Linear set theory with strict comprehension

Masaru Shirahata*
School of Information Science
Japan Advanced Institute of Science and Technology
sirahata@jaist.ac.jp

December 2, 1995

Abstract

In this paper, we study the extensionality axiom in the set theory with the unrestricted comprehension based on linear logic. We first review Grishin's result which shows the imcompatibility of the extensionality axiom and the unrestricted comprehesion in linear set theory. As one way to rectify this situation, we introduce the notion of "strict comprehension" and formulate a system of linear set theory which contains the extensionality and the strict comprehesion. The consistency of such a system is then proved by a simple cut-elimination argument.

1 Introduction

The set theory based on contraction-free logics [1, 5] has been studied for the reason that the use of unrestricted comprehension does not necessarily yield the contradiction in such logics [2, 4, 6, 7, 9]. However, it has been also noticed that the standard extetensionality axiom cannot be added to such a set theory without causing inconsistency [3, 10]. In this paper, we show one way to remedy this situation, using the notion of "strict comprehension."

The basic idea is as follows. Contraction and weakening of a formula in a sequent calculus correspond to *copying* and *discarding* resources, or proofs of that formula (terms of that type). In linear logic [1], those structural rules are not available and, as a result, we can dischage (abstract) exactly one occurrence of the assumption A (variable $x : A$) in forming the implication $A \multimap B$ ($(\lambda x : A)\ M$). We apply the similar restriction to the formation of set terms $\{x : A\}$ so that we have the logical equivalence between the formulas $A[s/x]$ and $s \in \{x : A\}$ only when the variable x has exactly one occurrence in the formula A. Such a comprehension rule will be called *strict (or linear) comprehension.*

We first review Grishin's result which shows the imcompatibility of the extensionality axiom and the unrestricted comprehesion in linear set theory. We then introduce the notion of strict comprehension and formulate a system of linear set theory which contains the extensionality and the strict comprehesion. Finally, the consistency of such a system is proved by a simple cut-elimination argument.

*JSPS research fellow. This work is supported by JSPS research fellowship for young scientists.

2 The extensionality and linear set theory

The extensionality principle says that two sets are identical if they have the same members. This is a very natural principle, given our intuitive notion of sets. However, it turned out that the set theory with the unrestricted comprehension based on affine (BCK) logic becomes inconsistent with the extensionality added.

The proof is due to Grishin [3]. We assume the standard axioms for the equality:

$$\vdash s = s \qquad\qquad s = t, A \vdash A[t/s]$$

where the formula $A[t/s]$ is obtained from the formula A by replacing the term t for some of the occurrences of the term s. We work in the set theory with the unrestricted comprehension

$$\frac{\Gamma, A[s/x] \vdash \Delta}{\Gamma, s \in \{x : A\} \vdash \Delta} \qquad \frac{\Gamma \vdash A[s/x], \Delta}{\Gamma \vdash s \in \{x : A\}, \Delta}$$

which is based on affine logic, *i.e.* the logic without contraction.

First, let's recall that the left contraction rule for the formula A is equivalent to the axiom $A \vdash A \otimes A$, which may be called *duplication* or *copying*. Then, note that we can easily recover the contraction rule for the equations $s = t$ as follows:

$$\dfrac{\dfrac{\vdash s = s \quad \vdash s = s}{\vdash s = s \otimes s = s} \qquad s = t, s = s \otimes s = s \vdash s = t \otimes s = t}{s = t \vdash s = t \otimes s = t}$$

Secondly, we define the terms $s \equiv \{x : x = q\}$ and $t \equiv \{x : x = q \otimes A\}$ for the formula A and show that A is logically equivalent to the equation $s = t$. For this, we use the extensionality written as the rule

$$\frac{\Gamma, x \in s \vdash x \in t, \Delta \quad \Gamma, x \in t \vdash x \in s, \Delta}{\Gamma \vdash s = t, \Delta}$$

where the variable x has no free occurrence in the multi-sets Γ and Δ. We state the proof of the logical equivalence in the informal fashion, form which one can easily recover the derivation in the sequent calculus.

Let's assume $x \in t$. We unwind the term and obtain $x = q \otimes A$. Discarding (left weakening) A, we have $x = q$ and hence, $x \in s$. This amounts to the derivation of the sequent $x \in t \vdash x \in s$.

