Validating Back-links of FOLID Cyclic Pre-proofs

Sorin Stratulat

Validating Back-links of $\mathrm{FOL_{ID}}$ Cyclic Pre-proofs

Sorin Stratulat
Université de Lorraine, CNRS, LORIA
F-57000 Metz, France
sorin.stratulat@univ-lorraine.fr

Cyclic pre-proofs can be represented as sets of finite tree derivations with back-links. In the frame of the first-order logic with inductive definitions ($\mathrm{FOL_{ID}}$), the nodes of the tree derivations are labelled by sequents and the back-links connect particular terminal nodes, referred to as *buds*, to other nodes labelled by a same sequent. However, only some back-links can constitute sound pre-proofs. Previously, it has been shown that special ordering and derivability conditions, defined along the minimal cycles of the digraph representing a particular normal form of the cyclic pre-proof, are sufficient for validating the back-links. In that approach, a same constraint could be checked several times when processing different minimal cycles, hence one may require additional recording mechanisms to avoid redundant computation in order to downgrade the time complexity to polynomial.

We present a new approach that does not need to process minimal cycles. It based on a normal form that allows to define the validation conditions by taking into account only the root-bud paths from the non-singleton strongly connected components of its digraph.

1 Introduction

In [4, 5, 7], Brotherston and Simpson introduced the notion of cyclic (pre-)proof in the frame of first-order logic with inductive definitions (for short $\mathrm{FOL_{ID}}$ and detailed, e.g., in [1]) and including equality. In this setting, the cyclic *pre-proofs* are sequent-based proof derivations usually presented in the form of finite trees. Some of their terminal nodes, called *buds*, are labelled by 'not-yet proved' sequents that already labelled other nodes, called *companions*. For each bud there is only one companion and the bud-companion relations are referred to as *back-links*.

Not all back-links may constitute sound pre-proofs. Indeed, a pre-proof can be constructed for any false sequent S by applying a stuttering inference step[1] that creates a copy of S. This terminal node is a bud whose companion is the root of the pre-proof. [5, 7] also introduced the CLKID^{ω} inference system for building cyclic pre-proofs and defined a sufficient criterion for checking their soundness in terms of a *global trace condition*. This condition is an ω-regular property that can be checked as an inclusion between two Büchi automata. The inclusion test includes an automata complementation procedure [9] whose time complexity is exponential in the number of states of the automaton to be complemented.

A more effective soundness criterion was given in [11] for pre-proofs generated by $\mathrm{CLKID}^{\omega}_{N}$, a restricted version of CLKID^{ω}. Inspired from a previous method [10, 12] for checking the soundness of cyclic proofs built using the Noetherian induction principle for reasoning on conditional specifications, its time complexity can be downgraded to polynomial. To do this, a $\mathrm{CLKID}^{\omega}_{N}$ pre-proof is normalised to some set of finite tree derivations which can be represented as a directed graph (for short, digraph) having some arrows labelled by substitutions. The soundness criterion asks that some derivability and ordering constraints hold along the paths leading root nodes to bud nodes in the *minimal* cycles of the digraph, i.e., cycles that do not include other cycles.

[1] for example, by applying the LK's (*Subst*) rule with an identity substitution (see the definition of (*Subst*) in Definition 2).

Submitted to:
CL&C'18

In general, the number of minimal cycles in a digraph with n nodes can be much greater than the number of its buds (which is always smaller than n). For complete digraphs, i.e., digraphs for which every pair of distinct nodes is connected by arrows in the two ways, one can define the number of minimal cycles built by $k \in [2..n]$ nodes, as follows. We take one of the n nodes as the starting node in the cycle, then the next one from the remaining $n-1$ nodes, and so on for $k-1$ times. So there are $n \times (n-1) \times \ldots \times (n-k+1)$ ways to do it. Since the cycle consisting of the k nodes can be built in k different times, depending which is the starting node among its nodes, this number is $\frac{n!}{(n-k)!k}$. Hence, the total number of minimal cycles in a complete digraph with n nodes is

$$\sum_{k=2}^{n} \frac{n!}{(n-k)!k}$$

Fortunately, the number of arrows in any digraph built with the approach from [11] is smaller than that for the complete digraphs because each bud node has only one companion. However, a ordering-derivability constraint can be checked several times as it may be defined w.r.t. different minimal cycles. In [11], it was already suggested that their number can be reduced to the number of buds from the minimal cycles, hence smaller than n. This redundancy can be avoided, for example, by using recording mechanisms.

In this paper, we present an improved version of the soundness criterion for validating $\text{CLKID}^{\omega}_{N}$ pre-proofs. The advantage is that the computation of minimal cycles is not needed and there is no redundancy in the computation of the ordering-derivability constraints. In order to do this, we propose a new normal form of $\text{CLKID}^{\omega}_{N}$ pre-proofs and define ordering and derivability constraints for every root-bud path that occurs in a non-singleton strongly connected component (SCC) of the digraphs associated to the new normal forms. We show that the number of constraints is that of the buds from the non-singleton SCCs.

The rest of the paper is structured as follows. Section 2 gives a brief presentation of FOL_{ID} and $\text{CLKID}^{\omega}_{N}$. Section 3 introduces the soundness criterion for $\text{CLKID}^{\omega}_{N}$ pre-proofs by detailing the normalisation procedure, the digraph construction and the definition of the ordering and derivability conditions. A comparison is made with the soundness criterion from [11]. The conclusions and future work are given in the last section.

2 Induction-based sequent calculus

Syntax. The logical setting is that presented in [7], based on FOL_{ID} with equality using a standard (countable) first-order language Σ. The predicate symbols are labelled either as *ordinary* or *inductive*, and we assume that there is an arbitrary but finite number of inductive predicate symbols. The terms are defined as usual. By $\bar{t}$, we denote a vector of terms $(t_1, \ldots, t_n)$ of length n, the value of n being usually deduced from the context.

