index.html

[13] Lawrence S. Moss. Completeness Theorems for Syllogistic Fragments, in F. Hamm and S. Kepser (eds.), *Logics for Linguistic Structures*, Mouton de Gruyter, 143–173, 2008.

[14] Lawrence S. Moss. Syllogistic Logic with Complements, Ms. Indiana University.

[15] Lawrence S. Moss. Syllogistic Logics with Verbs, *Journal of Logic and Computation*, to appear.

[16] N. Nishihara, and K. Morita. An extended syllogistic system with conjunctive, disjunctive and complementary terms, and its completeness proof (in Japanese). *Trans. IEICE Japan*, Vol. J71-D, No.4. 693–704, 1988.

[17] N. Nishihara, K. Morita, and S. Iwata. An extended syllogistic system with verbs and proper nouns, and its completeness proof (in Japanese). *Trans. IEICE Japan*, Vol J72-D-II, No.5, 96–111, 1989.

[18] Ian Pratt-Hartmann and Lawrence S. Moss. Logics for the relational syllogistic, *Review of Symbolic Logic*, to appear.

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

[20] V. Sánchez Valencia. *Studies on Natural Logic and Categorial Grammar*, PhD thesis, University of Amsterdam, 1991.

[21] Y. Sato, K. Mineshima, and R. Takemura. The efficacy of Euler and Venn diagrams in deductive reasoning: empirical findings. to appear in *Proceedings of Diagrams 2010*, *Lecture notes in computer science*, Springer.

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

[23] T. Smiley. What is a syllogism?, *Journal of Philosophical Logic*, Vol. 1, 136-154, 1974.

[24] Johan van Benthem. *Essays in logical semantics.* Reidel, Dordrecht, 1986.

[25] Johan van Benthem. A brief history of natural logic, Technical Report PP-2008-05, Institute for Logic, Language and Computation, 2008.

[26] Jan van Eijck. Natural logic for natural language, *Lecture Notes in Computer Science*, Volume 4363, Springer, 216-230, 2007.

Appendix. List of Cateorical Syllogisms

The following is the list of the categorical syllogisms and their labels used in this paper (in particular, in Section 3).

1st Figure

$$\frac{\text{All } B \text{ are } C \quad \text{All } A \text{ are } B}{\text{All } A \text{ are } C}\ \text{Barbara}$$

$$\frac{\text{No } B \text{ are } C \quad \text{All } A \text{ are } B}{\text{No } A \text{ are } C}\ \text{Celarent}$$

$$\frac{\text{All } B \text{ are } C \quad \text{Some } A \text{ are } B}{\text{Some } A \text{ are } C}\ \text{Darii}$$

$$\frac{\text{No } B \text{ are } C \quad \text{Some } A \text{ are } B}{\text{Some } A \text{ are not } C}\ \text{Ferio}$$

2nd Figure

$$\frac{\text{No } C \text{ are } B \quad \text{All } A \text{ are } B}{\text{No } A \text{ are } C}\ \text{Cesare}$$

$$\frac{\text{All } C \text{ are } B \quad \text{No } A \text{ are } B}{\text{No } A \text{ are } C}\ \text{Camestres}$$

$$\frac{\text{No } C \text{ are } B \quad \text{Some } A \text{ are } B}{\text{Some } A \text{ are not } C}\ \text{Festino}$$

$$\frac{\text{All } C \text{ are } B \quad \text{Some } A \text{ are not } B}{\text{Some } A \text{ are not } C}\ \text{Baroco}$$

3rd Figure

$$\frac{\text{Some } B \text{ are } C \quad \text{All } B \text{ are } A}{\text{Some } A \text{ are } C}\ \text{Disamis}$$

$$\frac{\text{All } B \text{ are } C \quad \text{Some } B \text{ are } A}{\text{Some } A \text{ are } C}\ \text{Datisi}$$

$$\frac{\text{Some } B \text{ are not } C \quad \text{All } B \text{ are } A}{\text{Some } A \text{ are not } C}\ \text{Bocardo}$$

$$\frac{\text{No } B \text{ are } C \quad \text{Some } B \text{ are } A}{\text{Some } A \text{ are not } C}\ \text{Ferison}$$

4th Figure

$$\frac{\text{All } C \text{ are } B \quad \text{No } B \text{ are } A}{\text{No } A \text{ are } C}\ \text{Camenes}$$

$$\frac{\text{Some } C \text{ are } B \quad \text{All } B \text{ are } A}{\text{Some } A \text{ are } C}\ \text{Dimaris}$$

$$\frac{\text{No } C \text{ are } B \quad \text{Some } B \text{ are } A}{\text{Some } A \text{ are not } C}\ \text{Fresison}$$