On the other hand, let's assume A and $x \in s$. We can immediately obtain $x = q \otimes A$. Winding this up, we have $x \in t$. Hence, we have the derivation of $A, x \in s \vdash x \in t$. Then, combining the two directions by the weakening and extensionality, we have $A \vdash s = t$.

For the other direction, assume $s = t$. Since $q = q$ always holds, we have $q \in s$. Hence, $q \in t$. Unwinding this, we obtain $q = q \otimes A$. Discarding $q = q$, we have A. Hence, $s = t \vdash A$ is derivable.

Since the contraction is admissible for $s = t$, it is admissible for the formula A as well. Therefore, we can recover the contraction for any formula. This implies that one can again derive the set-theoretic paradoxes.

Note that there are two critical points in the above proof:

- One of the standard axioms for the equality allows the substitution of the multiple occurrences of one term by another term, and makes it possible to recover the contraction for equations. This seems to be against the spirit of linear logic, which is supposed to be resource conscious.

- The weakening is used in establishing the logical equivalence between $s = t$ and A. It is not clear if one can do the same in the logic without weakening, *i.e.* in linear logic.

In this paper, we pursue the way to modify the system with respect to the first point, which leads to the notion of "strict comprehension." The second point should also be investigated, but we postpone it for another occasion.

3 The system STCOM of linear set theory

Our guiding idea is to restrict substitution $A[t/s]$ under the equation $s = t$ so that exactly one occurrence of s is replaced by t. To achieve this effect, however, we need to modify the comprehension principle as well. To see the point, let's consider the formula $A \equiv q \in s \otimes s \in p$. Under the assumption $s = t$, we can perform the substitution to obtain $q \in t \otimes s \in p$, but we should not be able to obtain $q \in t \otimes t \in p$. However, A is logically equivalent to $s \in \{x : q \in x \otimes x \in p\}$ by the comprehension, and one can use

$$s = t, s \in \{x : q \in x \otimes x \in p\} \vdash t \in \{x : q \in x \otimes x \in p\}$$

to derive

$$s = t, q \in s \otimes s \in p \vdash q \in t \otimes t \in p.$$

We therefore restrict the comprehension rule so that we can form the formula $s \in \{x : A\}$ from the formula $A[s/x]$ if and only if the variable x occurs exactly once in A. This type of comprehension will be called *strict comprehension*.

Definition 1 *The terms, formulas and free variables in them are defined by simultanous induction:*

1. *The variables, denoted x, y, z ..., are terms. They are free in themselves.*
2. *If s and t are terms, then $s \in t$ and $s = t$ are (atomic) formulas. They inherit the free variables from the terms.*
3. *If A is a formula with exactly one free occurrence of the variable x, then $\{x : A\}$ is a term. This term inherits the free variables of A except that x is no longer free.*
4. *If A and B are formulas, then $A \otimes B$, $A \parr B$ and $A^{\perp}$ are formulas. They inherit the free variables from A and B.*

We write $s \notin t$ and $s \neq t$ for $(s \in t)^{\perp}$ and $(s = t)^{\perp}$, respectively.

In the formulation of the deductive system, we use the Gentzen style two-sided sequent calculus. We write $e[t/s]_i$ for the expression obtained from the expression e by replacing the term t for the i-th occurrence of the term s, assuming that e contains such an occurrence. However, we suppress the subscript i as much as possible to simplify the notation.

Axioms:

$$s \in t \vdash s \in t \qquad s = t, p \in q \vdash (p \in q)[t/s] \qquad s = t, p \in q \vdash (p \in q)[s/t]$$

Rules of inference:

$$\frac{\Gamma, A \vdash \Delta}{\Gamma \vdash A^{\perp}, \Delta} \qquad \frac{\Gamma \vdash A, \Delta}{\Gamma, A^{\perp} \vdash \Delta}$$

$$\frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \otimes B \vdash \Delta} \qquad \frac{\Gamma \vdash A, \Delta \quad \Theta \vdash B, \Sigma}{\Gamma, \Theta \vdash A \otimes B, \Delta, \Sigma}$$

$$\frac{\Gamma, A \vdash \Delta \quad \Theta, B \vdash \Sigma}{\Gamma, \Theta, A \parr B \vdash \Delta, \Sigma} \qquad \frac{\Gamma \vdash A, B, \Delta}{\Gamma \vdash A \parr B, \Delta}$$

$$\frac{\Gamma, A[s/x] \vdash \Delta}{\Gamma, \forall x A \vdash \Delta} \qquad \frac{\Gamma \vdash A, \Delta}{\Gamma \vdash \forall x A, \Delta}$$

where the variable x has no free occurrence in Γ or Δ in the right.

