logics

MDPI

Article

A Decidable Ground Fragment of the Monotonicity Calculus

Daniel Li

Group in Logic and the Methodology of Science, University of California, Berkeley, CA 94720, USA; li.daniel2001@berkeley.edu

Abstract

Monotonicity is a prevalent feature in natural language. Notably, determiners such as every, some, and no induce monotonic (upward) or antitonic (downward) entailments. The Monotonicity Calculus is a proof system that formalizes such reasoning in natural language inferences through order statements (inequalities) involving higher-order (typed) terms. In this paper, we prove that with a slight modification of the conditions for term formation, it is decidable over ground terms, i.e., terms that contain no variables.

Keywords: monotonicity; decidability; subterm property; natural logic

1. Introduction

Monotonicity is a prevalent feature in natural language. Notably, determiners such as every, some, and no induce monotonic (upward) or antitonic (downward) entailments. For instance, run entails move, and Every man runs entails Every man moves. Meanwhile, man entails person, and Every person runs entails Every man runs. These can be formalized as

$$\frac{\texttt{run} \leq \texttt{move}}{\texttt{every}(\texttt{man})(\textbf{run}) \leq \texttt{every}(\texttt{man})(\textbf{move})} \qquad \frac{\texttt{man} \leq \texttt{person}}{\texttt{every}(\textbf{person})(\texttt{run}) \leq \texttt{every}(\textbf{man})(\texttt{run})}$$

We say that every is monotonic, or order-preserving, in its second argument and antitonic, or order-reversing, in its first argument.

Monotonicity reasoning is a central component in the program of *natural logic*, which aims to formalize fragments of natural language in systems that resemble linguistic surface forms, enabling direct inference without intermediate translation languages such as first-order logic. Special emphasis is placed on decidable systems, motivated by both theoretical concerns and applications. (See [1] for further discussion of the relevance of decidability to natural language semantics.)

This paper is based on the *Monotonicity Calculus* presented by Icard and Moss in [2] and Icard et al. in [3], a system for monotonicity reasoning in a higher-order (typed) setting. In this system, terms have their monotonicity information directly encoded in their types, and the logic consists of order statements (inequalities). It is known that the full Monotonicity Calculus is undecidable [4], but there has been relatively little investigation of its decidable fragments. In this paper, we show that with a slight modification of the definition of well-formed terms in [2,3], the Monotonicity Calculus is decidable over ground terms, i.e., terms that contain no variables. The focus on this ground fragment is motivated by Kozen's work on equational logic, which, similarly, is undecidable in general, but decidable over ground terms [5].

Fyodorov et al. have done similar work [6], but our presentation is different, and we also strengthen their result in the fragment of interest by removing an assumption they

Check for updates

Academic Editors: Valentin Goranko and Wesley H. Holliday

Received: 10 August 2025
Revised: 27 February 2026
Accepted: 14 May 2026
Published: 4 June 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

make. Our methods are proof-theoretic and rely on establishing the subterm property, to be defined below.

Beyond natural logic, similar inequational logics have been studied in the programming language theory literature [7,8].

2. Background

We present some relevant background information, focusing on syntactic components in [2,3]. We begin with base types $b \in \mathcal{B}$. The set of types is defined inductively by

$$\tau ::= b \mid (\tau \xrightarrow{+} \tau) \mid (\tau \xrightarrow{-} \tau) \mid (\tau \xrightarrow{\cdot} \tau)$$

where $\{+, -, \cdot\}$ is the set of *polarity markings*.

Definition 1. *The* ***valence*** *of a type is defined inductively as follows: valence*$(b) = 0$ *for* $b \in \mathcal{B}$, *and valence*$(\sigma \xrightarrow{m} \tau) =$ *valence*$(\tau) + 1$ *for any* $m \in \{+, -, \cdot\}$.

Each type τ has an associated set of constants $\mathcal{C}_\tau$, and we assume that these sets are pairwise disjoint. In this paper, we will only work with **ground terms**, which are terms that do not contain any variables.

Definition 2. *All constants are ground terms, and for any ground terms* F *with type* $\sigma \xrightarrow{m} \tau$ *and* A *with type* σ', *where* σ' *is the same type as* σ *disregarding polarity markings,* $F(A)$ *is a ground term with type* τ.

Consequently, every ground term has a unique type (including polarity markings).

Remark 1. *In Definition 2, we deviate from [2,3], which has a stricter requirement on the markings in* σ' *for term formation. Instead, we follow [6], which ignores the markings in the type of the argument, though the setup there is slightly different. This seems to be a reasonable stipulation if we are only dealing with the basic types e and t from natural language semantics (and the types that can be formed from them). For example, the type of* `every` *might be given as* $(e \xrightarrow{+} t) \xrightarrow{-} (e \xrightarrow{+} t) \xrightarrow{+} t$, *but it is unproblematic for* `every` *to take an antitonic argument. This is because all functions of type* $e \to t$ *are trivially monotonic (and antitonic), as the natural order on type e (assuming no mereological structure) is the trivial order:* $x \leq y \iff x = y$.

For any type $\sigma \to \tau$, *call* σ *an argument type within* $\sigma \to \tau$. *An alternative to the discussion above is to explicitly specify that argument types have no monotonicity information by redefining our types as*

$$\tau ::= b \mid (\sigma \xrightarrow{+} \tau) \mid (\sigma \xrightarrow{-} \tau) \mid (\sigma \xrightarrow{\cdot} \tau)$$
$$\sigma ::= b \mid (\sigma \to \sigma)$$

Definition 3. *If terms S and T have the same type disregarding their polarity markings, then we can write the* ***inequality*** $S \leq T$. *(We will generally not be explicit about types—we just assume that each complex term written is well-formed and that the terms in each inequality written satisfy the condition in this definition.)*

Definition 4. *If* F *has type* $\sigma \xrightarrow{+} \tau$, *then we say* F *is* ***monotonic*** *(denoted by* $F : +$ *or* F^+*). If* F *has type* $\sigma \xrightarrow{-} \tau$, *then we say* F *is* ***antitonic*** *(denoted by* $F : -$ *or* F^-*). If* F *has type* $\sigma \xrightarrow{\cdot} \tau$, *then it is neither monotonic nor antitonic.*

We now define **subterms**.

https://doi.org/10.3390/logics4020006

Definition 5. *For a constant c, $\mathrm{Subterm}(c) = \{c\}$, and for a term $F(A)$,*

$$\mathrm{Subterm}(F(A)) = \{F(A)\} \cup \mathrm{Subterm}(F) \cup \mathrm{Subterm}(A)$$

(In general, we use lowercase letters for constants and uppercase letters for (complex) terms.)

Given a set Γ of inequalities and terms S, T, the subterm set of (Γ, S, T) is

$$\mathrm{Subterm}(S) \cup \mathrm{Subterm}(T) \cup \bigcup_{(U \leq V) \in \Gamma} \mathrm{Subterm}(U) \cup \mathrm{Subterm}(V)$$

Instead of speaking of *subterms of* (Γ, S, T), we will often refer to them simply as *subterms* when it is understood what Γ, S, T are.

Definition 6. *The Monotonicity Calculus $\mathcal{M}$ (over ground terms) is the proof system with the following rules:*

$$(F:+)\ \frac{A \leq B}{F(A) \leq F(B)}\ \text{(Mono)} \qquad (G:-)\ \frac{A \leq B}{G(B) \leq G(A)}\ \text{(Anti)}$$

$$\frac{F \leq G}{F(A) \leq G(A)}\ \text{(Point)} \qquad \frac{A \leq B \qquad B \leq C}{A \leq C}\ \text{(Trans)} \qquad \frac{}{A \leq A}\ \text{(Refl)}$$

We will view formal derivations as Gentzen-style inverted trees, and we will use standard terminology involving trees. Let $\mathcal{D}$ be a derivation. If the conclusion of $\mathcal{D}$ is $S \leq T$, we say that $\mathcal{D}$ is a *derivation of* $S \leq T$. A *subderivation* within $\mathcal{D}$ is a subtree of $\mathcal{D}$.

Definition 7. *The **size** of a derivation $\mathcal{D}$ is the number of nodes in $\mathcal{D}$.*

In this paper, we prove that given a finite set Γ of ground inequalities and ground terms S, T, it is decidable whether $\Gamma \vdash_{\mathcal{M}} S \leq T$. Incidentally, the fragment we investigate coincides with a fragment of the logic in [6]. Fyodorov et al. provide a decision algorithm for this fragment, but they impose the restriction that derivations cannot have any occurrences of

$$\frac{F_1^+ \leq F_2^- \qquad F_2^- \leq F_3^+}{F_1^+ \leq F_3^+}\ \text{(Trans)} \qquad \frac{F_1^- \leq F_2^+ \qquad F_2^+ \leq F_3^-}{F_1^- \leq F_3^-}\ \text{(Trans)}$$

which they term *mixed monotonicity* (this is a reformulation of Definition 5.3 in [6]), conjecturing that such patterns do not occur in natural language. The main result in this paper extends the work of Fyodorov et al. in the fragment of interest by removing this restriction.

Remark 2. *An example of $F_1^- \leq F_2^+$ is* `every` $\leq$ `some`*, but it is unclear if there are any natural examples of $F_1^+ \leq F_2^-$. We present an argument which seems to cast doubt on this in the case where F_1, F_2 are interpreted as quantifiers.*

Let $Q_1 : (e \to t) \xrightarrow{+} (e \to t) \xrightarrow{\cdot} t$ and $Q_2 : (e \to t) \xrightarrow{-} (e \to t) \xrightarrow{\cdot} t$. Let $\top$ be a top element of type $e \to t$; i.e., $A \leq \top$ for any $A : e \to t$. Fix $B : e \to t$. Let $A, C : e \to t$ be arbitrary. Then,

$$\begin{aligned} Q_1(A)(B) &\leq Q_1(\top)(B) && \textit{(Mono)} \\ &\leq Q_2(\top)(B) && \textit{(Point)} \\ &\leq Q_2(C)(B) && \textit{(Anti)} \end{aligned}$$

This means either $Q_1(A)(B)$ holds for no A or $Q_2(C)(B)$ holds for all C. In other words, either Q_1 or Q_2 is trivial. (Thanks to Josef v. H. Doyle for pointing out this result to me.)

https://doi.org/10.3390/logics4020006

Whether the top element assumption is desirable for natural language applications is up for debate, though one might suggest, for example, that `thing` *or* `entity` *plays the role of the top element among nouns. We also point out that, in any case, we may have to consider other semantic types and domains for relevant examples. One promising direction to examine is monotonicity over the domain of pluralities with the parthood relation as its order.*

The issue of whether it is linguistically innocuous to eliminate mixed monotonicity remains an open question, but regardless, the formalism presented in this paper can also be applied to higher-order mathematical reasoning in other domains. Monotonicity is ubiquitous in logic, mathematics and computer science, and formal properties of monotonicity reasoning are of theoretical interest.

3. A Special Case

To motivate the solution to the full problem, we first consider a restricted system $\mathcal{M}_+$ with only monotonic functions and without the proof rule (Anti). Many of the ideas from this section will carry over to the general case, and we develop the intuition and details necessary for the full solution so that we can just highlight the differences instead of repeating these details later.

Definition 8. *$\mathcal{M}_+$ has the proof rules* (Mono), (Point), (Trans), and (Refl) from $\mathcal{M}$.

Definition 9. *Let Γ be a set of inequalities, and S, T be terms. A* ***subterm derivation*** *of $S \leq T$ from axioms in Γ is a derivation in which every inequality contains only subterms of (Γ, S, T).*

We say that a proof system $\mathcal{P}$ has the *subterm property* if $\Gamma \vdash_{\mathcal{P}} S \leq T$ implies $\Gamma \vdash_{\mathcal{P}} S \leq T$ via a subterm derivation. If we can show that a proof system has the subterm property, then this essentially makes the search space finite for proof search, and decidability follows closely after.

First note that $\mathcal{M}_+$ does not have the subterm property. As a very simple counterexample,

$$\{f \leq g, a \leq b\} \vdash f(a) \leq g(b)$$

but it is not possible to obtain a subterm derivation. The simplest derivations we can find are

$$\dfrac{\dfrac{a \leq b}{f(a) \leq f(b)}\,(\text{Mono}) \quad \dfrac{f \leq g}{f(b) \leq g(b)}\,(\text{Point})}{f(a) \leq g(b)}\,(\text{Trans}) \qquad \dfrac{\dfrac{f \leq g}{f(a) \leq g(a)}\,(\text{Point}) \quad \dfrac{a \leq b}{g(a) \leq g(b)}\,(\text{Mono})}{f(a) \leq g(b)}\,(\text{Trans}) \tag{1}$$

which contain non-subterms $f(b)$ and $g(a)$ respectively.

To that end, consider a new proof system:

Definition 10. *$\mathcal{M}'_+$ has the proof rules:*

$$\dfrac{F \leq G \qquad A \leq B}{F(A) \leq G(B)}\,(\text{App}) \qquad \dfrac{A \leq B \qquad B \leq C}{A \leq C}\,(\text{Trans}) \qquad \dfrac{}{A \leq A}\,(\text{Refl})$$

Essentially, this system resolves the issue above. We claim that $\mathcal{M}'_+$ is equivalent to $\mathcal{M}_+$.