New terms and formulas are built by instantiating variables by terms via substitutions. A *substitution* is a mapping from variables to terms, of the form $\{x_1 \mapsto t_1; \ldots; x_p \mapsto t_p\}$, for some $p > 0$, which can be written in a more compact form as $\{\bar{x} \mapsto \bar{t}\}$, where $\bar{x} \equiv (x_1, \ldots, x_p)$, $\bar{t} \equiv (t_1, \ldots, t_p)$, and $\equiv$ is the syntactic equality. The *composition* of σ_1 with σ_2 is denoted by $\sigma_1\sigma_2$, for all substitutions σ_1 and σ_2. A term t is an *instance* of t', or t *matches* t', if there is a substitution σ such that $t \equiv t'\sigma$. Similarly, the notion of matching can be extended to vector of terms, atoms, and formulas. For any substitution σ applied to a formula F, we use the notation $F[\sigma]$ instead of $F\sigma$.

Deductive sequent-based inference rules. The proof derivations are built from sequents [8] of the form $\Gamma \vdash \Delta$, where Γ and Δ are finite multisets of formulas called *antecedents* and *succedents*, respectively. $FV(\Gamma \vdash \Delta)$ denotes its set of free variables. An inference rule is represented by a horizontal line followed by the name of the rule. The line separates the lower sequent, called *conclusion*, from a (potentially empty) multiset of upper sequents, called *premises*. Most of the rules *introduce* an explicitly represented formula from the conclusion, called *principal* formula. In this case, the rules are annotated by L (resp., R) if the rule is introduced on the left (resp, right) of the $\vdash$ symbol from the conclusion.

A specification is built from a finite inductive definition set of axioms Φ consisting of formulas of the form

$$\bigwedge_{m=1}^{h} Q_m(\bar{u}_m) \wedge \bigwedge_{m=1}^{l} P_m(\bar{t}_m) \Rightarrow P(\bar{t}), \tag{1}$$

where h, l are naturals, $Q_1, \ldots, Q_h$ are ordinary predicate symbols, $P_1, \ldots, P_l, P$ are inductive predicate symbols. $\bigwedge_{m=1}^{0}$ is a shortcut for the 'true' boolean constant and can be ignored.

The deductive part of the sequent-based reasoning about FOL_{ID} is performed using the Gentzen's LK rules [8] and an 'unfold' rule. The unfold rule $(R.(rname))$ replaces an atom $P(\bar{t}')$ using the axiom $(rname)$ defining P. E.g., (1) can be applied on $\Gamma \vdash P(\bar{t}'), \Delta$ if $P(\bar{t}') \equiv P(\bar{t})[\sigma]$ for some substitution σ, as:

$$\frac{seq\_Q\_inst \quad seq\_P\_inst}{\Gamma \vdash P(\bar{t}'), \Delta}\ (R.(1)) \ ,$$

where *seq_Q_inst* (resp., *seq_P_inst*) is the multiset of sequents $\bigcup_{m=1}^{h}\{\Gamma \vdash Q_m(\bar{u}_m)[\sigma], \Delta\}$ (resp., $\bigcup_{m=1}^{l}\{\Gamma \vdash P_m(\bar{t}_m)[\sigma], \Delta\}$).

Semantics. The standard interpretation of inductive predicates is built from prefixed points of a monotone operator issued from the set of axioms representing Φ [1]. Its least prefixed point, approached by an iteratively built *approximant* sequence, helps defining a *standard model* for (Σ, Φ) (see, e.g., [7] for details).

Definition 1 (validity)**.** *Let M be a standard model for (Σ, Φ), $\Gamma \vdash \Delta$ a sequent and ρ a valuation which interprets in M the variables from $FV(\Gamma \vdash \Delta)$. We write $\Gamma \models^{M}_{\rho} \Delta$ if every $G \in \Gamma$ holds in M there is some $D \in \Delta$ that also holds in M. We say that $\Gamma \vdash \Delta$ is M-*true *and write $\Gamma \models^{M} \Delta$ if $\Gamma \models^{M}_{\rho} \Delta$, for any ρ.*

When M is implicit from the context, we use *true* instead of M-true. A rule is *sound*, or preserves the validity, if its conclusion is true whenever its premises are true. Hence, the conclusion of every 0-premise sound rule is true.

2.1 The CLKID$^{\omega}_{N}$ cyclic inference system

CLKID$^{\omega}$ [7] includes the LK rules, the rules from Figure 1 that process equalities, the 'unfold' rule and the $(Case)$ rule which represents a left-introduction operation for inductive predicate symbols:

$$\frac{\text{case distinctions}}{\Gamma, P(\bar{t}') \vdash \Delta}\ (Case\ P(\bar{t}'))$$

For each axiom of the form (1),

$$\Gamma, \bar{t}' = \bar{t}, Q_1(\bar{u}_1), \ldots, Q_h(\bar{u}_h), P_1(\bar{t}_1), \ldots, P_l(\bar{t}_l) \vdash \Delta \tag{2}$$

$$\frac{}{\Gamma \vdash t = t, \Delta}\ (=R) \qquad\qquad \frac{\Gamma[\{x \mapsto u; y \mapsto t\}] \vdash \Delta[\{x \mapsto u; y \mapsto t\}]}{\Gamma[\{x \mapsto t; y \mapsto u\}], t = u \vdash \Delta[\{x \mapsto t; y \mapsto u\}]}\ (=L)$$

Figure 1: Sequent-based rules for equality reasoning.

is the *case distinction* for which each free variable y from (1) is fresh w.r.t. the *free variables* from the conclusion of the rule (y can be renamed to a fresh variable, otherwise). $P_1(\bar{t}_1), \ldots, P_l(\bar{t}_l)$ are *case descendants* of $P(\bar{t}')$.

The inference system $\text{CLKID}^{\omega}_{N}$, introduced in [11], is the restricted version of CLKID^{ω} for which $(=L)$ is replaced by the generalization rule (Gen) that substitutes a term by a variable:

$$\frac{\Gamma'[\{t \mapsto u\}] \vdash \Delta'[\{t \mapsto u\}]}{\Gamma', t = u \vdash \Delta'}\ (Gen)$$

where t is a free variable that does not occur in u.