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, \exists x A \vdash \Delta} \qquad \frac{\Gamma \vdash A[s/x], \Delta}{\Gamma \vdash \exists x A, \Delta}$$

where the variable x has no free occurrence in Γ or Δ in the left.

$$\frac{\Gamma, A[s/x] \vdash \Delta}{\Gamma, s \in \{x : A\} \vdash \Delta} \qquad \frac{\Gamma \vdash A[s/x], \Delta}{\Gamma \vdash s \in \{x : A\}, \Delta}$$

where the variable x freely occurs exactly once in the formula A.

$$\frac{\Gamma, x \in s \vdash x \in t, \Delta \quad \Gamma, x \in t \vdash x \in s, \Delta}{\Gamma \vdash s = t, \Delta}$$

where the variable x has no free occurrence in Γ or Δ.

$$\frac{\Gamma \vdash A, \Delta \quad \Theta, A \vdash \Sigma}{\Gamma, \Theta \vdash \Delta, \Sigma}$$

As optional rules of inference, we may add to our system the weakenings:

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta} \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash A, \Delta}$$

Proposition 2 *The sequents*

1. $\vdash s = s$
2. $s = t \vdash t = s$
3. $s = t, p = q \vdash (p = q)[t/s]$

are derivable.

Proof
We have the derivations:

$$\frac{x \in s \vdash x \in s \quad x \in s \vdash x \in s}{\vdash s = s} \qquad \frac{s = t, x \in t \vdash x \in s \quad s = t, x \in s \vdash x \in t}{s = t \vdash t = s}$$

$$\frac{\dfrac{p = q, x \in p \vdash x \in q \quad s = t, x \in q \vdash x \in q[t/s]}{s = t, p = q, x \in p \vdash x \in q[t/s]} \quad \dfrac{s = t, x \in q[t/s] \vdash x \in q \quad p = q, x \in q \vdash x \in p}{s = t, p = q, x \in q[t/s] \vdash x \in p}}{s = t, p = q \vdash p = q[t/s]}$$

$$\frac{\dfrac{s = t, x \in p[t/s] \vdash x \in p \quad p = q, x \in p \vdash x \in q}{s = t, p = q, x \in p[t/s] \vdash x \in q} \quad \dfrac{p = q, x \in q \vdash x \in p \quad s = t, x \in p \vdash x \in p[t/s]}{s = t, p = q, x \in q \vdash x \in p[t/s]}}{s = t, p = q \vdash p[t/s] = q}$$

∎

In order to prove the properties of the system, we need to introduce the *size* of proofs.

Definition 3 *The size $\sigma(\pi)$ of a proof π is defined inductively:*

1. *If π is an axiom, then its size is 1;*
2. *If π is obtained from τ by an application of one of the one-premise rules, then $\sigma(\pi)$ is $\sigma(\tau)+1$;*
3. *If π is obtained from π_1 and π_2 by an application of one of the two-premise rules other than the extensionality, then $\sigma(\pi) = \sigma(\pi_1) + \sigma(\pi_2)$;*
4. *If π is obtained from π_1 and π_2 by an application of the extensionality, then $\sigma(\pi) = max(\sigma(\pi_1), \sigma(\pi_2)) + 1$.*

Proposition 4 *Suppose that we have a proof π of the sequent $\Gamma, A \vdash \Delta$ or $\Gamma \vdash A, \Delta$. Then, there exists a proof τ of the sequent $\Gamma, s = t, A[t/s] \vdash \Delta$ or $\Gamma, s = t \vdash A[t/s], \Delta$, respectively, such that the size of τ is at most $\sigma(\pi) + 1$. The same statement holds if we use $A[s/t]$ instead of $A[t/s]$.*

Proof
The proof is by induction on the size of π. When π is obtained from π_1 (and π_2) by an application of one of the inference rules, we call π_1 (and π_2) the immediatel subproof(s) of π. If A is not the main formula of the last inference of π, we apply the inductve hypothesis to the immediate subproof in which A occurs. If A is obtained by one of the weakenings, so is $A[t/s]$. Otherwise, we have three cases.