Proposition 1. *$\Gamma \vdash_{\mathcal{M}_+} S \leq T$ if and only if $\Gamma \vdash_{\mathcal{M}'_+} S \leq T$*

https://doi.org/10.3390/logics4020006

Proof. For both proof systems, we show that each of the proof rules can be derived in the other system. For the forward direction, (Mono) and (Point) can be derived, respectively, through

$$\dfrac{\dfrac{}{F \leq F}\,(\text{Refl}) \qquad A \leq B}{F(A) \leq F(B)}\,(\text{App}) \qquad\qquad \dfrac{F \leq G \qquad \dfrac{}{A \leq A}\,(\text{Refl})}{F(A) \leq G(A)}\,(\text{App})$$

For the backwards direction, (App) can be derived by the examples above. □

We now introduce a few technical notions in order to show that $\mathcal{M}'_+$ has the subterm property.

Definition 11. *Let $\mathcal{D}$ be a derivation of $S \leq T$. The* ***bottom-most chain*** *of $\mathcal{D}$ is the maximal sequence $(S = S_1, S_2, \ldots, S_k = T)$ such that $S_1 \leq S_2; \ldots; S_{k-1} \leq S_k$ all appear as premises of (Trans) in $\mathcal{D}$. In particular, for each $i = 1, \ldots, k-1$, the subderivation of $S_i \leq S_{i+1}$ does not have (Trans) as its last step.*

We will also refer to the sequence of inequalities $S_1 \leq S_2; \ldots; S_{k-1} \leq S_k$ as the bottom-most chain.

Example 1. *The bottom-most chain of the following derivation consists of the bolded inequalities.*

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\boldsymbol{S_1 \leq S_2}}\,(App) \quad \dfrac{\vdots}{\boldsymbol{S_2 \leq S_3}}\,(App)}{S_2 \leq S_3}\,(Trans) \quad \dfrac{\vdots}{\boldsymbol{S_3 \leq S_4}}\,(App)}{S_1 \leq S_4}\,(Trans) \quad \dfrac{\dfrac{\vdots}{\boldsymbol{S_4 \leq S_5}}\,(App) \quad \dfrac{\vdots}{\boldsymbol{S_5 \leq S_6}}\,(App)}{S_4 \leq S_6}\,(Trans)}{S_1 \leq S_6}$$

Definition 12. *Let $(S_1, \ldots, S_n)$ be the bottom-most chain of $\mathcal{D}$. We say that $\mathcal{D}$ is* ***left-transitive*** *if it has the form*

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{S_1 \leq S_2} \quad \dfrac{\vdots}{S_2 \leq S_3}}{S_1 \leq S_3}\,(\text{Trans}) \quad \dfrac{\vdots}{S_3 \leq S_4}}{S_1 \leq S_4}\,(\text{Trans}) \quad \ddots \quad \dfrac{\vdots}{S_{n-1} \leq S_n}}{S_1 \leq S_n}\,(\text{Trans})$$

Lemma 1. *Let $\mathcal{D}$ be a derivation of $S \leq T$ from Γ in $\mathcal{M}'_+$. Then $\Gamma \vdash_{\mathcal{M}'_+} S \leq T$ via a left-transitive derivation with the same size as $\mathcal{D}$.*

Proof. By induction on the size of the bottom-most chain. Rearrange the order in which (Trans) is applied. □

Remark 3. *Extending Lemma 1, for any derivation, we can convert the portion below the bottom-most chain to any form, without changing its size, by rearranging the order in which (Trans) is applied.*

We now discuss the main result of this section.

Theorem 1. *If $\Gamma \vdash_{\mathcal{M}'_+} S \leq T$, then $\Gamma \vdash_{\mathcal{M}'_+} S \leq T$ via a subterm derivation.*

Here is an example that illustrates the main idea of the proof.

Example 2. *Let* $\Gamma = \{f_1 \leq f_2,\ f_2 \leq f_3,\ f_3 \leq f_4,\ a_1 \leq a_2,\ a_2 \leq a_3,\ a_3 \leq a_4\}$. *Assume we begin with the following derivation.*

$$\dfrac{\dfrac{\dfrac{f_1 \leq f_2 \qquad a_1 \leq a_2}{f_1(a_1) \leq f_2(a_2)}\,(App) \qquad \dfrac{f_2 \leq f_3 \qquad a_2 \leq a_3}{f_2(a_2) \leq f_3(a_3)}\,(App)}{f_1(a_1) \leq f_3(a_3)}\,(Trans) \qquad \dfrac{f_3 \leq f_4 \qquad a_3 \leq a_4}{f_3(a_3) \leq f_4(a_4)}\,(App)}{f_1(a_1) \leq f_4(a_4)}\,(Trans)$$

Note that the non-subterms here are $f_2(a_2)$ *and* $f_3(a_3)$. *We apply the following transformations to eliminate these non-subterms.*

$$\dfrac{\dfrac{\dfrac{\boldsymbol{f_1 \leq f_2} \qquad \boldsymbol{f_2 \leq f_3}}{\boldsymbol{f_1 \leq f_3}}\,(Trans) \qquad \dfrac{\boldsymbol{a_1 \leq a_2} \qquad \boldsymbol{a_2 \leq a_3}}{\boldsymbol{a_1 \leq a_3}}\,(Trans)}{\boldsymbol{f_1(a_1) \leq f_3(a_3)}\,(App)} \qquad \dfrac{f_3 \leq f_4 \qquad a_3 \leq a_4}{f_3(a_3) \leq f_4(a_4)}\,(App)}{f_1(a_1) \leq f_4(a_4)}\,(Trans)$$

$$\dfrac{\dfrac{\dfrac{f_1 \leq f_2 \qquad f_2 \leq f_3}{\boldsymbol{f_1 \leq f_3}}\,(Trans) \qquad \boldsymbol{f_3 \leq f_4}}{\boldsymbol{f_1 \leq f_4}}\,(Trans) \qquad \dfrac{\dfrac{a_1 \leq a_2 \qquad a_2 \leq a_3}{\boldsymbol{a_1 \leq a_3}}\,(Trans) \qquad \boldsymbol{a_3 \leq a_4}}{\boldsymbol{a_1 \leq a_4}}\,(Trans)}{\boldsymbol{f_1(a_1) \leq f_4(a_4)}\,(App)}$$

The idea is to push applications of (Trans) upward.

In the derivation above, if each f_i was replaced with some complex term F_i and each a_i with some complex term A_i, and moreover, each $F_i \leq F_{i+1}$ and each $A_i \leq A_{i+1}$ were derived rather than being an axiom, we would repeat the process above on $F_1 \leq F_4$ and $A_1 \leq A_4$.

Proof. The structure of the proof is

Induction on derivation size
  Base case $n = 1$
    Induction on k
  Induction on n

where n and k are defined below. The main content, displayed as the bolded step, is in the Claim below, which generalizes Example 2. In the derivations displayed below, any step whose proof rule is omitted is by (Trans).

Fix Γ. The base case of derivation size 1 is trivial. Now assume that any derivation of size $s' < s$ can be converted to a subterm derivation (remember that what constitutes a subterm derivation is relativized to the conclusion of the derivation). Consider a derivation of size s. By Remark 3, we can convert it, without changing its size, to the form

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots \qquad \vdots}{T_0 \leq S_1 \qquad \boxed{S_1 \leq T_1}}}{T_0 \leq T_1} \qquad \vdots}{T_0 \leq S_2} \qquad \vdots}{T_0 \leq T_2} \qquad \begin{array}{c} \\ \\ \ddots \qquad \vdots \\ T_0 \leq S_i \qquad \boxed{S_n \leq T_n} \\ \hline T_0 \leq T_n \qquad T_n \leq S_{n+1} \\ \hline T_0 \leq S_{n+1} \end{array} \tag{2}$$

where the subderivation of each $T_i \leq S_{i+1}$ has only axioms in its bottom-most chain and the subderivation of each $S_i \leq T_i$ (boxed) has no axioms in its bottom-most chain. The subderivations of $T_0 \leq S_1$ and $T_n \leq S_{n+1}$ are possibly empty.

Consider the base case $n = 1$. Then derivation (2) reduces to the derivation of $T_0 \leq S_2$, so it suffices to show that the subderivation $\mathcal{D}$ of $S_1 \leq T_1$ can be converted to a subterm derivation. (The subderivations of $T_0 \leq S_1$ and $T_1 \leq S_2$ contain only axioms, so they do not need to be modified, and moreover, this means that S_1 and T_1 are subterms of (Γ, S, T).) We know that the last step of the subderivation of each $U_i \leq U_{i+1}$ in the bottom-most chain $(S_1 = U_1, \ldots, U_k = T_1)$ of $\mathcal{D}$ is not (Trans). We also have that each $U_i \leq U_{i+1}$ is not an axiom by definition of (2). Thus, the last step of the derivation of $U_i \leq U_{i+1}$ is either (App) or (Refl). If it is (Refl), we can simply remove $U_i \leq U_{i+1}$ and the immediate application of (Trans) from the derivation. Let $\mathcal{D}'$ be the new derivation with all such $U_i \leq U_{i+1}$ removed. Now, we can convert $\mathcal{D}'$ to a left-transitive derivation with the same size, so it will be of the form

$$\dfrac{\dfrac{\dfrac{\vdots \quad \vdots}{F_1 \leq F_2 \qquad A_1 \leq A_2}}{F_1(A_1) \leq F_2(A_2)}\,(\text{App}) \qquad \dfrac{\dfrac{\vdots \quad \vdots}{F_2 \leq F_3 \qquad A_2 \leq A_3}}{F_2(A_2) \leq F_3(A_3)}\,(\text{App})}{F_1(A_1) \leq F_3(A_3)} \qquad \cdots \qquad \dfrac{\dfrac{\vdots \quad \vdots}{F_{k-1} \leq F_k \qquad A_{k-1} \leq A_k}}{F_{k-1}(A_{k-1}) \leq F_k(A_k)}\,(\text{App})}{S_1 = F_1(A_1) \leq F_k(A_k) = T_1} \tag{3}$$

where $U_i = F_i(A_i)$.

We now prove the following claim: *For any k, if $F_1(A_1) \leq F_k(A_k)$ has a derivation of form (3), then it also has a derivation with the same size which has the form (4) below.*

$$\dfrac{\dfrac{\dfrac{\vdots \quad \vdots}{F_1 \leq F_2 \qquad F_2 \leq F_3}}{F_1 \leq F_3} \qquad \cdots \qquad \dfrac{\dfrac{\vdots \quad \vdots}{A_1 \leq A_2 \qquad A_2 \leq A_3}}{A_1 \leq A_3}}{F_1 \leq F_k} \qquad \cdots \qquad \dfrac{\dfrac{\vdots \quad \vdots}{A_1 \leq A_{k-1} \qquad A_{n-1} \leq A_k}}{A_1 \leq A_k}\,(\text{App})}{F_1(A_1) \leq F_k(A_k)} \tag{4}$$

The base cases $k = 1, 2$ are trivial. For the induction, assume the claim for $k - 1$. Now consider derivation (3). $F_1(A_1) \leq F_k(A_k)$ has a derivation of the following form with the same size, by applying the inductive hypothesis on $F_1(A_1) \leq F_{k-1}(A_{k-1})$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots \quad \vdots}{F_1 \leq F_2 \qquad F_2 \leq F_3}}{F_1 \leq F_3} \qquad \cdots \qquad \dfrac{\dfrac{\vdots \quad \vdots}{F_{k-2} \leq F_{k-1}}}{F_1 \leq F_{k-1}} \qquad \dfrac{\dfrac{\vdots \quad \vdots}{A_1 \leq A_2 \qquad A_2 \leq A_3}}{A_1 \leq A_3} \qquad \cdots \qquad \dfrac{\dfrac{\vdots \quad \vdots}{A_{k-2} \leq A_{k-1}}}{A_1 \leq A_{k-1}}\,(\text{App})}{F_1(A_1) \leq F_{k-1}(A_{k-1})} \qquad \dfrac{\dfrac{\vdots \quad \vdots}{F_{k-1} \leq F_k \qquad A_{k-1} \leq A_k}}{F_{k-1}(A_{k-1}) \leq F_k(A_k)}\,(\text{App})}{F_1(A_1) \leq F_k(A_k)}$$

Narrowing in on the bottom of the derivation, we have

$$\dfrac{\dfrac{\dfrac{\ddots}{F_1 \leq F_{k-1}}\,(\text{Trans}) \qquad \dfrac{\ddots}{A_1 \leq A_{k-1}}\,(\text{Trans})}{F_1(A_1) \leq F_{k-1}(A_{k-1})}\,(\text{App}) \qquad \dfrac{\dfrac{\vdots \quad \vdots}{F_{k-1} \leq F_k \qquad A_{k-1} \leq A_k}}{F_{k-1}(A_{k-1}) \leq F_k(A_k)}\,(\text{App})}{F_1(A_1) \leq F_k(A_k)}\,(\text{Trans})$$

and we can manipulate this to get

$$\dfrac{\dfrac{\ddots}{F_1 \leq F_{k-1}}\,(\text{Trans}) \qquad \dfrac{\vdots}{F_{k-1} \leq F_k}\,(\text{Trans}) \qquad \dfrac{\dfrac{\ddots}{A_1 \leq A_{k-1}}\,(\text{Trans}) \qquad \dfrac{\vdots}{A_{k-1} \leq A_k}\,(\text{Trans})}{A_1 \leq A_k}\,(\text{App})}{F_1(A_1) \leq F_k(A_k)}$$

which we can see to be of the desired form, after filling in the omissions from above. Also observe that the original subderivations of $F_1 \leq F_{k-1}; F_{k-1} \leq F_k; A_1 \leq A_{k-1}; A_{k-1} \leq A_k$ are preserved, so the new derivation of $F_1(A_1) \leq F_k(A_k)$ has the same size as the original. This completes the proof of the claim.