(Gen) is the particular instance of $(=L)$ from Figure 1 when $y \notin FV(\Gamma \vdash \Delta)$ and $t \notin FV(\Gamma \vdash \Delta)$ that also does not occur in u. By using the property that $\Phi[\{y \mapsto u\}] \equiv \Phi$, holding whenever y is a free variable not occurring in a formula Φ, the last condition can simplify $(=L)$ to a form equivalent to (Gen):

$$\frac{\Gamma[\{x \mapsto u\}] \vdash \Delta[\{x \mapsto u\}]}{\Gamma[\{x \mapsto t\}], t = u \vdash \Delta[\{x \mapsto t\}]}\ (=L)$$

$\text{CLKID}^{\omega}_{N}$ pre-proof trees. A *derivation tree* for a sequent S is built by successively applying inference rules starting from S. We consider only finite derivation trees whose terminal nodes can be either leaves or buds. The *leaves* are labelled by sequents that represent conclusions of 0-premise rule, e.g., the unfold rule using unconditional axioms. For each *bud* there is another node, called *companion* and having the same sequent labelling. The bud and its companion are annotated by the same sign, e.g., †. In addition, the buds having a same companion are labelled by the sign followed by a number that makes them unique, e.g., †1, †2,.... A *back-link* is a relation bud-companion.

Notation 1 (pre-proof tree, induction function for tree)**.** *The pair ($\mathscr{D}$, $\mathscr{R}$) denotes a* pre-proof tree*, where $\mathscr{D}$ is a finite derivation tree and $\mathscr{R}$ is a defined* induction function *assigning a companion to every bud in $\mathscr{D}$.*

Example 1. *To highlight the changes w.r.t. [11], we take the same running example (also presented in [6]). Let N and R be two inductive predicates defined by:*

$$\Rightarrow R(0,y) \quad (5)$$

$$\Rightarrow N(0) \quad (3) \qquad R(x,0) \Rightarrow R(sx,0) \quad (6)$$

$$N(x) \Rightarrow N(sx) \quad (4) \qquad R(ssx,y) \Rightarrow R(sx,sy) \quad (7)$$

where the parentheses around the argument of s are omitted. One can build the following pre-proof of $N(x), N(y) \vdash R(x,y)$:

$$\dfrac{\dfrac{}{\vdash R(0,0)}\,(R.(5)) \quad \dfrac{\dfrac{Nx' \vdash R(x',0)\ (\dagger\mathbf{1})}{Nx'' \vdash R(x'',0)}\,(Subst)}{Nx'' \vdash R(sx'',0)}\,(R.(6))}{\dfrac{\underline{Nx'} \vdash R(x',0)\ (\dagger)}{Nx' \vdash R(sx',0)}\,(R.(6))}\,(Case\ N) \quad \dfrac{\dfrac{Nx,Ny \vdash R(x,y)\ (*\mathbf{1})}{Nssx',Ny' \vdash R(ssx',y')}\,(Subst)}{\dfrac{Nx',Ny' \vdash R(ssx',y')}{Nx',Ny' \vdash R(sx',sy')}\,(R.(7))}\,(Cut)$$

$$\dfrac{\dfrac{}{Ny \vdash R(0,y)}\,(R.(5)) \quad \dfrac{}{Nx',\underline{Ny} \vdash R(sx',y)}\,(Case\ N)}{\underline{Nx},Ny \vdash R(x,y)\ (*)}\,(Case\ N)$$

where the double line means that (Gen) *was applied after* $(Case)$ *and the principal formulas of the* $(Case)$ *steps are underlined.* (Cut) *is applied as in [6]. For lack of space, the parentheses around the argument of N are omitted.*

We denote by $S(N)$ the sequent labelling any node N. A *path* is a list $[N^0, N^1, \ldots]$ of nodes in a pre-proof tree such that, for all $i \geq 0$, $S(N^{i+1})$ is either one of the premises of the rule applied on $S(N^i)$ if N^i is not a terminal node, or $S(\mathscr{R}(N^i))$ if N^i is a bud.

Definition 2 (Trace, Progress point [11])**.** *Let* $(\mathscr{D}, \mathscr{R})$ *be a CLKID*$^\omega_N$ *pre-proof tree and let* $[N^0, N^1, \ldots]$ *be one of its infinite paths and denoted by l. A* trace *following l is a sequence* $(\tau_i)_{i\geq 0}$ *of inductive antecedent atoms (IAAs) such that, for all i, we have that* N^i *is labelled by* $\Gamma_i \vdash \Delta_i$ *and:*

1. τ_i *is some* $P_{j_i}(\overline{t_i}) \in \Gamma_i$*;*
2. *if* $\Gamma_i \vdash \Delta_i$ *is the conclusion of* $(Subst)$ *then* $\tau_i = \tau_{i+1}[\theta]$*, where* θ *is the substitution used by the LK's* $(Subst)$ *rule defined as:*

$$\dfrac{\Gamma \vdash \Delta}{\Gamma[\theta] \vdash \Delta[\theta]}\,(Subst)$$

3. *if* $\Gamma_i \vdash \Delta_i$ *is the conclusion of* (Gen) *having* $t = u$ *as principal formula, there is a formula F such that* $\tau_i = F$ *and* $\tau_{i+1} = F[\{t \mapsto u\}]$*;*
4. *if* $\Gamma_i \vdash \Delta_i$ *is the conclusion of a* $(Case)$ *rule then either a)* $\tau_{i+1} = \tau_i$*, if* τ_i *is not the principal formula of the rule instance, or b)* τ_i *is the principal formula and* τ_{i+1} *is a case descendant of* τ_i*. In the latter case, i is said to be a* progress point *of the trace;*
5. *if* $\Gamma_i \vdash \Delta_i$ *is the conclusion of any other rule then* $\tau_{i+1} = \tau_i$*.*

Remark 1. *Non-equality relations between (instances of)* τ_i *and* τ_{i+1} *in the above definition are possible only when i is a progress point.*

Remark 2. *Condition 3 is an abbreviated form of the case dealing with* $(= L)$ *in Definition 5.4 from [7], by applying the discussed restrictions to* $(= L)$*, i.e., if* $\Gamma_i \vdash \Delta_i$ *is the conclusion of* $(= L)$*, of the form* $\Gamma[\{x \mapsto t; y \mapsto u\}], t = u \vdash \Delta[\{x \mapsto t; y \mapsto u\}]$ *and having* $t = u$ *as principal formula, there is a formula* F' *such that* $\tau_i = F'[\{x \mapsto t; y \mapsto u\}]$ *and* $\tau_{i+1} = F'[\{x \mapsto u; y \mapsto t\}]$ *under the following conditions:* $y \notin FV(\Gamma_i \backslash \{t = u\} \vdash \Delta_i)$*, t is a free variable not occurring in u and* $t \notin FV(\Gamma \vdash \Delta)$*.*

An *infinitely progressing* trace is a trace with infinitely many progress points.