Case 1 A has the form $B \otimes C$ or $B \parr C$. Then, it suffices to perform the substitution in exactly one of the subformulas B and C since exactly one occurrence of s is replaced by t in $A[t/s]$. Hence, we only need to apply the inductive hypothesis to the appropriate immediate subproof. Similarly for the case where A has the form $B^{\perp}$, $\forall x A$ or $\exists x A$.

Case 2 A has the form $p \in q$. Then, we use the axiom $s = t, (p \in q)[t/s] \vdash p \in q$ or $s = t, p \in q \vdash (p \in q)[t/s]$ and perform the cut. This increases the size of a proof by 1.

Case 3 A has the form $p = q$. If $s = t$ is obtained as the main formula of an appplication of extensionality, then apply the inductive hypothesis to both of the immediate subproofs of π. This increases the sizes of the immediate subproofs at most by 1. By definition, the size of π also increases at most by 1. If the formula $p = q$ is on the lefthand side, then it is from the axiom:

$$p = q, a \in b \vdash (a \in b)[q/p] \qquad \text{or} \qquad p = q, a \in b \vdash (a \in b)[p/q]$$

Without the loss of generality, we can concentrate on the first case. Suppose that the substituton of t for s ocurrs in the term p. We then have

$$\frac{s = t, a \in b \vdash (a \in b)[t/s] \quad p[t/s] = q, (a \in b)[t/s] \vdash (a \in b)[t/s][q/p[t/s]]}{s = t, p[t/s] = q, a \in b \vdash (a \in b)[q/p]}$$

where $(a \in b)[t/s]$ is obtained from $a \in b$ by replacing the term s which is inside the term p in $a \in b$. Hence, we have the occurrence of the term $p[t/s]$ in $(a \in b)[t/s]$, which is to be replaced by q. Therefore, the final result $(a \in b)[t/s][q/p[t/s]]$ is identical to $a \in b$. If the substitution of t for s occurs in the term q, we have

$$\frac{p = q[t/s], a \in b \vdash (a \in b)[q[t/s]/p] \quad s = t, (a \in b)[q[t/s]/p] \vdash (a \in b)[q[t/s]/p][s/t]}{s = t, p = q[t/s], a \in b \vdash (a \in b)[q/p]}$$

where $(a \in b)[q[t/s]/p][s/t]$ is obtained by replacing the term t which is inside the term $q[t/s]$ so that the latter becomes q again. In any on those cases, we only add one application of cut with an axiom. Hence, the size of a proof increases by 1. ∎

We often write $\hat{\pi}$ for the proof obtained in Proposition 4 from π.

4 The cut-elimination theorem

We use three types of reduction rules, *i.e.* 1) permutative reductions, 2) symmetric reductions and 3) axiom reductions. For the full exposition of those rules which do not involve the equality, we refer the reader to [6, 7].

Permutative reductions This type of reductions are applied when one of the cut formulas is not the main formula of the immediate subproof. They are mostly standard and we only give the rules for the extensionality:

$$\dfrac{\Gamma \vdash A, \Delta \quad \dfrac{A, \Theta, x \in s \vdash x \in t, \Sigma \;\; A, \Theta, x \in t \vdash x \in s, \Sigma}{A, \Theta \vdash s = t, \Sigma}}{\Gamma, \Theta \vdash s = t, \Delta, \Sigma} \rightsquigarrow \dfrac{\dfrac{\Gamma \vdash A, \Delta \;\; A, \Theta, y \in s \vdash y \in t, \Sigma}{\Gamma, \Theta, y \in s \vdash y \in t, \Sigma} \quad \dfrac{\Gamma \vdash A, \Delta \;\; A, \Theta, y \in t \vdash y \in s, \Sigma}{\Gamma, \Theta, y \in t \vdash y \in s, \Sigma}}{\Gamma, \Theta \vdash s = t, \Delta, \Sigma}$$

where y is a fresh new variable. The case when the cut-formula is on the righthand side in the extensionality is similarly handled. Since $(max(m, n) + 1) + l = max(m + l, n + l) + 1$, this transformation does not change the size of a proof. In general, one can establish by observation that the permutative reductions preserve the size of a proof. In particular, we can freely permute two consecutive applications of cut without changing the size.