Now we can obtain a derivation $\mathcal{D}''$ of $F_1(A_1) \leq F_k(A_k)$ with the same size as $\mathcal{D}'$:

$$\dfrac{\dfrac{\ddots}{F_1 \leq F_k}\,(\text{Trans}) \qquad \dfrac{\ddots}{A_1 \leq A_k}\,(\text{Trans})}{F_1(A_1) \leq F_k(A_k)}\,(\text{App})$$

Since the subderivations of $F_1 \leq F_k$ and $A_1 \leq A_k$ have size less than that of $\mathcal{D}''$, and $\text{size}(\mathcal{D}'') = \text{size}(\mathcal{D}') \leq \text{size}(\mathcal{D}) \leq s$, by the inductive hypothesis on derivation size, we can obtain subterm derivations for both of them. Furthermore, F_1, F_k, A_1, A_k are all subterms of $F_1(A_1)$ or $F_k(A_k)$, so this also gives us a subterm derivation of $F_1(A_1) \leq F_k(A_k)$, which was originally labeled as $S_1 \leq T_1$.

The result follows by induction on n, by replicating the reasoning above. (In other words, for each $i = 1, \ldots, n$, we can perform the conversion described above on $S_i \leq T_i$ in (2).) □

Corollary 1. *It is decidable, given a finite set* Γ *of ground inequalities and ground terms* S, T*, whether* $\Gamma \vdash_{\mathcal{M}_+} S \leq T$*.*

Proof. From Proposition 1 and Theorem 1, it follows that $\Gamma \vdash_{\mathcal{M}_+} S \leq T$ if and only if $\Gamma \vdash_{\mathcal{M}'_+} S \leq T$ via a subterm derivation. Also, note that for any derivation $\mathcal{D}$ of $S \leq T$, if any branch in $\mathcal{D}$ contains two separate instances of some $A \leq B$, then we can obtain a new derivation of $S \leq T$ by replacing the entire subderivation of the lower instance of $A \leq B$ with the subderivation of the higher instance, thereby removing the redundancy. We can repeat this process until there are no more repeated inequalities in any branch. Therefore, $\Gamma \vdash_{\mathcal{M}_+} S \leq T$ if and only if $\Gamma \vdash_{\mathcal{M}'_+} S \leq T$ via a subterm derivation in which no branch has any inequality appearing more than once in it. We see that there are finitely many such derivations, since there are finitely many pairs of subterms of (Γ, S, T) and finitely many proof rules in $\mathcal{M}_+$. Thus, we can enumerate all such derivations. □

4. The General Case

We now return to discussing the Monotonicity Calculus $\mathcal{M}$. The goal is still to define a new proof system that is equivalent to $\mathcal{M}$ in some sense and prove that it has the subterm property. We begin by pointing out a key difference from the previous section and providing some intuition. In $\mathcal{M}$, we may have derivations of the form.

$$\dfrac{\dfrac{\dfrac{\vdots}{A \leq B}}{F(A) \leq F(B)}\,(\text{Mono}) \qquad \dfrac{\dfrac{\vdots}{F^+ \leq G^-}}{F(B) \leq G(B)}\,(\text{Point})}{F(A) \leq G(B)}\,(\text{Trans}) \qquad \dfrac{\dfrac{\vdots}{C \leq B}}{G(B) \leq G(C)}\,(\text{Anti})}{F(A) \leq G(C)}\,(\text{Trans}) \tag{5}$$

where B is not a subterm of $(\Gamma, F(A), G(C))$. (Recall that uppercase letters stand for complex terms). Then $F(B)$ is also not a subterm, assuming it is not in an axiom. In the

previous section (when we considered only monotonic functions), if we had the conclusion $F(A) \leq G(C)$, for instance via a derivation of the form

$$\dfrac{\dfrac{\dfrac{\vdots}{A \leq B}}{F(A) \leq F(B)}\,(\text{Mono}) \quad \dfrac{\dfrac{\vdots}{F^+ \leq G^+}}{F(B) \leq G(B)}\,(\text{Point})}{F(A) \leq G(B)}\,(\text{Trans}) \quad \dfrac{\dfrac{\vdots}{\boldsymbol{B \leq C}}}{\boldsymbol{G(B) \leq G(C)}}\,\textbf{(Mono)}}{F(A) \leq G(C)}\,(\text{Trans}) \tag{6}$$

then we could convert it to

$$\dfrac{\dfrac{\vdots}{F^+ \leq G^+} \quad \dfrac{\dfrac{\vdots}{A \leq B} \quad \dfrac{\vdots}{B \leq C}}{A \leq C}\,(\text{Trans})}{F(A) \leq G(C)}\,(\text{App}) \tag{7}$$

in the proof system $\mathcal{M}_+$ that we defined. Here, $F \leq G$ and $A \leq C$ both appear in the derivation, thus essentially eliminating $F(B)$ from the derivation; B can then be eliminated, if necessary, by induction. But if we try to do something similar for derivation (5), it is not clear what should take the place of $A \leq C$ in (7):

$$\dfrac{\dfrac{\vdots}{F^+ \leq G^-} \quad \dfrac{\dfrac{\vdots}{A \leq B} \quad \dfrac{\vdots}{C \leq B}}{???}}{F(A) \leq G(C)}$$

Let us rewrite this as

$$\dfrac{\dfrac{\vdots}{F^+ \leq G^-} \quad \dfrac{\dfrac{\vdots}{A \leq B} \quad \dfrac{\vdots}{B \geq C}}{???}}{F(A) \leq G(C)}$$

so that it is more suggestive of the form of (7). We want to be able to go directly from $F^+ \leq G^-$ and $A \leq B \geq C$ (where $A \leq B \geq C$ is shorthand for the two separate inequalities $A \leq B$; $C \leq B$) to $F(A) \leq G(C)$, without the term $F(B)$ appearing. But it appears that, due to the problem just mentioned, we do not have an immediate analogue for the step in the proof of Theorem 1 that allowed us to eliminate non-subterms.

To address this issue, we will introduce a new relation symbol $\lesseqgtr$ called a *zigzag*. Intuitively, $S_1 \lesseqgtr S_n$ in this new language represents a sequence of inequalities

$$S_1 \leq S_2 \geq \cdots \leq S_n \tag{8}$$

in the original system $\mathcal{M}$, for some terms $S_2, \ldots, S_{n-1}$. (This is somewhat of a simplification. In general, the direction of the inequalities do not need to alternate as in the pattern displayed in (8). For example, $S_1 \lesseqgtr S_4$ could represent $S_1 \leq S_2 \leq S_3 \geq S_4$ or even $S_1 \leq S_2 \leq S_3 \leq S_4$.). Now, after introducing new (Trans) and (App) rules, we can get a derivation like

$$\dfrac{\dfrac{\vdots}{F^+ \leq G^-} \quad \dfrac{\dfrac{\vdots}{A \leq B} \quad \dfrac{\vdots}{B \geq C}}{A \lesseqgtr C}\,(\text{Trans})}{F(A) \leq G(C)}\,(\text{App}) \tag{9}$$

because antitonic G reverses the direction of the inequality in $B \geq C$, as seen in (5). (To be clear, the previous and following derivations are merely schematic and do not always follow the official syntax to be defined)

We should emphasize that from the perspective of $\mathcal{M}$, the inequalities in (8) should be viewed as separate inequalities $S_1 \leq S_2$; $S_3 \leq S_2$; …; $S_{n-1} \leq S_n$ perhaps obtained through separate derivations—they are merely connected by the fact that they have terms in common, and they cannot be combined in a single derivation. We have already seen above that the inequalities $A \leq B$ and $B \geq C$, corresponding to $A \lesseqgtr C$, cannot be combined in $\mathcal{M}$, but to further illustrate this point, consider this slight modification of the previous example (in the new system):

$$\dfrac{\dfrac{\vdots}{F^+ \leq \boldsymbol{G^+}} \quad \dfrac{\dfrac{\vdots}{A \leq B} \quad \dfrac{\vdots}{B \geq C}}{A \lesseqgtr C}\,(\text{Trans})}{\boldsymbol{F(A) \lesseqgtr G(C)}}\,(\text{App}) \tag{10}$$

Here, $F(A) \lesseqgtr G(C)$ can be thought of as $F(A) \leq F(B) \leq G(B) \geq G(C)$, corresponding to the separate derivations (in $\mathcal{M}$)

$$\dfrac{\dfrac{\dfrac{\vdots}{A \leq B}}{F(A) \leq F(B)}\,(\text{Mono}) \quad \dfrac{\dfrac{\vdots}{F^+ \leq G^+}}{F(B) \leq G(B)}\,(\text{Point})}{F(A) \leq G(B)}\,(\text{Trans}) \qquad \dfrac{\dfrac{\vdots}{C \leq B}}{G(C) \leq G(B)}\,(\text{Mono}) \tag{11}$$

whose conclusions cannot be combined (contrast this with (5)). However, with the syntactic flexibility introduced by zigzags, we can then combine them in a derivation, as seen in (10).

Now, having converted (5) to (9), just as we converted (6) to (7), we have completed the preliminary motivation of the new proof system. However, our examples so far have been rather simple, suppressing certain details that we must consider. In general, we cannot completely ignore the intermediate terms in the sequence (8) that a zigzag $S_1 \lesseqgtr S_n$ corresponds to. (Contrast this with $\mathcal{M}_+$. There, if we have derived the sequence of inequalities $S_1 \leq S_2 \leq \cdots \leq S_n$, then we can derive $S_1 \leq S_n$ by repeatedly applying (Trans) and then forget about the intermediate terms.) These intermediate terms contain some information which we must retain to ensure that the new proof rules we are defining are sound. We discuss the relevant information here below.

Call the left premise of (App) *function position* and the right premise *argument position*. Observe that, as suggested by (10) and (11), derivation (5) would not go through if G were monotonic. The idea is that we would no longer have anything that could flip the direction of the second inequality in $A \leq B \geq C$. So, if we replace $F^+ \leq G^-$ in (9) with $F^+ \leq G^+$, we should not be able to derive $F(A) \leq G(C)$ in our new system:

$$\dfrac{\dfrac{\vdots}{F^+ \leq G^+} \quad \dfrac{\dfrac{\vdots}{A \leq B} \quad \dfrac{\vdots}{B \geq C}}{A \lesseqgtr C}\,(\text{Trans})}{\cancel{F(A) \leq G(C)}}\,(\text{App}) \tag{12}$$

Likewise, in argument position, if we had $A \lesseqgtr C$ representing $A \leq B_1 \geq B_2 \leq C$ instead of $A \leq B \geq C$ in (9), we should also not be able to derive $F(A) \leq G(C)$:

$$\dfrac{\dfrac{\vdots}{F^+ \leq G^-} \quad \dfrac{\dfrac{\dfrac{\vdots}{A \leq B_1} \quad \dfrac{\vdots}{B_1 \geq B_2}}{A \lesseqgtr B_2}\,(\text{Trans}) \quad \dfrac{\vdots}{B_2 \leq C}}{A \lesseqgtr C}\,(\text{Trans})}{\cancel{F(A) \leq G(C)}}\,(\text{App}) \tag{13}$$

because $\{F^+ \leq G^-, A \leq B_1, B_2 \leq B_1, B_2 \leq C\} \nvdash_{\mathcal{M}} F(A) \leq G(C)$. The idea here is that G^- flips either (a) all the inequalities, (b) the last two inequalities, or (c) the last inequality in

$A \leq B_1 \geq B_2 \leq C$, but none of these cases results in inequalities which are all in the same direction. As an example, case (b) yields $F(A) \leq F(B_1) \leq G(B_1) \leq G(B_2) \geq G(C)$.

Thus, we see that, intuitively, fewer polarity marking changes in function position (as in (12), compared to (9)) and more inequality direction flips in argument position (as in (13), compared to (9)) will yield weaker conclusions; in other words, more polarity marking changes and fewer inequality direction flips will yield stronger conclusions. (Intuitively, a statement is *stronger* than another if the sequence of inequalities it corresponds to has fewer flips. That is, it is closer to becoming an inequality.) In particular, $A \lesseqgtr C$ in (13) should not be treated the same as $A \lesseqgtr C$ in (9). So, in function position, we need to keep track of the number of polarity marking changes, and in argument position, the number of inequality direction flips.

In all the examples above, there was an inequality in function position of (App), but we might want to consider the more general case where there is a zigzag there. For example (with some details omitted),

$$\dfrac{\dfrac{\vdots \quad \vdots}{F \leq G \qquad G \geq H}\,(\text{Trans}) \qquad \vdots}{\dfrac{F \lesseqgtr H}{F(A) \lesseqgtr G(B)}\,(\text{App})} \qquad \dfrac{}{A \lesseqgtr B}$$

Although this would be irrelevant if it were a standalone derivation (since the conclusion is not an inequality), it becomes relevant when it appears as a subderivation within another derivation whose conclusion is an inequality:

$$\dfrac{\vdots \qquad \dfrac{\dfrac{\vdots \quad \vdots}{F \leq G \qquad G \geq H}\,(\text{Trans}) \qquad \vdots}{\dfrac{F \lesseqgtr H}{F(A) \lesseqgtr H(B)}\,(\text{App})} \qquad \dfrac{}{A \lesseqgtr B}}{\dfrac{M \leq N}{M(F(A)) \leq N(H(B))}\,(\text{App})} \tag{14}$$

The number of flips in function position ($F \lesseqgtr H$ in (14)) affects the number of flips in the immediate conclusion ($F(A) \lesseqgtr H(B)$). In particular, if $F \lesseqgtr H$ has any flips at all (i.e., it is not an inequality), then $F(A) \lesseqgtr H(B)$ will have some flips as well. Previously we have seen, through (13), that we need to keep track of the number of flips in argument position. Since $F(A) \lesseqgtr H(B)$ occurs in argument position, we need to keep track of the number of flips it has, which then means we have to keep track of the number of flips in $F \lesseqgtr H$. So in general, we also have to encode the number of inequality direction flips in function position.