3 The criterion for validating the soundness of CLKID$^\omega_N$ pre-proofs

The proof that some cyclic pre-proof is sound is done by using a *Descente Infinie* argument. The general technique is to assume, by contradiction, that the root of a pre-proof is labelled by a false sequent. Then,

we have to show that there is an infinite path of nodes in the pre-proof for which there is an infinite progressing trace following some tail of it. This means that all successive steps in the tail are decreasing and the steps corresponding to the progress points are strictly decreasing w.r.t. some semantic ordering over ordinals. We get a contradiction because it is not possible to built an infinite strictly decreasing sequence of ordinals.

Since the inference rules are sound, an infinite path of nodes labelled by false sequents should exist in the pre-proof whenever its root sequent is false. A sufficient criterion for validating the soundness of CLKID^{ω} pre-proofs is the *global trace condition* [4, 5, 7]: for every infinite path, there is an infinitely progressing trace following some tail. A different sufficient criterion for validating the soundness of $\text{CLKID}^{\omega}_{N}$ pre-proofs was given in [11]; it defines ordering and derivability conditions to be satisfied by the digraph representing some normal form of the pre-proof. The normalisation procedure transforms the pre-proof into a set of pre-proof trees, for short *pre-proof tree-sets*, such that the root of the pre-proof is among the roots of the trees from the normal form. If the sequent labelling the root of the pre-proof is false, one can build an infinite path in the digraph, whose nodes are labelled by false sequents and for which there is an infinite progressing trace following some tail of it.

In the following, we present an improved version of the criterion from [11].

3.1 Normalising pre-proof trees

The normalisation process consists in the exhaustive application of the following three operations. The first operation applies on an internal node labelled by some premise of (Subst), of the form

$$\dfrac{\dfrac{\vdots}{\Gamma \vdash \Delta}}{\Gamma[\sigma] \vdash \Delta[\sigma]}\,(\mathit{Subst})$$
$$\vdots$$

The result is displayed in Figure 2. The internal node is duplicated and the subtree derivation rooted by it is detached to become a new tree derivation. At the end, we get two distinct pre-proof trees. The two occurrences of the duplicated node establish a new bud-companion relation.

$$\dfrac{\Gamma \vdash \Delta\ (*1)}{\Gamma[\sigma] \vdash \Delta[\sigma]}\,(\mathit{Subst}) \qquad\qquad \dfrac{\vdots}{\Gamma \vdash \Delta\ (*)}\ \text{(new tree)}$$
$$\vdots$$

Figure 2: The result of the first operation.

The second operation applies on a non-root companion which is duplicated and the subtree derivation rooted by it becomes a new pre-proof tree. The result is displayed in Figure 3. The sequent labelling the copy of the companion $(*)$ becomes the conclusion of a new (Subst) rule. The substitution used by the new (Subst) rule is chosen such that its premise labels a new bud node labelled by the same sequent as the conclusion, e.g., the *empty* substitution. The new bud node will have (*) assigned as companion.

The last operation applies on a bud node labelled by some sequent that is the premise of a rule r different from (Subst) such that

$$\dfrac{\dfrac{\Gamma \vdash \Delta\ (*1)}{\Gamma \vdash \Delta}\ (\mathit{Subst})}{\vdots} \qquad \begin{array}{c}\vdots\\ \Gamma \vdash \Delta\ (*)\\ \text{(new tree)}\end{array}$$

Figure 3: The result of the second operation.

$$\dfrac{\dfrac{\Gamma \vdash \Delta\ (*1)}{\Gamma' \vdash \Delta'}\ r}{\vdots} \qquad \text{is transformed to} \qquad \dfrac{\dfrac{\Gamma \vdash \Delta\ (*1)}{\Gamma \vdash \Delta}\ (\mathit{Subst})}{\dfrac{\Gamma \vdash \Delta}{\Gamma' \vdash \Delta'}\ r}\ \vdots$$

Let $(*)$ denote the companion of the bud node. A new application of (Subst) with the empty substitution was performed on the bud sequent such that the node labelled by its premise becomes the new bud node whose companion is $(*)$.

Compared with the normalisation procedure from [11], the two procedures share only the first operation. The procedure from [11] also includes an operation that applies on non-root companions but does not include the (Subst)-step from Figure 3. It does not have an equivalent transformation for the third operation.

The following properties, related to the normalisation process and the resulted normal form as given by Lemmas 1 and 2, are satisfied.

Lemma 1 (termination)**.** *The normalisation process terminates.*

Proof. The number of nodes that can be processed by the three operations is finite, for every pre-proof tree. In addition, it decrements after applying each operation. □

The induction function is extended to allow new bud-companion relations between nodes from different pre-proof trees.

Definition 3 (rb-path, IH-node)**.** *An* rb-path *is a path of the form* $[R,\ldots,H,B]$ *that leads the root R to a bud B in some pre-proof tree of a pre-proof tree-set such that B is the only bud in the path. We will call H an inductive hypothesis node (for short,* IH-node*).*

A path in a pre-proof tree-set $(\mathscr{MD}, \mathscr{MR})$ is a list $[N^0, N^1, \ldots]$ of nodes in $\mathscr{MD}$ such that, for all $i \geq 0$, $S(N^{i+1})$ is one of the premises of the rule applied on $S(N^i)$ if N^i is an internal node, or $S(\mathscr{MR}(N^i))$ if N^i is a bud.

Lemma 2. *The normalisation of any pre-proof* $(\mathscr{D},\mathscr{R})$ *of a sequent S builds a pre-proof tree-set* $(\mathscr{MD}, \mathscr{MR})$

1. *that has a pre-proof tree rooted by a node labelled by S, and*
2. *for which each of its rb-paths* $[R,\ldots,B]$ *has B as the only node that is labelled by the premise of a* (Subst) *rule. A node is a* (Subst)*-node if and only if it is an* (IH)*-node.*

Proof. Claim 1) holds because the first operation duplicates only non-root nodes and the third operation expands bud nodes, so the root nodes do not change. If S labels the root node of a pre-proof tree t having a non-root companion n, t will be processed by the second operation applied on n but will still have its root labelled by S.