Symmetric reductions This type of reductions are applied when both of the cut-formulas are the main formulas of the respective immediate subproofs. In the presence of weakening, the symmetric reduction for it is applied when one of the cut-formula is the main formula. They are all standard. Hence, we only state two exemplary cases.

$$\dfrac{\dfrac{\Gamma_1 \vdash A, \Delta_1 \quad \Gamma_2 \vdash B, \Delta_2}{\Gamma_1, \Gamma_2 \vdash A \otimes B, \Delta_1, \Delta_2} \quad \dfrac{\Theta, A, B \vdash \Sigma}{\Theta, A \otimes B \vdash \Sigma}}{\Gamma_1, \Gamma_2, \Theta \vdash \Delta_1, \Delta_2, \Sigma} \quad \rightsquigarrow \quad \dfrac{\Gamma_2 \vdash B, \Delta_2 \qquad \dfrac{\Gamma_1 \vdash A, \Delta_1 \quad \Theta, A, B \vdash \Sigma}{\Gamma_1, \Theta, B \vdash \Delta_1, \Sigma}}{\Gamma_1, \Gamma_2, \Theta \vdash \Delta_1, \Delta_2, \Sigma}$$

$$\dfrac{\dfrac{\Gamma \vdash A[s/x], \Delta}{\Gamma \vdash s \in \{x : A\}, \Delta} \quad \dfrac{\Theta, A[s/x] \vdash \Sigma}{\Theta, s \in \{x : A\} \vdash \Sigma}}{\Gamma, \Theta \vdash \Delta, \Sigma} \quad \rightsquigarrow \quad \dfrac{\Gamma \vdash A[s/x], \Delta \quad \Theta, A[s/x] \vdash \Sigma}{\Gamma, \Theta \vdash \Delta, \Sigma}$$

We note that the size of a proof is strictly decreased by an application of any of the symmetric reductions.

Axiom reductions This type of reductions are for the cases where one of the cut-formulas is in one of the axioms. There are some cases of this type of cut for which no reduction is defined. When we substitute the term t for the term s in $p \in s$ to obtain $p \in t$, we call such a substitution *critical*. When there are two reduction rules which only differ in the position (*i.e.* the lefthand side or righthand side of the symbol $\vdash$) of the cut formulas, we only state one of them explicitly.

$$\dfrac{s \in t \vdash s \in t \quad s \in t, \Gamma \vdash \Delta}{s \in t, \Gamma \vdash \Delta} \quad \rightsquigarrow \quad s \in t, \Gamma \vdash \Delta$$

The size of a proof is strictly decreased by this reduction.

$$\dfrac{s = t, (p \in \{x : A\})[s/t] \vdash p \in \{x : A\} \quad \dfrac{A[p/x], \Gamma \vdash \Delta}{p \in \{x : A\}, \Gamma \vdash \Delta}}{s = t, (p \in \{x : A\})[s/t], \Gamma \vdash \Delta} \quad \rightsquigarrow \quad \dfrac{s = t, A[p/x][s/t], \Gamma \vdash \Delta}{s = t, (p \in \{x : A\})[s/t], \Gamma \vdash \Delta}$$

where $\{x : A\}$ is not identical to s, *i.e.* the substitution is not critical, and we replace the right immediate subproof π by the subproof $\hat{\pi}$. Since $\sigma(\hat{\pi}) \leq \sigma(\pi) + 1$, this transformation preserves the size of a proof.

$$\dfrac{\dfrac{\Gamma, x \in s \vdash x \in t, \Delta \quad \Gamma, x \in t \vdash x \in s, \Delta}{\Gamma \vdash s = t, \Delta} \qquad s = t, p \in s \vdash p \in t}{\Gamma, p \in s \vdash p \in t, \Delta} \quad \rightsquigarrow \quad \Gamma, p \in s \vdash p \in t, \Delta$$

where the new proof is obtained by replacing the term p for the variable x in the leftmost subproof. Note that the substitution through the axiom is critical. By this transformation, the size of a proof is strictly decreased.

Each reduction rule gives a one-step reduction. We say that a proof π *reduces to* a proof τ if there is a finite sequence of one-step reductions which transforms π into τ.