In the language of the new proof system $\mathcal{M}_\infty$, we will replace all inequalities with zigzags and encode all the information discussed above in *info tables* to be defined below. Every zigzag $S_1 \lesseqgtr S_n$ is associated with an info table χ (denoted $S_1 \lesseqgtr_\chi S_n$) which records the sequence (8) of inequalities, but with the intermediate terms replaced by their polarity markings. For example, if $S_2 : e \xrightarrow{-} (e \xrightarrow{+} e)$, then χ will lose access to the term S_2 but retain the information

$$\begin{bmatrix} - \\ + \end{bmatrix}$$

Inequalities can be seen as a special case of zigzags.

The idea is that in an $\mathcal{M}_\infty$-derivation, info tables record intermediate computation which would have appeared in the corresponding $\mathcal{M}$-derivation(s), using finite resources (polarity markings instead of terms). Note that in argument position, polarity changes at which there is no inequality flip are also encoded but can be ignored intuitively.

The definitions below formalize the details discussed above.

Symbol	Meaning
ϕ, χ, ψ, ω	Info table $[v_0 *_1 v_1 *_2 \cdots *_n v_n]$
$\chi' \sqsubseteq \chi$	χ' deletes zero or more occurrences of $*_i v_i$ from χ for which $*_i = *_{i+1}$
$\chi' \subseteq \chi$	χ' is a subsequence of χ
$\chi' \preceq \chi$	χ' is a subsequence of χ, considering only the inequality symbols
$i(\chi)$	The number of inequality symbols in χ
$f(\chi)$	The number of inequality direction flips in χ
$S \leqslant_1 T$	$S \leqslant_\chi T$ for some χ whose inequalities are all $\leq$

Definition 13. *An* ***info table*** χ *is a sequence that alternates between polarity marking columns* $v \in \{+,-\}^k$ *(for some k) and inequality symbols* $* \in \{\leq, \geq\}$*, and ends with a polarity marking column. We write* $\chi = [v_0 *_1 v_1 *_2 \cdots *_n v_n]$.

Remark 4. *For ease of presentation, we omit the* $\cdot$ *marking. Doing this does not substantially change the technical details in the paper. Hereafter, think of* $\mathcal{M}$ *as consisting of only monotonic and antitonic functions.*

Let $S_0 *_1 S_1 *_2 \cdots *_n S_n$ be shorthand for the sequence of inequalities $S_0 *_1 S_1$; $S_1 *_2 S_2$; …; $S_{n-1} *_n S_n$.

Definition 14. *For any term* S *with type* $\sigma_1 \overset{m_1}{\to} \ldots \overset{m_{i-1}}{\to} \sigma_i \overset{m_i}{\to} \ldots \overset{m_l}{\to} \sigma_{l+1}$ *(right-associative), where* $l = \text{valence}(S)$*, let* v_S *denote the polarity marking column such that the ith entry of* v_S *is* m_i*. A sequence of inequalities* $S_0 *_1 S_1 *_2 \cdots *_n S_n$ *in* $\mathcal{M}$ ***respects*** χ *if* $\chi = [v_{S_0} *_1 v_{S_1} *_2 \cdots *_n v_{S_n}]$.

While the definition of info tables theoretically allows for arbitrary contents, their information will generally arise from terms involved a derivation.

Example 3. *The info table*

$$\chi = \begin{bmatrix} + & \leq & - & \geq & + \\ + & & + & & - \end{bmatrix}$$

indicates that $S_0 \leq S_1 \geq S_2$ *for some terms*

$$S_0 : e \overset{+}{\to} (e \overset{+}{\to} e),\ S_1 : e \overset{-}{\to} (e \overset{+}{\to} e),\ S_2 : e \overset{+}{\to} (e \overset{-}{\to} e).$$

That is, $S_0 \leq S_1 \geq S_2$ *respects*

$$\chi = [v_{S_0} \leq v_{S_1} \geq v_{S_2}].$$

Definition 15. *If terms* S *and* T *have the same type disregarding their polarity markings, then we can write the* ***zigzag*** $S \leqslant_\chi T$*, where* χ *is an info table.*

To translate an axiom $S \leq T$ in $\mathcal{M}$ into the new language, we simply encode the inequality into the info table, either as $S \leq T$ or as $T \geq S$.

Definition 16. *Given a set* Γ *of inequalities, the set of zigzags* $\Gamma_{\leqslant}$ *is*

$$\{(S \leqslant_{[v_S \leq v_T]} T) \colon (S \leq T) \in \Gamma\} \cup \{(T \leqslant_{[v_T \geq v_S]} S) \colon (S \leq T) \in \Gamma\}$$

We present the new proof rules in three parts, as Definitions 17–19.

Definition 17. $\mathcal{M}_\infty$ *is the proof system with the following rules:*

$$(\chi \in \phi(\psi))\ \frac{F \leqslant_\phi G \qquad A \leqslant_\psi B}{F(A) \leqslant_\chi G(B)}\ \text{(App)} \qquad \frac{A \leqslant_\phi B \qquad B \leqslant_\psi C}{A \leqslant_{\phi+\psi} C}\ \text{(Trans)}$$

https://doi.org/10.3390/logics4020006

$$\frac{}{A \leqslant_{[v_A \leq v_A]} A}\ (\text{Refl})$$

where

- $\phi + \psi$ *is the concatenated sequence* $[\phi\ \psi']$ *where* ψ' *is* ψ *with its first polarity column deleted;*
- $\phi(\psi)$ *is a set of info tables to be defined in Definition 18 (so (App) represents a family of proof rules).*

(App′) and (Trans′) are to be defined in Definition 19.

Example 4. *Let*

$$f_0 : e \xrightarrow{+} (e \xrightarrow{+} e),\ f_1 : e \xrightarrow{-} (e \xrightarrow{+} e),\ f_2 : e \xrightarrow{+} (e \xrightarrow{-} e),\ f_3 : e \xrightarrow{-} (e \xrightarrow{-} e)$$

and let $\Gamma = \{f_0 \leq f_1, f_2 \leq f_1, f_2 \leq f_3\}$. *Then,* $\Gamma_{\leqslant}$ *contains* $f_0 \leqslant_{\phi_1} f_1$, $f_1 \leqslant_{\phi_2} f_2$, *and* $f_2 \leqslant_{\psi} f_3$ *where*

$$\phi_1 = \begin{bmatrix} + & \leq & - \\ + & & + \end{bmatrix},\ \phi_2 = \begin{bmatrix} - & \geq & + \\ + & & - \end{bmatrix},\ \psi = \begin{bmatrix} + & \leq & - \\ - & & - \end{bmatrix}.$$

Then we have

$$\dfrac{\dfrac{f_0 \leqslant_{\phi_1} f_1 \qquad f_1 \leqslant_{\phi_2} f_2}{f_0 \leqslant_{\phi} f_2}\ (\textit{Trans}) \qquad f_2 \leqslant_{\psi} f_3}{f_0 \leqslant_{\chi} f_3}\ (\textit{Trans})$$

where

$$\phi = \begin{bmatrix} + & \leq & - & \geq & + \\ + & & + & & - \end{bmatrix}$$

$$\chi = \begin{bmatrix} + & \leq & - & \geq & + & \leq & - \\ + & & + & & - & & - \end{bmatrix}.$$

Definition 18 ((App) Rule). *Let* $F * \cdots * G$ *be a sequence of k (hypothetical) axioms in* $\mathcal{M}$ *that respects* ϕ. *Let* $A * \cdots * B$ *be a sequence of* k' *(hypothetical) axioms that respects* ψ. *We say that* $\chi \in \boldsymbol{\phi(\psi)}$ *iff some sequence* $F(A) * \cdots * G(B)$ *of* $k + k'$ *inequalities that respects* χ *is derivable from those axioms in* $\mathcal{M}$, *each via a single application of (Mono), (Anti), or (Point) (note that since we are only encoding polarity markings rather than specific terms, the set* $\phi(\psi)$ *is finite). In particular, the polarity columns in info tables in* $\phi(\psi)$ *have one less entry than those in* ϕ.

To provide some intuition for Definition 18, we consider (10) again, now with proper notation (and treating the inequalities in (10) as axioms instead of derived).

Example 5. *Let*

$$f : e \xrightarrow{+} (e \xrightarrow{+} e),\ g : e \xrightarrow{+} (e \xrightarrow{-} e),\ a, b, c : e$$

and let $\Gamma = \{f \leq g, a \leq b, c \leq b\}$. *Consider*

$$\dfrac{f \leqslant_{\phi} g \qquad \dfrac{a \leqslant_{\psi_1} b \qquad b \leqslant_{\psi_2} c}{a \leqslant_{\psi} c}\ (\textit{Trans})}{f(a) \leqslant_{\chi} g(c)}\ (\textit{App})$$

where

$$\psi_1 = [\varnothing \leq \varnothing]$$

$$\psi_2 = [\varnothing \geq \varnothing]$$

$$\psi = [\varnothing \leq \varnothing \geq \varnothing]$$

Note that a, b, c are of base type, so ψ_1, ψ_2, ψ do not contain any polarity markings—we use $\varnothing$ to denoted the empty tuple.

$$\phi = \begin{bmatrix} + & \leq & + \\ + & & - \end{bmatrix}$$

$$\chi = [+ \leq + \leq - \geq -]$$

We will verify that this is a proper application of (App), i.e., $\chi \in \phi(\psi)$. In this case, k and k' from Definition 18 are 1 *and* 2*, respectively.*

The first polarity column of ψ corresponds to a, and the last polarity column corresponds to c. Intuitively, the middle polarity column comes from b earlier in the derivation, but once we have concluded $a \leqslant_\psi c$, we lose access to the term and only retain its polarity information in the info table. Now Definition 18 tells us to introduce a temporary term—call it a_1 to emphasize that we have lost access to the term b—and suppose that we have the inequalities $a \leq a_1$ and $a_1 \geq c$ in $\mathcal{M}$, as encoded by ψ. Also, from the info table ϕ, we have that $f^+ \leq g^+$. Then by replicating the derivations in (11), we get the three inequalities

$$f(a) \leq f(a_1) \leq g(a_1) \geq g(c) \tag{15}$$

in $\mathcal{M}$: the first by (Mono), the second by (Point), and the third by (Mono). Given the types of f and g, we see that

$$f(a) : e \xrightarrow{+} e,\ f(a_1) : e \xrightarrow{+} e,\ g(a_1) : e \xrightarrow{-} e,\ g(c) : e \xrightarrow{-} e,$$

so (15) respects χ. Then by Definition 18, $\chi \in \phi(\psi)$.

Note that from $f \leq g$ and $a \leq a_1 \geq c$, we can also derive the three inequalities

$$f(a) \leq g(a) \leq g(a_1) \geq g(c) \tag{16}$$

with $g(a) : e \xrightarrow{-} e$. Thus, (16) respects

$$\omega = [+ \leq - \leq - \geq -],$$

so $\omega \in \phi(\psi)$, and $f(a) \leqslant_\omega g(c)$ would also be a valid conclusion of this derivation.

Definition 19. *Here are modified versions of (App) and (Trans):*

$$(\chi' \sqsubset \chi \in \phi(\psi))\ \frac{F \leqslant_\phi G \qquad A \leqslant_\psi B}{F(A) \leqslant_{\chi'} G(B)}\ (\text{App}')$$

$$(\chi' \sqsubset \phi + \psi)\ \frac{A \leqslant_\phi B \qquad B \leqslant_\psi C}{A \leqslant_{\chi'} C}\ (\text{Trans}')$$

where $\chi' \sqsubset \chi$ means χ' is the result of deleting some occurrences of $*_i\, v_i$ from $\chi = [v_0 *_1 v_1 *_2 \cdots *_n v_n]$ for which $*_i = *_{i+1}$. The side condition for (App′) is that $\chi' \sqsubset \chi$ for some $\chi \in \phi(\psi)$.

Remark 5. *It will be convenient to speak of and think of (App′) as the same rule as (App) but followed by an application of a separate* ***deletion rule*** *on the conclusion, and likewise for (Trans′).*

Example 6. *Consider the same setup in Example 5.*

$$\frac{f \leqslant_\phi g \qquad \dfrac{a \leqslant_{\psi_1} b \qquad b \leqslant_{\psi_2} c}{a \leqslant_\psi c}\ (\textit{Trans})}{f(a) \leqslant_{\chi'} g(c)}\ (\textit{App}')$$

where $\chi' = [+ \leq - \geq -]$. *Note, however, that deleting the last inequality in* χ *would reduce the number of flips and is not allowed by (App′).*

We can verify that deleting a polarity column which has the same inequality direction immediately after it in info tables does not allow us to derive additional stronger statements. That is, roughly speaking, the deletion rule is admissible. To prove this, we show that anything that is derivable from an info table with such deletions would essentially also have been derivable from the info table without any deletions. The intuitive idea is that the number of inequality flips remains the same, and deleting polarity columns amounts to deleting information.