Claim 2) holds by the construction of the normal forms. □

Example 2. *The second operation can be applied on the non-root companion from Example 1, denoted by* $(*)$*, to give the following normalised pre-proof tree-set:*

$$\dfrac{\dfrac{}{Ny \vdash R(0,y)}\,(R.(5)) \qquad \dfrac{\dfrac{\dfrac{\underline{Nx'} \vdash R(x',0)\ (\dagger\mathbf{1})}{\underline{Nx'} \vdash R(x',0)}\,(Subst)}{Nx' \vdash R(sx',0)}\,(R.(6)) \qquad \dfrac{\dfrac{Nx, Ny \vdash R(x,y)\ (*)}{Nssx', Ny' \vdash R(ssx',y')}\,(Subst)}{Nx', Ny' \vdash R(ssx',y')}\,(Cut)}{Nx', \underline{Ny} \vdash R(sx',y)}\,(R.(7))\ (Case\ N)}{\underline{Nx}, Ny \vdash R(x,y)\ (*)}\,(Case\ N)$$

$$\dfrac{\dfrac{}{\vdash R(0,0)}\,(R.(5)) \qquad \dfrac{\dfrac{Nx' \vdash R(x',0)\ (\dagger)}{Nx'' \vdash R(x'',0)}\,(Subst)}{Nx'' \vdash R(sx'',0)}\,(R.(6))}{\underline{Nx'} \vdash R(x',0)\ (\dagger)}\,(Case\ N)$$

3.2 Building the digraph of a pre-proof tree-set

Any pre-proof tree-set can also be represented as a *digraph* of sequents built from the nodes of its tree-set. The digraph associated to a pre-proof tree-set $(\mathscr{MD}, \mathscr{MR})$ is crucial in our setting to check whether $(\mathscr{MD}, \mathscr{MR})$ is a proof tree-set. Its edges are arrows built as follows:

- a *forward* arrow leads a node N^1 to a node N^2 if there is a rule that was applied on the sequent labelling N^1 and the sequent labelling N^2 is a premise of the rule;
- a *back-link* (or backward arrow) starts from a bud and ends to its companion.

Some arrows will be annotated by substitutions. Each forward arrow, starting from a (Gen)-node whose principal formula is $x = u$, is annotated by the *equality substitution* $\{x \mapsto u\}$. The forward arrow starting from a node N that is different from (Gen)- and $(Subst)$-nodes is annotated with the *identity substitution* for $S(N)$, which maps the free variables from $S(N)$ to themselves. Finally, the forward arrows starting from $(Subst)$-nodes and the back-links are not annotated. They help to build infinite paths but do not play any role when defining the soundness constraints.

By abuse of notation, a *path* in a digraph is a (potentially infinite) list of nodes built by following the arrows in the digraph. An *rb-path* is any path leading a root to some bud node and does not have other bud nodes. **Unless otherwise stated, we will consider only rb-paths in the digraphs associated to *normalised* pre-proof tree-sets.**

Remark 3. *According to Lemma 2, the bud node B of any such rb-path is the only node in the rb-path for which $S(B)$ is the premise of a $(Subst)$ rule.*

Definition 4 (cumulative substitution)**.** *An rb-path $[N^1, \ldots, N^n, B]$ $(n > 0)$ can be annotated by the* cumulative substitution $\sigma_{id}^{all}\sigma_1 \cdots \sigma_{n-1}$*, where σ_i is the substitution annotating the forward arrow leading N_i to N_{i+1}, for each $i \in [1..n-1]$, and σ_{id}^{all} is the* overall *identity substitution* $\cup_{N \in [N^1, \ldots, N^{n-1}]}\{x \mapsto x \mid x \in FV(S(N))\}$.

A list of sequents $[S_1, \ldots, S_n]$ $(n > 0)$ is *admissible* if either i) it is a singleton $(n = 1)$, or ii) for every $i \in [2..n]$, S_i is the premise of some rule whose conclusion is S_{i-1}. By construction, the list of sequents labelling the nodes from every path from the digraph associated to a pre-proof tree-set is admissible.

Lemma 3. *Let* $[N^1,\ldots,N^{n-1},N^n,B]$ *be an rb-path. We define its* cumulative *list* l_c *as* $[S(N^1)[\theta^c_{(1,n)}],\ldots,S(N^{n-1})[\theta^c_{(n-1,n)}],S(N^n),S(B)]$, *where* $\theta^c_{(i,n)}$ *is the cumulative substitution for* $[N^i,\ldots,N^{n-1},N^n]$. *Then, the following properties hold:*

1. l_c *is admissible, and*
2. *the rule applied on each* $S(N^i)$ *is also applicable on* $S(N^i)[\theta^c_{(i,n)}]$, $\forall i \in [1..n-1]$, *if it is different from* (Gen). *If the rule is* (Gen), *the* (Gen)*-step can be replaced by a* (Wk)*-step, where the LK's* (Wk) *rule is defined as*

$$\frac{\Gamma' \vdash \Delta'}{\Gamma \vdash \Delta}\ (Wk)\ \textit{if}\ \Gamma' \subseteq \Gamma, \Delta' \subseteq \Delta$$

Proof. We will perform induction on n. If $n = 1$, then $N^n \equiv N^1$ and $[S(N^1)]$ is a singleton, hence it is admissible.

If $n > 1$, let p denote the path $[N^1,\ldots,N^{n-1},N^n,B]$. By induction hypothesis, we assume that $[S(N^1)[\theta^c_{(1,n-1)}],\ldots,S(N^{n-2})[\theta^c_{(n-2,n-1)}],S(N^{n-1})]$ is admissible, where $\theta^c_{(i,n-1)}$ ($i \in [1..n-2]$) is the cumulative substitution annotating $[N^i,\ldots,N^{n-1}]$ and the rules applied on $S(N^i)[\theta^c_{(i,n-1)}]$ and $S(N^i)$ are the same. We denote by $\theta^c_{(n-1,n-1)}$ the identity substitution for $S(N^{n-1})$.