In order to prove the (partial) cut-elimination theorem, we need to define what the results of reductions are. For this, we first describe certain configurations of cut in a given proof π. The

subproof τ of π will be called a *chain* if τ is a consecutive application of cut with the axioms for substitution, *i.e.* it has the form

$$\dfrac{s_0 = t_0, p_0 \in q_0 \vdash p_1 \in q_1 \qquad \dfrac{\dfrac{}{s_n = t_n, p_n \in q_n \vdash p_{n+1} \in q_{n+1}}}{\vdots}}{s_1 = t_1, \ldots, s_n = t_n, p_1 \in q_1 \vdash p_{n+1} \in q_{n+1}}$$
$$s_0 = t_0, \ldots, s_n = t_n, p_0 \in q_0 \vdash p_{n+1} \in q_{n+1}$$

where $p_{i+1} \in q_{i+1}$ is obtained from $p_i \in q_i$ by the substitution of s_i (or t_i) for t_i (s_i).

The equations involved in the chain are called the *gates*. They are either *critical* or *non-critical* according to whether the substitutions with them are critical or not. The formulas $p_o \in q_o$ and $p_{n+1} \in q_{n+1}$ are the *end-formulas* of the chain τ. They are called *critical* as well if the substitutions for them are critical.

Now suppose that the chain in a proof π is followed by consecutive applications of cuts such that

1. the cut formulas for each cut are
   (a) one of the non-critical gates and the main formula of a subproof ended with the extensionality, or
   (b) one of the end-formulas, which is critical, and the main formula of a subproof ended with one of the comprehension rules,
2. none of the critical gates and non-critical end-formulas are used as cut-formulas in π.

Such a configuration of cuts will be called a *cluster* in the proof π. To be precise, a configuration consists of a chain and the cut-formulas involved in the cuts following the chain. A cluster α is *closed* if one of the end-formulas of the chain is used as a cut-formula in α. If both of the end-formulas are used as cut-formulas in α, it is *closed at both ends*. If a cluster is not closed, then it is *open*.

Definition 5 *The proof π is called normal if the cuts in π appears only as part of clusters in π.*

Theorem 6 *For any proof π, there exists a normal proof τ such that π reduces to τ and $\sigma(\tau) \leq \sigma(\pi)$.*

Proof
The proof is by induction on the size of a proof. If π does not end with a cut, we apply the inductive hypothesis to the immediate subproof(s) of π. Otherwise, we can assume by the inductive hypothesis that π is obtained from the normal subproofs π_1 and π_2.

Let's suppose that one of the subproofs, π_i, is not an axiom. Furthermore, we suppose that π_i does not end with a cut and the cut-formula of the last cut in π is not the main formula of the last inference rule of π_i. In this case, we can apply the permutative reduction and use the inductive hypothesis. If π_i is an axiom $s \in t \vdash s \in t$, then we use the symmetric reduction. If both of the cut-formulas are the main formulas, then we apply the symmetric reduction with the decrease in size and use the inductive hypothesis. If both of π_1 and π_2 are axioms for substitution, then π itself is a chain.

Next, suppose that one of π_i is an axiom for substitution and the cut-formula A is the main formula in the other subproof. If A has the form $s = t$ and the substitution is critical, then we apply the symmetric reduction with the decrease in size and use the inductive hypothesis. If the substitution is not critical, the configuration already is a cluster. If A has the form $s \in t$ and the substitution is non-critical, then we apply the axiom reduction and use the inductive hypothesis to the immediate subproof of the result. If the substitution is critical, we already hava a cluster.

Finally, consider the case when one of the subproofs, π_i, ends with a cut. Since π_i is normal by the assumption, this cut must be part of a cluster in π_i. If the cut-formula A of the last cut in π does not come from the chain, it is from one of the subproofs, let's say π', which is used in a cut with the chain. We then use the permutative reduction between cuts and move the last cut up to π'. We then use the permutative reduction in order to preserve the last inference rule of π' and apply the inductive hypothesis to the immediate subproof of the result.