Definition 20. *Let* $\chi = [v_0 *_1 v_1 *_2 \cdots *_n v_n]$. *Recall from Definition 19 that* $\chi' \sqsubset \chi$ *means* χ' *is the result of deleting some occurrences of* $*_i v_i$ *from* χ *for which* $*_i = *_{i+1}$. *Let* $\chi' \sqsubseteq \chi$ *denote* $\chi' \sqsubset \chi$ *or* $\chi' = \chi$, *and let* $\chi' \sqsubseteq_1 \chi$ *denote* $\chi' \sqsubset \chi$ *where exactly one such* $*_i v_i$ *is deleted.*

Lemma 2. *Let* $\phi' \sqsupseteq_1 \phi$, *and let* $\mathcal{D}$ *be a derivation of* $S \lesseqgtr_\chi T$ *from* $\Gamma_\lesseqgtr$ *that contains* $A \lesseqgtr_\phi B$. *If the subderivation of* $A \lesseqgtr_\phi B$ *is replaced with* $A \lesseqgtr_{\phi'} B$, *then there is a derivation* $\mathcal{D}'$ *of* $S \lesseqgtr_{\chi'} T$ *(from* $\Gamma_\lesseqgtr$*) for some* $\chi' \sqsupseteq_1 \chi$ *that differs from* $\mathcal{D}$ *at most by having each info table* α *below* $A \lesseqgtr_\phi B$ *in* $\mathcal{D}$ *modified to some* $\alpha' \sqsupseteq_1 \alpha$ *in* $\mathcal{D}'$ *(but keeping the terms the same).*

Definition 21. *Let* $\chi = [v_0 *_1 v_1 *_2 \cdots *_n v_n]$. *An* ***inequality flip****, or simply* ***flip****, occurs at* $*_i$ *if* $*_i \neq *_{i+1}$. *The number of flips in* χ *(including* $*_n$*) is denoted by* $f(\chi)$.

Note that we count flips starting from 1; i.e., if all the inequalities in χ have the same direction, then $f(\chi) = 1$.

Another consequence of Definitions 17 and 18 is

Lemma 3. *For any* $\chi \in \phi(\psi)$, $f(\chi) \geq f(\phi)$.

The idea is that in $\mathcal{M}$, (Point) does not have an order-reversing analogue like (Mono) does with (Anti). See Example A1 in the Appendix A for an illustration.

The following shorthand will occasionally be used: $S \lesseqgtr T$ means $S \lesseqgtr_\chi T$ for some χ. $S \lesseqgtr_1 T$ means $S \lesseqgtr_\chi T$ for some χ whose inequalities are all $\leq$.

We have just defined an infinite proof system $\mathcal{M}_\infty$: we have not placed any bounds on info tables so far. But given Γ, we can systematically eliminate unnecessary rules by limiting the info tables allowed. Eventually, we will end up with a finite proof system $\mathcal{M}_\Gamma$, to be defined below, in order to demonstrate decidability.

An outline of the proof of decidability is

I. $\Gamma \vdash_\mathcal{M} S \leq T$ if and only if $\Gamma_\lesseqgtr \vdash_{\mathcal{M}_\infty} S \lesseqgtr_1 T$.
II. $\Gamma_\lesseqgtr \vdash_{\mathcal{M}_\infty} S \lesseqgtr_1 T$ if and only if $\Gamma_\lesseqgtr \vdash_{\mathcal{M}_\infty} S \lesseqgtr_1 T$ via a subterm derivation.
III. $\Gamma_\lesseqgtr \vdash_{\mathcal{M}_\infty} S \lesseqgtr_1 T$ via a subterm derivation if and only if $\Gamma_\lesseqgtr \vdash_{\mathcal{M}_\Gamma} S \lesseqgtr_1 T$ via a subterm derivation.

with each roughly corresponding to one of the next three subsections. Part I is the analogue of Proposition 1, and part II is the analogue of Theorem 1. In part III, we bound the size of info tables.

4.1. Part I: Derivational Equivalence

We first show that $\mathcal{M}$ embeds into $\mathcal{M}_\infty$. The following result is analogous to the forward direction of Proposition 1.

Proposition 2. *If* $\Gamma \vdash_\mathcal{M} S \leq T$, *then* $\Gamma_\lesseqgtr \vdash_{\mathcal{M}_\infty} S \lesseqgtr_1 T$.

Conversely, we have

Proposition 3. *If* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} S \leqslant_\chi T$, *then* $\Gamma \vdash_{\mathcal{M}} S * \cdots * T$ *via a sequence of derivations that respects* χ.

Proof. By structural induction on the proof rules. Use Lemma 2. □

4.2. Part II: Subterm Property

Remark 6. *The definitions of* ***subterm derivation***, ***subterm property***, *and* ***bottom-most chain*** *are analogous to those in Section 3. Then, we also have an analogue for Lemma 1.*

Theorem 2. *If* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} S \leqslant_\chi T$, *then* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} S \leqslant_\chi T$ *via a subterm derivation.*

The proof is almost identical to the proof of Theorem 1, but there is an additional consideration with respect to info tables. See the Appendix A.

4.3. Part III: Bounding Info Tables

Since this section is rather dense, we first provide a guiding example that highlights the ideas in the results that follow. Throughout this section, we use $S \leq T$ as shorthand for $S \leqslant_{[v_S \leq v_T]} T$, and $S \geq T$ for $S \leqslant_{[v_S \geq v_T]} T$.

Example 7. *Let* Γ *consist of* $a \leq b \geq c : e$, $f^+ \leq g^- \geq h^+ : e \to e$, *and* $p^+ \leq q^+ \leq r^- : e \to e$. *Consider the derivation from axioms in* $\Gamma_{\leqslant}$ *(all the omitted rules are (Trans)):*

$$\dfrac{\dfrac{p^+ \leq q^+ \qquad q^+ \leq r^-}{p \leqslant_\phi r} \qquad \dfrac{\dfrac{f^+ \leq g^- \qquad g^- \geq h^+}{f \leqslant h} \qquad \dfrac{\dfrac{a \leq b \qquad b \geq c}{a \leqslant_{\psi_1} c} \qquad \dfrac{}{c \leq c}\,(Refl)}{a \leqslant_{\psi_2} c}\,(App)}{f(a) \leqslant_\omega h(c)}\,(App)}{p(f(a)) \leqslant_\chi r(h(c))}\,(App)$$

where ω *consists of three* $\leq$*'s followed by two* $\geq$*'s, and* χ *consists of seven* $\leq$*'s. The other info tables, some of which are omitted, are fixed by the definition of (Trans). In particular, we note that* ψ_1 *has two inequality flips, and* ψ_2 *has three, so* $a \leqslant_{\psi_1} c$ *is a stronger statement than* $a \leqslant_{\psi_2} c$.

We will convert this to a simpler derivation, both by reducing the number of nodes and by reducing the size of certain info tables using the deletion rules (App′) and (Trans′). First, we replace the subderivation of $a \leqslant_{\psi_2} c$ *with the smaller subderivation of* $a \leqslant_{\psi_1} c$. *Having changed one of the premises of that instance of (App), we then adjust the info table in* $f(a) \leqslant h(c)$ *so that it is still a proper application of (App), and then likewise for* $p(f(a)) \leqslant r(h(c))$, *resulting in*

$$\dfrac{\dfrac{p^+ \leq q^+ \qquad q^+ \leq r^-}{p \leqslant_\phi r}\,(Trans) \qquad \dfrac{\dfrac{f^+ \leq g^- \qquad g^- \geq h^+}{f \leqslant h} \qquad \dfrac{\boldsymbol{a \leq b \qquad b \geq c}}{\boldsymbol{a \leqslant_{\psi_1} c}}}{\boldsymbol{f(a) \leqslant_{\omega'} h(c)}\,(App)}}{\boldsymbol{p(f(a)) \leqslant_{\chi'} r(h(c))}\,(App)}$$

where

$$\omega' = [\varnothing \leq \varnothing \leq \varnothing \geq \varnothing \geq \varnothing].$$

This replacement may remind us of Corollary 1, but the difference is that here, $a \leqslant_{\psi_1} c$ *and* $a \leqslant_{\psi_2} c$ *are different statements. Nonetheless, we observed that* $a \leqslant_{\psi_1} c$ *is a stronger statement than* $a \leqslant_{\psi_2} c$, *so performing this replacement should be advantageous.*

Then, using the deletion rule, we delete one $\leq$ *and one* $\geq$ *from* ω', *since this is redundant information in argument position:*

$$\dfrac{\dfrac{p^+ \le q^+ \qquad q^+ \le r^-}{p \lesssim_{\phi} r}\,(Trans) \qquad \dfrac{\dfrac{f^+ \le g^- \qquad g^- \ge h^+}{f \lesssim h} \qquad \dfrac{a \le b \qquad b \ge c}{a \lesssim_{\psi_1} c}\,(App')}{f(a) \lesssim_{\omega''} h(c)}\,(App)}{p(f(a)) \lesssim_{\chi''} r(h(c))}$$

where

$$\omega'' = [\varnothing \le \varnothing \ge \varnothing].$$

Finally, we consider

$$\phi = [+ \le + \le -]$$

in function position of the last application of (App). Temporarily associate ϕ with $p^+ \le F^+ \le r^-$ and ω'' with $f(a) \le A \ge h(c)$. Suppose that

$$\chi'' = [\varnothing \le \varnothing \le \varnothing \le \varnothing \le \varnothing]$$

corresponds to $p(f(a)) \le p(A) \le F(A) \le r(A) \le r(h(c))$, obtained by (Mono), (Point), (Point), and (Anti), respectively. Observe, then, that this uses only the polarity information from p (for (Mono)) and r (for (Anti)), and not from F. That is, $p^+ \le r^-$ suffices to align all inequalities in the same direction, so the polarity information from F^+ is unnecessary. Hence, we delete the middle polarity entry in ϕ and then adjust the conclusion accordingly:

$$\dfrac{\dfrac{p^+ \le q^+ \qquad q^+ \le r^-}{p \lesssim_{\phi'} r}\,(Trans') \qquad \dfrac{\dfrac{f^+ \le g^- \qquad g^- \ge h^+}{f \lesssim h} \qquad \dfrac{a \le b \qquad b \ge c}{a \lesssim_{\psi_1} c}\,(App')}{f(a) \lesssim_{\omega''} h(c)}\,(App)}{p(f(a)) \lesssim_{\chi'''} r(h(c))}$$

where

$$\phi' = [+ \le -].$$

We can also apply a deletion rule in the last application of (App) to further simplify χ'''.

This concludes the example. The procedure of replacing $a \lesssim_{\psi_2} c$ with $a \lesssim_{\psi_1} c$ in the first step of our conversion hints at the proof idea for Lemma 4, and Lemma 5 ensures that adjusting the info tables below is unproblematic. Proposition 4 involves repeatedly applying these procedures. The second step in our conversion features as part of Proposition 5, which is completed by applying Lemma 6, which generalizes the third step in our conversion.

Fix a finite Γ. We first show that any derivation whose conclusion has an info table with a "pathologically" large number of inequality flips can be converted to a derivation where the terms in the conclusion remain the same but its info table (as well as the other info tables in the derivation) is bounded in the number of flips.

Definition 22. *Let $i(\chi)$ denote the number of inequality symbols in an info table χ.*

Also recall that $f(\chi)$ denotes the number of inequality flips in χ.

Lemma 4. *Let N be the size of the subterm set of (Γ, S, T). If $\Gamma_{\lesssim} \vdash_{\mathcal{M}_\infty} S \lesssim_{\chi} T$ via a subterm derivation with $f(\chi) > 2^{N^2}$, then $\Gamma_{\lesssim} \vdash_{\mathcal{M}_\infty} S \lesssim_{\chi'} T$ via a subterm derivation in which every info table has at most 2^{N^2} inequality flips.*

Proof. The reasoning here is similar to that in Corollary 1.

We first observe that if no branch in the derivation of $S \leqslant_\chi T$ contains any repeated pair of terms, i.e., at least two occurrences of $A \leqslant B$ for some pair A, B (where the info tables in the two instances are not necessarily the same), then $f(\chi) \leq 2^{N^2}$. This follows from a straightforward counting argument: the leaves of any derivation will have 1 inequality each in their info tables. For any step, if the info tables of the premises have k_1, k_2 inequalities each, then the conclusion of that step has at most $k_1 + k_2$ inequalities by Definitions 17 and 18. Now, since χ has height at most the number of pairs of subterms, which is N^2, $i(\chi) \leq 2^{N^2}$.