Let $\theta^c_{(i,n)}$ be the cumulative substitution annotating $[N^i,\ldots,N^{n-1},N^n]$, for all $i \in [1..n-1]$. Let also θ be the substitution annotating the forward arrow leading N^{n-1} to N^n, which can be either an identity substitution, or an equality substitution. In the first case, for every $i \in [1..n-1]$, $\theta^c_{(i,n)}$ is i) $\theta^c_{(i,n-1)} \cup \{x \mapsto x \mid x \in \bar{x}\}$ if the rule applied on $S(N^{n-1})$ is the LK's rule $(\forall R)$ or $(\exists L)$, defined below:

$$\frac{\Gamma \vdash F, \Delta}{\Gamma \vdash \forall \bar{x} F, \Delta}\ (\forall R)\ \text{if}\ \bar{x} \cap FV(\Gamma \cup \Delta) = \emptyset$$

$$\frac{\Gamma, F \vdash \Delta}{\Gamma, \exists \bar{x} F \vdash \Delta}\ (\exists L)\ \text{if}\ \bar{x} \cap FV(\Gamma \cup \Delta) = \emptyset$$

and $\bar{x}$ is the vector of new free variables introduced by these rules, or ii) $\theta^c_{(i,n-1)}$, otherwise. Since $S(N^i)[\theta^c_{(i,n-1)}] \equiv S(N^i)[\theta^c_{(i,n)}]$ by induction hypothesis, we can apply the same rules on $S(N^i)[\theta^c_{(i,n)}]$ and $S(N^i)$, hence the list $[S(N^1)[\theta^c_{(1,n)}],\ldots,S(N^{n-1})[\theta^c_{(n-1,n)}],S(N^n)]$ is admissible. $[S(N^1)[\theta^c_{(1,n)}],\ldots,S(N^{n-1})[\theta^c_{(n-1,n)}],S(N^n),S(B)]$ is also admissible since $S(B)$ is the premise of a $(Subst)$ rule whose conclusion is $S(N^n)$, by property 2) from Lemma 2.

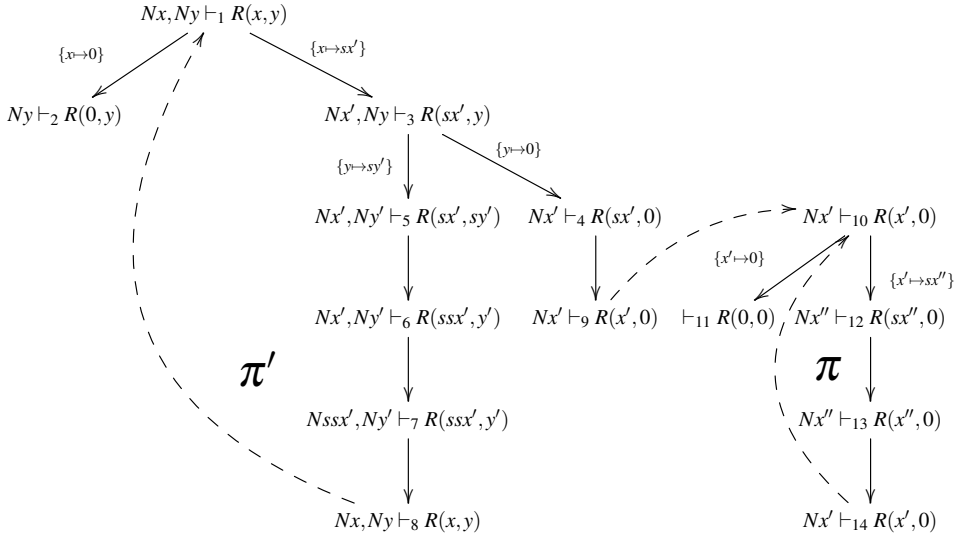
For the second case, θ is an equality substitution. We have that $\theta^c_{(i,n)}$ equals $\theta^c_{(i,n-1)}\theta$, for all $i \in [1..n-1]$. Since the rule applied on a sequent can also be applied on every instance of it, we have that $[S(N^1)[\theta^c_{(1,n)}],\ldots,S(N^{n-1})[\theta^c_{(n-1,n)}],S(N^n)]$ is admissible; the rule applied on $S(N^i)$ can also be applied on $S(N^i)[\theta^c_{(i,n)}]$, for all $i \in [1..n-1]$. Notice that the (Gen) rule has $S(N^n)$ as premise when applied on $S(N^{n-1})[\theta^c_{(n-1,n)}\theta]$. Let us assume that $x = u$ is the principal formula of $S(N^{n-1})[\theta^c_{(n-1,n)}]$. Then, θ is $\{x \mapsto u\}$. On the one hand, (Gen) cannot be applied on $S(N^{n-1})[\theta^c_{(n-1,n)}\theta]$, whose principal formula is $u = u$, when u is a non-variable term. On the other hand, the generalised form of (Gen) from CLKID$^\omega$, displayed in Figure 1, would replace u by u and delete $u = u$. If $S(N^{n-1})[\theta^c_{(n-1,n)}\theta]$ is of the form $\Gamma, u = u \vdash \Delta$, the same result can be achieved with CLKID$^\omega_N$ by applying (Wk) instead:

$$\frac{\Gamma \vdash \Delta}{\Gamma, u = u \vdash \Delta}\ (Wk)$$

So, the list $[S(N^1)[\theta^c_{(1,n)}],\ldots,S(N^{n-1})[\theta^c_{(n-1,n)}],S(N^n)]$ is admissible.
$[S(N^1)[\theta^c_{(1,n)}],\ldots,S(N^{n-1})[\theta^c_{(n-1,n)}],S(N^n),S(B)]$ is also admissible, as shown for the first case. □

A path has *cycles* if some nodes are repeated in the path. The set of *strongly connected components* (SCCs) of a digraph $\mathscr{P}$ of some pre-proof tree-set $(\mathscr{MD}, \mathscr{MR})$ is a partition of $\mathscr{P}$, where each SCC is a maximal sub-graph for which any two different nodes are linked in each direction by following only arrows from the sub-graph. Therefore, every non-singleton SCC has at least one cycle. Additionally, if $\mathscr{P}$ is acyclic, each of its nodes is a singleton SCC.

Example 3. *The digraph of the normalised pre-proof tree-set from Example 2 is:*



The sequent labelling a node is annotated by the number of the node in the digraph. The digraph has two non-singleton SCCs: i) π:$\{N^{10},N^{12},N^{13},N^{14}\}$, and ii) π' : $\{N^1,N^3,N^5,N^6,N^7,N^8\}$.