If the cut-formula A comes from the chain, we first move the last cut upward by the permutative reduction between cuts so that the cut comes right after the chain. If the other subproof is an axiom for substitution, the last cut is simply extending the chain. Hence, suppose otherwise. If A is the main formula in the other subproof, we further move the cut upward by the permutation to create a cut with an axiom for substitution and use the same arguments as before. In the case when A is $s \in t$, we use the permutation again to recover a chain at the final step. If A is not the main formula in the other subproof, it suffices to consider the case when the last inference rule is a cut and A comes from the chain. We then use the permutation of cuts again which results in adjoining the two chains and the new cluster. ∎

We introduce the ordering among clusters in a given proof π. Suppose that the closed cluster α is formed around the chain τ. By the definition of a cluster, none of the critical gates and non-critical end-formulas in τ are used as cut-formulas in π. However, they may appear as subformulas of formulas $s \in t$ or $s = t$ which are used in cuts within other clusters. If one of the critical gates or non-critical end-formulas of the cluster α appears as a subformula of one of the cut-formulas in the cluster β, we say that β is *lower than* α.

In order to construct a formula which contains a gate or an end-formula A as a subformula, some inference rules other than cut need to be applied to A. Since a cluster is a consecutive application of cuts, such a formula is formed outside and below the cluster itself. Therefore, it is not possible that a gate or an end-formula of the lower cluster appears as a subformula of a cut-formula in the upper cluster. Hence, the ordering thus defined among clusters does not contain a loop. Then, there are clusters which are *minima* under this ordering.

Corollary 7 *There is no proof of the empty sequent.*

Proof
By Theorem 7, it suffices to consider normal proofs. Let π be a normal proof. If π does not contain a cut, then the end sequent of π is not empty. Otherwise, choose one of the clusters, α, which is a *minimum* under the ordering. If the cluster is not closed at both ends, then one of the end-formulas remains as a subformula in the end sequent. If the cluster is closed at both ends, then the chain contains at least one critical gate, which remains as a subformula in the end sequent. ∎

5 Concluding remarks

There are two issues which may be interesting to pursue, *i.e.* semantics and decidability. The models for linear set theory in terms of Kripke semantics and phase semantics have been studied [4, 8]. However, they are not easy to construct except as the term models. In [8], we give a much simpler model in terms of hereditarily finite sets to a system similar to STCOM but without the extensionality. The idea was to interpret different *occurrences* of the same term by different finite sets. The same approach does not immediately work with the extensionality since the formula occurrences in the left and right premisses are identified in the conclusion.

The decidability of many of propositional logics can be established through the subformula or substerm property, even in the presence of cuts. However, this does not immediately apply to STCOM. If we have the cluster

$$\dfrac{\Gamma \vdash \Delta, p \in s \qquad \dfrac{s = t, p \in s \vdash p \in t \quad p \in t, \Theta \vdash \Sigma}{p \in s, \Theta, s = t \vdash \Sigma}}{\Gamma, \Theta, s = t \vdash \Sigma}$$

then the term p may never appear as a subterm of any of the formulas below the cluster.

References

[1] J.Y. Girard. "Linear logic." *Theoretical Computer Science*, 50, 1987, 1-102.

[2] V.N. Grishin. "A nonstandard logic and its application to set theory," (Russian). *Studies in Formalized Languages and Nonclassical Logics (Russian)*, Izdat, "Nauka," Moskow. 1974, 135-171.

[3] V.N. Grishin. "Predicate and set theoretic calculi based on logic without contraction rules," (Russian). *Izvestiya Akademii Nauk SSSR Seriya Matematicheskaya*, 45, no.1, 1981, 47-68. 239. *Math. USSR Izv.*, 18, no.1, 1982, 41-59 (English translation).

[4] Y. Komori. "Illative combinatory logic based on BCK-logic." *Math. Japonica*, 34, No. 4, 1989, 585-596.

[5] H. Ono and Y. Komori. "Logics without the contraction rule." *Journal of Symbolic Logic*, 50, 1985, 169-201.

[6] M. Shirahata. *Linear Set Theory*. Dissertation, Department of Philosophy, Stanford University, 1994.

[7] M. Shirahata. "A linear conservative extension of Zermelo-Fraenkel set theory." *Studia Logica*, to appear.

[8] M. Shirahata. "Two simple models of linear set theory." *manuscript*.

[9] R.B. White. "A demonstrably consistent type-free extension." *Mathematica Japonica*, 32, 1987, 149-169.

[10] R.B. White. "A consistent theory of attributes in a logic without contraction." *Studia Logica*, 52, 1993, 113-142.