Thus, if $f(\chi) > 2^{N^2}$, the derivation $\mathcal{D}$ of $S \leqslant_\chi T$ must contain two separate instances $A \leqslant_\phi B, A \leqslant_\psi B$ in some branch. WLOG, assume $A \leqslant_\psi B$ is below $A \leqslant_\phi B$ in $\mathcal{D}$. Then we obtain a new derivation by replacing the entire subderivation of $A \leqslant_\psi B$ with the subderivation of $A \leqslant_\phi B$ and then adjusting the info tables below, as dictated by the proof rules, in $\mathcal{D}$ (the details of how the info tables are adjusted do not matter; all that matters is that in the end, every info table will have at most 2^{N^2} flips). Repeat this process until there are no more repeated pairs of terms in any branch. The resulting derivation will be as desired. □

Note that, strictly speaking, the lemma above is true no matter what $f(\chi)$ is. We state the lemma as such to emphasize that we will only apply it when $f(\chi)$ is very large. In particular, the sizes of info tables in Example 7 are already safely within the given bound, but in that specific case, it was unproblematic to perform the transformation described in the lemma. However, in other cases, performing this transformation on a derivation whose conclusion has few flips (but contains a pair of repeated terms in some branch) could increase the number of flips in the conclusion, making it weaker. Consider

$$\dfrac{\dfrac{a \leq f(a) \qquad \dfrac{f^+ \leq g^- \qquad \dfrac{a \leq b \qquad b \geq c}{a \leqslant_\psi c}\,(\text{Trans})}{f(a) \leqslant_1 g(c)}\,(\text{App})}{a \leqslant_1 g(c)}\,(\text{Trans}) \qquad g(c) \leq c}{a \leqslant_1 c}\,(\text{Trans}) \tag{17}$$

We would not want to replace the derivation with the subderivation of $a \leqslant_\psi c$, which has two flips.

Hereafter, let $K_\Gamma := 2^{N^2}$ where N is the size of the subterm set of (Γ, S, T). We sometimes omit the subscript Γ when no confusion arises.

Definition 23. *Let $\chi' \subseteq \chi$ denote that χ' is a subsequence of χ. Let $\chi' \subseteq_1 \chi$ denote that χ' is the result of deleting exactly one $*_i v_i$ from χ.*

Definition 24. *Let $\chi' \preceq \chi$ denote that the sequence of inequalities within χ' is a subsequence of the sequence of inequalities within χ, i.e., χ' is a subsequence of χ disregarding the polarity columns. Let $\chi' \preceq_1 \chi$ denote that $\chi' \preceq \chi$ and χ' has one fewer $*_i v_i$ than χ.*

Example 8. $[- \leq - \geq -] \preceq [+ \leq + \geq + \leq + \geq +]$.

Note that $\chi' \subseteq \chi$ or $\chi' \preceq \chi$ means that $S \leqslant_{\chi'} T$ is at least as strong of a statement as $S \leqslant_\chi T$.

Proposition 4. *If $\Gamma_\leqslant \vdash_{\mathcal{M}_\infty} S \leqslant_\chi T$ via a subterm derivation where $f(\chi) \leq K_\Gamma$, then $\Gamma_\leqslant \vdash_{\mathcal{M}_\infty} S \leqslant_{\chi'} T$ for some $\chi' \subseteq \chi$ via a subterm derivation in which every info table has at most K_Γ inequality flips.*

Proof. Consider the bottom-most node $A \leqslant_\psi B$ with $f(\psi) > K$ in the original derivation. (If there are multiple bottom-most such nodes, pick any one of them.) We claim that it must occur in argument position of (App). If it were in function position or as a premise of (Trans), then the conclusion of that step would also have more than K flips by Lemma 3, contradicting that $A \leqslant_\psi B$ is the bottom-most.

By Lemma 4, we can replace the subderivation of $A \leqslant_\psi B$ with a derivation of $A \leqslant_{\psi'} B$ that contains at most K inequality flips in each of its info tables. If ψ' contains any $*_i\, v_i$ where $*_i = *_{i+1}$, apply the deletion rule on them to get $\psi'' \preceq \psi$. Then we finish the proof with Lemma 5 below, which implies that, in argument position, replacing an info table with one that has fewer inequalities allows us to derive a conclusion that is at least as strong as the original, without significantly changing the structure of the derivation.

Applying Lemma 5 (repeatedly) to $A \leqslant_{\psi''} B$, we get a new derivation with conclusion $S \leqslant_{\chi'} T$ for some $\chi' \subseteq \chi$. Repeat this process—using Lemmas 4 and 5—until there are no more info tables with more than K flips. □

Lemma 5. *Let $\mathcal{D}$ be a derivation of $S \leqslant_\chi T$ that contains $A \leqslant_\psi B$ in argument position of (App), and let $\psi' \preceq_1 \psi$. If the subderivation of $A \leqslant_\psi B$ is replaced with $A \leqslant_{\psi'} B$, then there is a derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ for some $\chi' \subseteq_1 \chi$ that differs from $\mathcal{D}$ at most by having each info table α below $A \leqslant_\psi B$ in $\mathcal{D}$ modified to some $\alpha' \subseteq_1 \alpha$ in $\mathcal{D}'$.*

We have just established that we only need to encode at most K inequality flips. But in function position, we may need to keep inequalities in the same direction (between flips) to derive stronger statements. For example, consider

$$\dfrac{\dfrac{f^+ \le g^- \qquad g^- \le h^+}{f \leqslant_\phi h}\,(\text{Trans}) \qquad \dfrac{\dfrac{a \le b \qquad b \ge c}{a \leqslant_{\psi_1} c}\,(\text{Trans}) \qquad c \le d}{a \leqslant_{\psi_2} d}\,(\text{Trans})}{f(a) \leqslant_\chi h(d)}\,(\text{App}) \tag{18}$$

where

$$\phi = [+ \le - \le +]$$
$$\psi_1 = [\varnothing \le \varnothing \ge \varnothing]$$
$$\psi_2 = [\varnothing \le \varnothing \ge \varnothing \le \varnothing]$$

and $f(\chi) = 1$. We see that if we deleted "$\le -$" from ϕ, then $f(\chi) \neq 1$, so in function position, we have to be more careful when applying the deletion rule, compared to in argument position.

We now bound the number of these intermediate inequalities and polarity columns that need to be retained.

Lemma 6. *Let $\mathcal{D}$ be a derivation of the form*

$$\dfrac{\dfrac{\dfrac{\vdots \qquad \vdots}{F_l \leqslant_{\phi_l} G_l \qquad A_l \leqslant_{\psi_l} B_l}\,(\text{App})}{F_l(A_l) \leqslant_{\omega_l} G_l(B_l)}}{\vdots}$$
$$\dfrac{\dfrac{\ddots \qquad\qquad \vdots}{F_1 \leqslant_{\phi_1} G_1 \qquad\qquad A_1 \leqslant_{\psi_1} B_1}\,(\text{App})}{F_1(A_1) \leqslant_{\omega_1} G_1(B_1)}$$
$$\vdots$$
$$S \leqslant_\chi T$$

where the zigzags in the path downward from $F_l \leqslant_{\phi_l} G_l$ are all in either function position of (App) or a premise of (Trans), and l is the number of such zigzags that are in function position. Then there exists $\phi' \sqsubseteq \phi_l$ with at most $f(\phi_1) + i(\psi_1) + \cdots + f(\phi_l) + i(\psi_l)$ inequalities such that if the subderivation of $F_l \leqslant_{\phi_l} G_l$ in $\mathcal{D}$ is replaced with $F_l \leqslant_{\phi'} G_l$, then there is a derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ for some $\chi' \sqsubseteq \chi$ that differs from $\mathcal{D}$ at most by having each info table α below $F_l \leqslant_{\phi_l} G_l$ in $\mathcal{D}$ modified to some $\alpha' \sqsubseteq \alpha$ in $\mathcal{D}'$.

Definition 25. *$\mathcal{M}_\Gamma$ is the restriction of $\mathcal{M}_\infty$ to zigzags whose info tables have at most $2K_\Gamma L_\Gamma$ inequalities, where L_Γ is the maximum valence (defined in Definition 1 and used in Definition 13) of any term in Γ.*

Proposition 5. *If $\Gamma_\leqslant \vdash_{\mathcal{M}_\infty} S \leqslant_\chi T$ via a subterm derivation where every info table has at most K_Γ flips, then $\Gamma_\leqslant \vdash_{\mathcal{M}_\Gamma} S \leqslant_{\chi'} T$ for some $\chi' \subseteq \chi$ via a subterm derivation.*

Proof. Throughout this proof, we will treat the conclusion as argument position.

Let $\mathcal{D}$ be a subterm derivation where every info table has at most K flips. First, for every $A \leqslant_\psi B$ in $\mathcal{D}$ which is in argument position, delete all $*_i v_i$ in ψ for which $*_i = *_{i+1}$, and also delete the corresponding $*_i v_i$ in any info tables in the bottom-most chain of the subderivation of $A \leqslant_\psi B$. Now, each of these info tables will have at most K inequalities. By Lemma 5, we get a new derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ for $\chi' \subseteq \chi$ without adding inequalities at any info table.

Now consider the top-most node $F \leqslant_\phi G$ in $\mathcal{D}'$ that is in function position, and let $A_0 \leqslant_{\psi_0} B_0$ (to be thought of as $S \leqslant_\chi T$ in the statement of Lemma 6) be the nearest (top-most) zigzag in argument position in the path downward from $F \leqslant_\phi G$. Let l be the number of zigzags in function position in this path. By Lemma 6, we can replace ϕ with $\phi' \sqsubseteq \phi$ which has at most $f(\phi_1) + i(\psi_1) + \cdots + i(\phi_l) + f(\psi_l)$ inequalities—where each ϕ_i, ψ_i is as depicted in the lemma—to obtain a new derivation $\mathcal{D}''$ of $S \leqslant_{\chi''} T$ for $\chi'' \sqsubseteq \chi' \subseteq \chi$ without adding inequalities at any info table. Since $l \leq L$, where L is the maximum valence of any term in Γ, and $f(\phi_i), i(\psi_i) \leq K$, it follows that $i(\phi') \leq 2KL$. Repeat this process on all info tables in function position with more than $2KL$ inequalities. Finally, make deletions accordingly in the bottom-most chain of these subderivations ending in function position, just as we did for argument position, so that all info tables have at most $2KL$ inequalities. □

4.4. Putting It Together

Finally, here is the result we want.

Theorem 3. *$\Gamma \vdash_{\mathcal{M}} S \leq T$ if and only if $\Gamma_\leqslant \vdash_{\mathcal{M}_\Gamma} S \leqslant_1 T$ via a subterm derivation.*

Proof. Refer back to the outline at the beginning of the section. For the forwards direction, (I) is by Proposition 2, (II) is by Theorem 2, and (III) is by Propositions 4 and 5. For the backwards direction, (I) is by Proposition 3, and (II) and (III) are trivial. □

Corollary 2. *It is decidable, given a finite set Γ of ground inequalities and ground terms S, T, whether $\Gamma \vdash_{\mathcal{M}} S \leq T$.*

Proof. There are finitely many distinct info tables, and consequently finitely many zigzags and finitely many proof rules, in $\mathcal{M}_\Gamma$ by Definition 25. So we can replicate the argument in the proof of Corollary 1. □

https://doi.org/10.3390/logics4020006

5. Conclusions and Future Work

We have shown that the Monotonicity Calculus in [2,3] is decidable over a modified set of ground terms: in the setting where polarity markings are ignored when forming terms via function application, amounting to the assumption that argument types within a function type do not carry substantive monotonicity information.

In addition, by considering weaker fragments and then moving to stronger fragments, we have identified the sources of complexity within the system. This can be seen from both the move from $\mathcal{M}_+$ (Section 3) to $\mathcal{M}_\infty$ (Section 4) and the move from simpler constructions to more complicated ones within Section 4. Notably, the latter illuminates the fact that mixed monotonicity (defined in Section 2) introduces additional complexity and pinpoints where exactly: for instance, in (18). One takeaway is that systematic studies of decidability can give rise not only to interesting formal questions but also to substantive empirical questions about natural language—for example, those pertaining to mixed monotonicity.

We close by clarifying the scope of our results and suggesting directions for future work. One point concerns the restriction on term formation described at the beginning of the section. Despite our remark about types e and t in Remark 1, this assumption might be undesirable in the context of natural language when we expand our inventory of types to include other commonly used ones, such as v for events. In particular, it has been observed that measure adverbials such as `for an hour` only take *atelic* predicates [9]. For example, consider

(1) a. John searched his house for an hour.
b. #John found his keys for an hour.
c. John found his keys *in* an hour.

Atelicity can be seen as including an antitonicity constraint: P is atelic only if for any events e, f,

$$e \sqsubseteq f \implies P(f) \to P(e)$$

where $\sqsubseteq$ is the parthood relation on events. So we might want the type of `for an hour` to be $(v \xrightarrow{-} t) \xrightarrow{+} (v \xrightarrow{\cdot} t)$ where it cannot take arguments of type $v \xrightarrow{+} t$ or $v \xrightarrow{\cdot} t$.

This suggests a natural next step: to investigate the Monotonicity Calculus under the original stricter term formation condition. (Roughly, if $F : \sigma \xrightarrow{m} \tau$ and $A : \sigma'$, then $F(A)$ is a well-formed term only if σ' is a "subtype" of σ. See [2,3] for details.) This system would not be equivalent to $\mathcal{M}_\infty$ because $\mathcal{M}_\infty$ does not take into account information about polarity markings in argument types (for example, the "$-$" in the type $(v \xrightarrow{-} t) \xrightarrow{+} (v \xrightarrow{\cdot} t)$ from above would not get recorded in an info table). We can see this omission in Definition 13—such markings are hidden inside $\sigma_1, \ldots, \sigma_{l+1}$ and not encoded in the info tables we have defined. It would be interesting to explore whether we can modify $\mathcal{M}_\infty$ in a way that will allow us to prove results analogous to those in Section 4 to establish decidability.

Moreover, since our focus was on proving decidability rather than establishing more specific complexity results, we prioritized convenient presentations over tight bounds. Another interesting future direction would be to improve the bounds on info tables in Section 4.3.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A. The General Case: Omitted Proofs

In our previous discussion of the (App) rule, we relied on introducing temporary terms and inequalities, but it will be convenient to directly work with info tables in the proofs below.