3.3 Defining the ordering and derivability conditions

The premises for defining the new soundness criterion are similar to [11]. Let π be a SCC from $\mathscr{P}$ and $<_a$ an ordering *stable under substitutions* defined over the set $\mathscr{S}$ of instances of the IAAs from the sequents labelling nodes inside π, i.e., if $l <_a l'$ then $l[\sigma] <_a l'[\sigma]$, for all $l,l' \in \mathscr{S}$ and substitution σ. Given a path p in π, we say that an IAA τ_j *derives from* an IAA τ_i using the trace $(\tau_k)_{(k\geq 0)}$ along p if $i < j$. Also, given two arbitrary substitutions γ and δ, we say that $\tau_j[\gamma]$ derives from $\tau_i[\delta]$ using $(\tau_k)_{(k\geq 0)}$ along p. $<_\pi$ is the *multiset extension* [2] of $<_a$.

The ordering constraints from a multiset extension relation comparing two sequent instances can be combined with derivability constraints on IAAs to give the $<_\pi$-*derivability* relation, referred to as *ordering-derivability* when the ordering is not known. For this, we assume that every sequent S has associated a measure value (weight), denoted by A_S and represented by a multiset of IAAs of S.

Definition 5 ($<_\pi$-derivability)**.** *Let N^i and N^j be two nodes occurring in some path p from π, and θ, δ be two substitutions. We define $A'_{S(N^i)[\theta]}$ (resp., $A'_{S(N^j)[\delta]}$) as the multiset, resulting from $A_{S(N^i)[\theta]}$ (resp., $A_{S(N^j)[\delta]}$) after the pairwise deletion of all common IAAs from $A_{S(N^i)[\theta]}$ and $A_{S(N^j)[\delta]}$. In addition, we assume that for each $l \in A_{S(N^j)[\delta]} \backslash A'_{S(N^j)[\delta]}$, there is $l' \in A_{S(N^i)[\theta]} \backslash A'_{S(N^i)[\theta]}$ satisfying i) $l \equiv l'$, and ii) l is the unique literal from $A_{S(N^j)[\delta]}$ that derives from l' using some trace following p.*

*Then, $S(N^j)[\delta]$ is $<_\pi$-*derivable *from $S(N^i)[\theta]$ along p if for each $l \in A'_{S(N^j)[\delta]}$ there exists $l' \in A'_{S(N^i)[\theta]}$ such that $l' >_a l$ and l derives from l' using some trace following p.*

By the definition of $<_\pi$ as a multiset extension of $<_a$, the following results can be proved when considering some path in π.

Lemma 4. *If S is $<_\pi$-derivable from S' then $A_S <_\pi A_{S'}$.*

Proof. By the definition of the ordering constraint in the $<_\pi$-derivability relation. □

Lemma 5. *The '$<_\pi$-derivability' relation is stable under substitutions and transitive.*

Proof. Let S and S' be two sequents such that S is $<_\pi$-derivable from S' along some path p in π. By Lemma 4, $A_{S'} >_\pi A_S$. Since $<_\pi$ is stable under substitutions, we have that $A_{S'[\sigma]} >_\pi A_{S[\sigma]}$, for every substitution σ. According to Definition 5, the derivability relations between their IAAs do not change by instantiation operations. Therefore, $S[\sigma]$ is $<_\pi$-derivable from $S'[\sigma]$ along p. We conclude that the '$<_\pi$-derivability' relation is stable under substitutions.

To prove the transitivity property, let us assume three sequents S_1, S_2 and S_3 labelling nodes in a path p built by the concatenation of two paths p_1 and p_2 such that S_3 is $<_\pi$-derivable from S_2 along p_2 and S_2 is $<_\pi$-derivable from S_1 along p_1. We will try to prove that S_3 is $<_\pi$-derivable from S_1 along p.

Since S_3 is $<_\pi$-derivable from S_2 along p_2, by Definition 5 we have that

(i1) for each $l_3 \in A'_{S_3}$ there exists $l_2 \in A'_{S_2}$ such that $l_2 >_a l_3$ and l_3 derives from l_2 using some trace following p_2, and

(ii1) for each $l_3 \in A_{S_3} \backslash A'_{S_3}$, there is some $l_2 \in A_{S_2} \backslash A'_{S_2}$ such that $l_3 \equiv l_2$ and l_3 is the unique IAA that derives from l_2 using some trace following p_2,

where A'_{S_3} (resp., A'_{S_2}) is the multiset resulting from A_{S_3} (resp., A_{S_2}) after the pairwise deletion of all common IAAs from A_{S_3} and A_{S_2}. Also, since S_2 is $<_\pi$-derivable from S_1 along p_1, we have that

(i2) for each $l_2 \in A''_{S_2}$, there exists $l_1 \in A'_{S_1}$ such that $l_1 >_a l_2$ and l_2 derives from l_1 using some trace following p_1, and

(ii2) for each $l_2 \in A_{S_2} \backslash A''_{S_2}$, there is some $l_1 \in A_{S_1} \backslash A'_{S_1}$ such that $l_2 \equiv l_1$ and l_2 is the unique IAA that derives from l_1 using some trace following p_1,

where A''_{S_2} (resp., A'_{S_1}) is the multiset resulting from A_{S_2} (resp., A_{S_1}) after the pairwise deletion of all common IAAs from A_{S_2} and A_{S_1}. We have to check that for each $l_3 \in A''_{S_3}$, there exists $l_1 \in A''_{S_1}$ such that $l_1 >_a l_3$ and l_3 derives from l_1 using some trace following p, where A''_{S_3} (resp., A''_{S_1}) is the multiset resulting from A_{S_3} (resp., A_{S_1}) after the pairwise deletion of all common IAAs from A_{S_3} and A_{S_1}. Moreover, for each $l_3 \in A_{S_3} \backslash A''_{S_3}$, there is some $l_1 \in A_{S_1} \backslash A''_{S_1}$ such that $l_3 \equiv l_1$ and l_3 is the unique IAA that derives from l_1 using some trace following p. We consider the following cases:

1. If $l_3 \in A'_{S_3}$ there exists $l_2 \in A'_{S_2}$ such that $l_2 >_a l_3$ and l_3 derives from l_2 using some trace t_2 following p_2.
   (a) If $l_2 \in A''_{S_2}$ there exists $l_1 \in A'_{S_1}$ such that $l_1 >_a l_2$ and l_2 derives from l_1 by using some trace t_1 following p_1. Then $l_1 >_a l_3$ by the transitivity of $<_a$, so $l_1 \in A''_{S_1}$, $l_3 \in A''_{S_3}$ and l_3 derives from l_1 using the concatenation of t_1 and t_2 following p.
   (b) If $l_2 \in A_{S_2} \backslash A''_{S_2}$, there is $l_1 \in A_{S_1} \backslash A'_{S_1}$ such that $l_2 \equiv l_1$ and l_2 is the unique IAA that derives from l_1 by using some trace t_1 following p_1. Since $l_1 (\equiv l_2) >_a l_3$, we have that $l_1 \in A''_{S_1}$, $l_3 \in A''_{S_3}$ and l_3 derives from l_1 using the concatenation of t_1 and t_2 following p.