Example A1. *Here is a procedure for visualizing Definition 18 that works directly with info tables, rather than inequalities in $\mathcal{M}$. We will first introduce terms, as before, for transparency and then reformulate the setup with only info tables.*

Assume we have

$$\phi = \begin{bmatrix} + & \leq_1 & - & \geq_2 & + \\ + & & + & & - \end{bmatrix}$$

$$\psi = [\varnothing \leq_{\mathbf{3}} \varnothing \geq_{\mathbf{4}} \varnothing \leq_{\mathbf{5}} \varnothing \geq_{\mathbf{6}} \varnothing]$$

(We have labeled the inequality symbols to keep track of particular instances of these inequality symbols—the labels are not technically part of the info tables.)

Suppose ϕ and ψ come from $F_0 \leq F_1 \geq F_2$ and $A_0 \leq A_1 \geq A_2 \leq A_3 \geq A_4$, respectively, where

$$F_0 : b \xrightarrow{+} (b \xrightarrow{+} b),\ F_1 : b \xrightarrow{-} (b \xrightarrow{+} b),\ F_2 : b \xrightarrow{+} (b \xrightarrow{-} b)$$

and A_0, A_1, A_2, A_3 are all be of type b. Then, in $\mathcal{M}$, we can derive $F_0(A_0) \leq F_0(A_1)$ by (Mono), $F_0(A_1) \leq F_1(A_1)$ by (Point), $F_1(A_1) \leq F_1(A_2)$ by (Anti), and so on, resulting in

$$F_0(A_0) \leq F_0(A_1) \leq F_1(A_1) \leq F_1(A_2) \geq F_1(A_3) \geq F_2(A_3) \geq F_2(A_4) \tag{A1}$$

Now we describe the procedure, which can be seen as a reformulation of the setup above. First, insert each inequality symbol from ψ into ϕ after a polarity column and then replicate that column after the inserted inequality ($\leq_3$ must come before $\geq_4$, which must come before $\leq_5$, and so on). We call this an ***insertion of*** $*_i$ ***from*** ψ ***into*** ϕ ***at*** v_j*:*

$$\begin{bmatrix} + & \leq_{\mathbf{3}} & + & \leq_1 & - & \geq_{\mathbf{4}} & - & \leq_{\mathbf{5}} & - & \geq_2 & + & \geq_{\mathbf{6}} & + \\ + & & + & & + & & + & & + & & - & & - \end{bmatrix}$$

Since "$-$" appears as the top entry of the polarity columns that $\geq_{\mathbf{4}}$ and $\leq_{\mathbf{5}}$ were inserted into, reverse their directions:

$$\begin{bmatrix} + & \leq_3 & + & \leq_1 & - & \leq_4 & - & \geq_5 & - & \geq_2 & + & \geq_6 & + \\ + & & + & & + & & + & & + & & - & & - \end{bmatrix}$$

Remove the top entry of each polarity column:

$$[+ \leq_3 + \leq_1 + \leq_4 + \geq_5 + \geq_2 - \geq_6 -]$$

Call the info table above χ. Then $\chi \in \phi(\psi)$. Note that (A1) respects χ. We say that the black inequalities in χ are ***contributed by*** *ϕ, and the red inequalities are* ***contributed by*** *ψ.*

Moreover, by Definition 19 and Remark 5, we can reduce χ to

$$\chi' = [+ \leq + \geq + \geq -]$$

by using (App′). Note that we have reduced 4 flips in ψ to 2 flips in χ.

To get the entire set $\phi(\psi)$, we consider all possible insertions of the inequalities from ψ into ϕ.

https://doi.org/10.3390/logics4020006

Throughout the paper, we have used the term *below* (in a derivation tree) somewhat informally, since the details have not played an essential role. We now give a precise definition for the sake of the proofs below.

Definition A1. *Let x, y be nodes in a derivation. We say that x is* ***below*** *y if the distance from x to the conclusion is less than the distance from y to the conclusion and x, y are in the same branch.*

In the following discussion, for any polarity column v, let v^- denote v with its first (top) entry removed. Here is Lemma 2 reproduced.

Lemma A1. *Let $\phi' \sqsupseteq_1 \phi$, and let $\mathcal{D}$ be a derivation of $S \leqslant_\chi T$ from $\Gamma_\leqslant$ that contains $A \leqslant_\phi B$. If the subderivation of $A \leqslant_\phi B$ is replaced with $A \leqslant_{\phi'} B$, then there is a derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ (from $\Gamma_\leqslant$) for some $\chi' \sqsupseteq_1 \chi$ that differs from $\mathcal{D}$ at most by having each info table α below $A \leqslant_\phi B$ in $\mathcal{D}$ modified to some $\alpha' \sqsupseteq_1 \alpha$ in $\mathcal{D}'$ (but keeping the terms the same).*

Proof. By induction on the distance from $A \leqslant_\phi B$ to $S \leqslant_\chi T$ in $\mathcal{D}$.

There are three base cases: when $A \leqslant_\phi B$ is a premise of (Trans), when $A \leqslant_\phi B$ is in argument position (in which case we use the notation $A \leqslant_\psi B$), and when it is in function position (in which case we use the notation $F \leqslant_\phi G$).

The first case is clear. For the second case, assume we have

$$\begin{aligned} \psi' &= [\cdots *_{i-1}\, v_{i-1} \leq_i v_i \leq_{i+1} \cdots] \\ \psi &= [\cdots *_{i-1}\, v_{i-1} \qquad\quad \leq_{i+1} \cdots] \end{aligned}$$

and

$$\frac{\begin{array}{cc} \vdots & \vdots \\ F \leqslant_\phi G & A \leqslant_\psi B \end{array}}{F(A) \leqslant_\chi G(B)}\ (\text{App})$$

Then choose $\chi' \in \phi(\psi')$ so that it is obtained by

- Inserting $*_1, \ldots, *_{i-1}, \leq_{i+1}, \cdots *_n$ from ψ' into ϕ exactly where the corresponding inequalities from ψ were inserted into ϕ, respectively, to obtain $\chi \in \phi(\psi)$;
- Inserting $\leq_i$ from ψ' (the extra inequality) into ϕ at the same polarity column where $\leq_{i+1}$ was inserted.

More precisely, assuming we had

$$\begin{aligned} \phi &= [\cdots *_j \quad v_j \qquad\qquad\quad *_{j+1} \cdots] \\ \chi &= [\cdots *_j \cdots v_j^- \; *_{i+1}\, v_j^- \cdots *_{j+1} \cdots] \end{aligned}$$

where $*_{i+1}$ is the result of inserting $\leq_{i+1}$ from ψ into ϕ, then we choose

$$\chi' = [\cdots *_j \cdots v_j^- \; *_i\, v_j^- \; *_{i+1}\, v_j^- \cdots *_{j+1} \cdots]$$

with $*_i = *_{i+1}$, which means $\chi' \sqsupseteq_1 \chi$. Note that the argument is the same if we replace $\leq$ with $\geq$.

For the third case, assume we have

$$\begin{aligned} \phi' &= [\cdots *_{i-1}\, v_{i-1} \leq_i v_i \leq_{i+1} \cdots] \\ \phi &= [\cdots *_{i-1}\, v_{i-1} \qquad\quad \leq_{i+1} \cdots] \end{aligned}$$

and

$$\dfrac{\begin{array}{cc}\vdots & \vdots \\ F \leqslant_{\phi} G & A \leqslant_{\psi} B\end{array}}{F(A) \leqslant_{\chi} G(B)}\ (\text{App})$$

Then choose $\chi' \in \phi'(\psi)$ so that it is obtained by inserting every inequality from ψ into ϕ' exactly where it was inserted into ϕ to obtain $\chi \in \phi(\psi)$ (all the polarity columns available in ϕ are still available in ϕ'). It follows that $\chi' \sqsupseteq_1 \chi$.

The result follows by induction. □

Appendix A.1. Part I

Let $S \geqslant_1$ denote $S \leqslant_{\chi} T$ for some χ whose inequalities are all $\geq$. For Proposition 2, we prove the stronger result below.

Proposition A1. *If* $\Gamma \vdash_{\mathcal{M}} S \leq T$, *then* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} S \leqslant_1 T$ *and* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} T \geqslant_1 S$.

Proof. By induction on size.

The base case is by the definition of $\Gamma_{\leqslant}$. For the induction, consider a derivation from Γ whose last step uses (Anti):

$$\dfrac{\begin{array}{c}\vdots \\ A \leq B\end{array}}{G(B) \leq G(A)}\ (\text{Anti})$$

By the inductive hypothesis, $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} A \leqslant_{\psi} B$ and $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} B \leqslant_{\psi'} A$ where all of the inequalities in ψ are $\leq$ and all of the inequalities in ψ' are $\geq$. Then, since G is antitonic, we get derivations

$$\dfrac{\dfrac{}{G \leqslant_{[v_G \geq v_G]} G}\ (\text{Refl}) \qquad \begin{array}{c}\vdots \\ A \leqslant_{\psi} B\end{array}}{G(A) \leqslant_{\chi} G(B)}\ (\text{App})$$

where $\chi = [v_G^- \geq \cdots \geq v_G^-]$ and

$$\dfrac{\dfrac{}{G \leqslant_{[v_G \leq v_G]} G}\ (\text{Refl}) \qquad \begin{array}{c}\vdots \\ B \leqslant_{\psi'} A\end{array}}{G(B) \leqslant_{\chi'} G(A)}\ (\text{App})$$

where $\chi' = [v_G^- \leq \cdots \leq v_G^-]$.

The other cases can be verified in a similar manner. □

Appendix A.2. Part II

Definition A2. *Let* Φ, Ψ *be sets of info tables. We define*

$$\Phi + \Psi = \{\phi + \psi \colon \phi \in \Phi, \psi \in \Psi\}$$

Lemma A2. *Let* $\phi_1, \phi_2, \psi_1, \psi_2$ *be info tables. Then*

$$\phi_1(\psi_1) + \phi_2(\psi_2) \subseteq (\phi_1 + \phi_2)\Big(\psi_1 + \psi_2\Big)$$

Proof. This follows from the procedure in Example 2. The left-hand side involves inserting inequalities from ψ_1 into ϕ_1 and inserting inequalities from ψ_2 into ϕ_2, whereas the right-hand side involves inserting inequalities from ψ_1 into either ϕ_1 or ϕ_2, and likewise for ϕ_2. □

Theorem A1. *If* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} S \leqslant_{\chi} T$, *then* $\Gamma_{\leqslant} \vdash_{\mathcal{M}_\infty} S \leqslant_{\chi} T$ *via a subterm derivation.*

Proof. Revisit the proof of the claim in Theorem 1. We have an analogous setup here:

$$\dfrac{\dfrac{\dfrac{\ddots}{F_1 \leqslant_\phi F_{n-1}}\,(\text{Trans}) \qquad \dfrac{\ddots}{A_1 \leqslant_\psi A_{n-1}}\,(\text{Trans})}{F_1(A_1) \leqslant_\omega F_{n-1}(A_{n-1})}\,(\text{App}) \qquad \dfrac{F_{n-1} \leqslant_{\phi_n} F_n \qquad A_{n-1} \leqslant_{\psi_n} A_n}{F_{n-1}(A_{n-1}) \leqslant_{\omega_n} F_n(A_n)}\,(\text{App})}{F_1(A_1) \leqslant_\chi F_n(A_n)}\,(\text{Trans})$$

We claim we can manipulate this to get

$$\dfrac{\dfrac{\dfrac{\ddots}{F_1 \leqslant_\phi F_{n-1}}\,(\text{Trans}) \qquad F_{n-1} \leqslant_{\phi_n} F_n}{F_1 \leqslant_{\phi+\phi_n} F_n}\,(\text{Trans}) \qquad \dfrac{\dfrac{\ddots}{A_1 \leqslant_\psi A_{n-1}}\,(\text{Trans}) \qquad A_{n-1} \leqslant_{\psi_n} A_n}{A_1 \leqslant_{\psi+\psi_n} A_n}\,(\text{Trans})}{F_1(A_1) \leqslant_\chi F_n(A_n)}\,(\text{App})$$

Note that the conclusion of the new derivation has the same info table χ as in the original derivation.

We must verify that in the new derivation, χ is obtainable from $\phi + \phi_n$ and $\psi + \psi_n$ using (App). In the original derivation, $\omega \in \phi(\psi)$ and $\omega_n \in \phi_n(\psi_n)$, so $\chi \in \phi(\psi) + \phi_n(\psi_n)$. Now observe that the conclusion of the new derivation can have any info table in $(\phi + \phi_n)\Big(\psi + \psi_n\Big)$. By Lemma A2, $\chi \in (\phi + \phi_n)\Big(\psi + \psi_n\Big)$, so the last step is indeed a correct application of (App).

The rest of the proof is essentially identical to the proof of Theorem 1. □

Appendix A.3. Part III

For Lemma 5, we prove a stronger result by mutual induction. Let $\chi' \subseteq_1^v \chi$ denote that χ' is the result of deleting exactly one $v_{i-1} *_i$, where $v_{i-1} = v_i$, from χ.

Lemma A3. *(a) Let $\mathcal{D}$ be a derivation of $S \leqslant_\chi T$ that contains $A \leqslant_\psi B$ in argument position of (App), and let $\psi' \preceq_1 \psi$. If the subderivation of $A \leqslant_\psi B$ is replaced with $A \leqslant_{\psi'} B$, then there is a derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ for some $\chi' \subseteq_1^v \chi$ that differs from $\mathcal{D}$ at most by having each info table α below $A \leqslant_\psi B$ in $\mathcal{D}$ modified to some $\alpha' \subseteq_1^v \alpha$ in $\mathcal{D}'$.*

(b) Let $\mathcal{D}$ be a derivation of $S \leqslant_\chi T$ that contains $F \leqslant_\phi G$ in function position of (App), and let $\phi' \subseteq_1^v \phi$. If the subderivation of $F \leqslant_\phi G$ is replaced with $F \leqslant_{\phi'} G$, then there is a derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ for some $\chi' \subseteq_1^v \chi$ that differs from $\mathcal{D}$ at most by having each info table α below $F \leqslant_\phi G$ in $\mathcal{D}$ modified to some $\alpha' \subseteq_1^v \alpha$ in $\mathcal{D}'$.

As in Lemma A1, the main content of the proof is in the base case, stated below.

Lemma A4. *(a) Assume we have*

$$\dfrac{\begin{matrix}\vdots\\ F \leqslant_\phi G\end{matrix} \qquad \begin{matrix}\vdots\\ A \leqslant_\psi B\end{matrix}}{F(A) \leqslant_\chi G(B)}\,(\text{App})$$

and $\psi' \preceq_1 \psi$. Then we can get

$$\dfrac{\begin{matrix}\vdots\\ F \leqslant_\phi G\end{matrix} \qquad \begin{matrix}\vdots\\ A \leqslant_{\psi'} B\end{matrix}}{F(A) \leqslant_{\chi'} G(B)}\,(\text{App})$$

for some $\chi' \subseteq_1^v \chi$.

(b) Assume we have

$$\dfrac{\begin{matrix}\vdots\\ F \leqslant_\phi G\end{matrix} \qquad \begin{matrix}\vdots\\ A \leqslant_\psi B\end{matrix}}{F(A) \leqslant_\chi G(B)}\,(\text{App})$$

and $\phi' \subseteq_1^v \phi$. *Then we can get*

$$\frac{\begin{array}{cc}\vdots & \vdots \\ F \leqslant_{\phi'} G & A \leqslant_{\psi} B\end{array}}{F(A) \leqslant_{\chi'} G(B)}\ (\text{App})$$

for some $\chi' \subseteq_1^v \chi$.

Proof. (a) Assume we have

$$\begin{aligned}\psi &= [\cdots v_{j-1} *_j v_j \cdots] \\ \psi' &= [\cdots v'_{j-1} \quad \cdots]\end{aligned}$$

and

$$\chi = [\cdots v_i^- *_j v_i^- \cdots]$$

Then we can get

$$\chi' = [\cdots v_i^- \quad \cdots]$$

by inserting every inequality from ψ' into ϕ exactly where the corresponding inequality from ψ was inserted into ϕ to obtain $\chi \in \phi(\psi)$.

(b) Assume we have

$$\begin{aligned}\phi &= [\cdots v_{i-1} \quad *_i v_i \quad \cdots] \\ \phi' &= [\cdots \quad v_i \quad \cdots] \\ \chi &= [\cdots v_{i-1}^- *_j v_{i-1}^- \cdots *_n v_{i-1}^- *_i v_i^- *_{n+1} v_i^- \cdots]\end{aligned}$$

where $v_{i-1} = v_i$ and the red inequalities are contributed by ψ. Then we can get

$$\chi' = [\cdots v_i^- \quad *_j v_i^- \quad \cdots *_n \quad v_i^- *_{n+1} v_i^- \cdots]$$

by

- Inserting $*_j, \ldots, *_n$ at v_i in ϕ', instead of v_{i-1} which is now deleted;
- Inserting the remaining inequalities from ψ exactly where they were inserted into ϕ to obtain $\chi \in \phi(\psi)$.

The red inequalities are in the same direction in χ' as they were in χ because $v_{i-1} = v_i$. □

Here is Lemma 6 reproduced.

Lemma A5. *Let* $\mathcal{D}$ *be a derivation of the form*

$$\dfrac{\dfrac{\dfrac{\begin{array}{cc}\vdots & \vdots \\ F_l \leqslant_{\phi_l} G_l & A_l \leqslant_{\psi_l} B_l\end{array}}{F_l(A_l) \leqslant_{\omega_l} G_l(B_l)}\ (\text{App}) \quad \begin{array}{cc}\ddots & \\ & \vdots \end{array}}{\begin{array}{cc}F_1 \leqslant_{\phi_1} G_1 & A_1 \leqslant_{\psi_1} B_1\end{array}}\ (\text{App})}{F_1(A_1) \leqslant_{\omega_1} G_1(B_1)}\ (\text{App})$$

$$\vdots$$

$$S \leqslant_{\chi} T$$

where the zigzags in the path downward from $F_l \leqslant_{\phi_l} G_l$ are all in either function position or a premise of (Trans), and l is the number of such zigzags that are in function position. Then there exists $\phi' \sqsubseteq \phi_l$ with at most $f(\phi_1) + i(\psi_1) + \cdots + f(\phi_l) + i(\psi_l)$ inequalities such that if the subderivation of $F_l \leqslant_{\phi_l} G_l$ in $\mathcal{D}$ is replaced with $F_l \leqslant_{\phi'} G_l$, then there is a derivation $\mathcal{D}'$ of $S \leqslant_{\chi'} T$ for some $\chi' \sqsubseteq \chi$ that differs from $\mathcal{D}$ at most by having each info table α below $F_l \leqslant_{\phi_l} G_l$ in D modified to some $\alpha' \sqsubseteq \alpha$ in $\mathcal{D}'$.

Sketch. Intuitively, we define ϕ' to include every $*_i\, v_i$ from ϕ_l for which $*_i \neq *_{i+1}$ (which the deletion rule does not permit deleting) and every $*_i\, v_i$ at which any of $\psi_1, \ldots, \psi_l$ inserted at to get $\omega_1, \ldots, \omega_l$, respectively. □

Call the modification of a portion of a derivation **structure-preserving** if it satisfies the above property, i.e., each α becomes some $\alpha' \sqsubseteq \alpha$.

Proof. By induction on l. For the base case, assume we have

$$\dfrac{\begin{array}{cc}\vdots & \vdots \\ F \leqslant_\phi G & A \leqslant_\psi B\end{array}}{F(A) \leqslant_\omega G(B)}\ (\text{App})$$
$$\vdots$$
$$S \leqslant_\chi T$$

with the form described in the lemma, and say $\phi = [v_0 *_1 v_1 \cdots *_n v_n]$. WLOG, assume the deletion rule was not used at this instance of (App). Define ϕ' to be the subsequence of ϕ that contains:

- v_0;
- Every $*_i\, v_i$ where $*_i \neq *_{i+1}$ (including $*_n\, v_n$);
- Every $*_i\, v_i$ at which ψ inserted into to get $\omega \in \phi(\psi)$.

Indeed, $\phi' \sqsubseteq \phi$, since we have not deleted any $*_i\, v_i$ where $*_i \neq *_{i+1}$. Then choose $\omega' \in \phi'(\psi)$ so that it is obtained by inserting every inequality from ψ into ϕ' exactly where it was inserted in ϕ to obtain ω (we kept those columns). Then ω' is the result of deleting every $*_i\, v_i^-$ from ω for which $*_i\, v_i$ was removed from ϕ to get ϕ', so $\omega' \sqsubseteq \omega$. Since the portion of the derivation below $F(A) \leqslant_\omega G(B)$ consists of premises of (Trans) only, it follows that ω' induces a structure-preserving modification of that portion, resulting in a new derivation of $S \leqslant_{\chi'} T$ for $\chi' \sqsubseteq \chi$. Moreover, we see from the definition of ϕ' that it has at most $f(\phi) + i(\psi)$ inequalities.

Now, for the induction, consider

$$\dfrac{\dfrac{\begin{array}{cc}\vdots & \vdots \\ F_l \leqslant_{\phi_l} G_l & A_l \leqslant_{\psi_l} B_l\end{array}}{F_l(A_l) \leqslant_{\omega_l} G_l(B_l)}\ (\text{App})}{\vdots}$$
$$\dfrac{\dfrac{\begin{array}{cc}\ddots & \vdots \\ F_{l-1} \leqslant_{\phi_{l-1}} G_{l-1} & A_{l-1} \leqslant_{\psi_{l-1}} B_{l-1}\end{array}}{F_{l-1}(A_{l-1}) \leqslant_{\omega_{l-1}} G_{l-1}(B_{l-1})}\ (\text{App})}{\ddots}$$
$$S \leqslant_\chi T$$

By the inductive hypothesis, there is some $\phi'_{l-1} \sqsubseteq \phi_{l-1}$ with at most $f(\phi_1) + i(\psi_1) + \cdots + f(\phi_{l-1}) + i(\psi_{l-1})$ inequalities that induces a structure-preserving modification of the portion below $F_{l-1} \leqslant_{\phi_{l-1}} G_{l-1}$. Now, since the portion below $F_l(A_l) \leqslant_{\omega_l} G_l(B_l)$ to $F_{l-1} \leqslant_{\phi_{l-1}} G_{l-1}$ only consists of premises of (Trans), there is $\omega'_l \sqsubseteq \omega_l$, where $\omega'_l \subseteq \phi'_{l-1}$, that induces a structure-preserving modification of that portion such that it concludes with

$F_{l-1} \leqslant_{\phi'_{l-1}} G_{l-1}$. So it suffices to find $\phi'_l \sqsubseteq \phi_l$ with at most $f(\phi_1) + i(\psi_1) + \cdots + f(\phi_l) + i(\psi_l)$ inequalities that will allow us to derive $F_l(A_l) \leqslant_{\omega'_l} G_l(B_l)$.

Say $\phi_l = [v_0 *_1 v_1 \cdots *_n v_n]$. Define ϕ'_l to be the subsequence of ϕ_l that contains:

- v_0;
- Every $*_i v_i$ where $*_i \neq *_{i+1}$ (including $*_n v_n$);
- Every $*_i v_i$ at which ψ_l inserted into to get $\omega_l \in \phi_l(\psi_l)$;
- Every $*_i v_i$ for which ϕ_l contributed $*_i$ to ω'_l (technically ω'_l was not explicitly obtained by applying the procedure from Example A1 on $\phi_l(\psi_l)$, but we can view ω'_l as a copy of ω_l with fewer entries, so we can speak of $*_i$ in ω'_l being contributed by ϕ_l).

As in the case base, $\phi'_l \sqsubseteq \phi_l$, and similarly, we choose $\omega''_l \in \phi'_l(\psi_l)$ using the same procedure as described above so that $\omega''_l \sqsubseteq \omega_l$. By construction, $\omega''_l \supseteq \omega'_l$, so it follows that $\omega''_l \sqsupseteq \omega'_l$ (since we also have $\omega'_l \sqsubseteq \omega_l$). In other words, the $*_i v_i$ that appear in ω''_l but not ω'_l are deletable because if they were not, then they also would not have been deletable in ω'_l.

Thus, since $F_l(A_l) \leqslant_{\omega''_l} G_l(B_l)$ is derivable (by definition of ω''_l), it follows that $F_l(A_l) \leqslant_{\omega'_l} G_l(B_l)$ is derivable. Moreover, we see that in the definition of ϕ'_l, the second and third clauses contribute at most $f(\phi_l) + i(\psi_l)$ inequalities, and the last clause contributes at most $f(\phi_1) + i(\psi_1) + \cdots + f(\phi_{l-1}) + i(\psi_{l-1})$ inequalities, as $\omega'_l \subseteq \phi'_{l-1}$. □

References

1. Moss, L. Natural Logic. In *The Handbook of Contemporary Semantic Theory*, 2nd ed.; Lappin, E., Fox, C., Eds.; Wiley-Blackwell: Chichester, UK, 2015; pp. 561–592.
2. Icard, T.; Moss, L. Recent Progress on Monotonicity. *Linguist. Issues Lang. Technol.* **2014**, *9*, 167–194. [CrossRef]
3. Icard, T.; Moss, L.; Tune, W. A Monotonicity Calculus and Its Completeness. In *Proceedings of the 15th Meeting on the Mathematics of Language*; Association for Computational Linguistics: London, UK, 2017; pp. 75–87. [CrossRef]
4. Moss, L.; Icard, T. An Undecidability Result for Monotonicity Inference. Indiana University, Bloomington, IN, USA, 2024, *manuscript in preparation*.
5. Kozen, D. Complexity of Finitely Presented Algebras. In *Proceedings of the 9th Symposium on Theory of Computing*; Association for Computing Machinery: New York, NY, USA, 1977; pp. 165–177.
6. Fyodorov, Y.; Winter, Y.; Francez, N. Order-based Inference in Natural Logic. *Log. J. IGPL* **2003**, *11*, 385–416. [CrossRef]
7. Kennedy, A.; Pierce, B. On Decidability of Nominal Subtyping with Variance. In Proceedings of the 2007 International Workshop on Foundations and Developments of Object-Oriented Languages, Nice, France, 20 January 2007. Available online: http://foolwood07.cs.uchicago.edu/ (accessed on 13 May 2026).
8. Grigore, R. Java Generics are Turing Complete. In *POPL '17: Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages*; Association for Computing Machinery: New York, NY, USA, 2017; pp. 73–85.
9. Dowty, D. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*; D. Reidel Publishing Company: Dordrecht, The Netherlands, 1979.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.