2. If $l_3 \in A_{S_3} \backslash A'_{S_3}$ there exists $l_2 \in A'_{S_2}$ such that $l_3 \equiv l_2$ and l_3 is the unique IAA that derives from l_2 using some trace t_2 following p_2.
   (a) If $l_2 \in A''_{S_2}$ there exists $l_1 \in A'_{S_1}$ such that $l_1 >_a l_2$ and l_2 derives from l_1 by using some trace t_1 following p_1. Then, $l_1 >_a (l_2 \equiv) l_3$, so $l_1 \in A''_{S_1}$, $l_3 \in A''_{S_3}$ and l_3 derives from l_1 using the concatenation of t_1 and t_2 following p.
   (b) If $l_2 \in A_{S_2} \backslash A''_{S_2}$ there exists $l_1 \in A'_{S_1}$ such that $l_1 \equiv l_2$ and l_2 is the unique IAA that derives from l_1 by using some trace t_1 following p_1. This means that $l_3 \in A_{S_3} \backslash A''_{S_3}$, $l_1 \in A_{S_1} \backslash A''_{S_1}$ with $l_1 \equiv (l_2 \equiv) l_3$ and l_3 derives from l_1 using the concatenation of t_1 and t_2 following p. In addition, l_3 is the unique IAA in A_{S_3} that derives from l_1. □

The soundness criterion consists in checking if the sequents labelling (IH)-nodes from every non-singleton SCC, referred to as *induction hypotheses*, satisfy some constraints.

Definition 6 (induction hypothesis (IH), IH discharged by a SCC)**.** *Let π be a non-singleton SCC and $[R,\ldots,H,B]$ an rb-path p in π. We say that the* induction hypothesis *(IH) $S(H)$ is* discharged *by π if $S(H)$ is $<_\pi$-derivable from $S(R)[\theta^c]$ along p, where θ^c is the cumulative substitution annotating p.*

Theorem 1 (soundness)**.** *The sequents, labelling the roots from every normalised pre-proof tree-set whose non-singleton SCCs discharge their IHs, are true.*

Proof. Let M be a standard model for (Σ,Φ) and assume a normalised pre-proof tree-set. Let also $\mathscr{P}$ denote its digraph whose non-singleton SCCs discharge their IHs. By contradiction, we assume that there exists a root node N such that $S(N)$ is false. We define a partial (well-founded) ordering $<_{\mathscr{R}}$ over the (finite number of) root nodes from $\mathscr{P}$ such that, for every two distinct root nodes N^1 and N^2, we have $N^1 <_{\mathscr{R}} N^2$ if i) N^1 and N^2 are not in the same SCC, and ii) N^1 can be joined from N^2 in $\mathscr{P}$.

By induction on $<_{\mathscr{R}}$, we consider the base case when N is a $<_{\mathscr{R}}$-minimal node. (The step case, when N is not a $<_{\mathscr{R}}$-minimal node, will not be detailed since it can be treated similarly by assuming that all $<_{\mathscr{R}}$-smaller root nodes are labelled by true sequents.) If N is included in a one-node SCC, N is also a leaf node. The only 0-premise rules are the LK's (Ax) rule as well as $(R.)$ when unfolding with unconditional axioms. In both cases, $S(N)$ is true which leads to a contradiction.

Let us now assume that N is a $<_{\mathscr{R}}$-minimal node from some non-singleton SCC π. We will analyse all possible scenarios and show that each of them leads to a contradiction. The tree t from $\mathscr{P}$ and rooted by N should have buds labelled by false sequents, otherwise $S(N)$ would be true. Let B be such a bud such that N^h is its companion and $[N,\ldots,H,B]$ is an rb-path in π. N^h should be a root node from π because N is $<_{\mathscr{R}}$-minimal; it is labelled by the false sequent $S(B)$. Since the CLKID^ω_N rules are sound, by Lemma 3, we conclude that the *cumulative instance* $S(N)[\theta_c]$ is false, where θ_c is the cumulative substitution for $[N,\ldots,H,B]$. π discharges its IHs, so we have that $S(B)[\delta_h](\equiv S(H))$ is $<_\pi$-derivable from $S(N)[\theta_c]$, where δ_h is the substitution used by the $(Subst)$-step whose conclusion is $S(H)$. By Lemma 4, we have that $A_{S(N^h)[\delta_h]} <_\pi A_{S(N)[\theta_c]}$.

We perform a similar reasoning on N^h as for N. There is an rb-path $[N^h,\ldots,H',N^{f'}]$ such that the companion of $N^{f'}$ (in π) is $N^{h'}$ and $S(N^h)[\delta_h]$ shares false instances with $S(N^h)[\theta^c_1]$, where θ^c_1 is the cumulative substitution annotating $[N^h,\ldots,H',N^{f'}]$. By contradiction, we assume that no false instance of $S(N^h)[\delta_h]$ is shared. Then, one can build a finite bud-free pre-proof tree of $S(N^h)[\delta_h]$, by using only sound rules. Hence, $S(N^h)[\delta_h]$ is true, so contradiction. Therefore, there are two substitutions ε and τ such that $S(N^h)[\delta_h\varepsilon] \equiv S(N^h)[\theta^c_1\tau]$ and $S(N^h)[\theta^c_1\tau]$ is false. Let $S(N^{h'})[\delta'_h](\equiv S(H'))$ be the instance of $S(N^{h'})$ used as IH. Since it is discharged by π, we have that $A_{S(N^h)[\theta^c_1]} >_\pi A_{S(N^{h'})[\delta'_h]}$. From $A_{S(N)[\theta^c]} >_\pi A_{S(N^h)[\delta_h]}$

and the previous ordering constraint, we get $A_{S(N)[\theta^c\varepsilon]} >_\pi A_{S(N^h)[\delta_h\varepsilon]}$ and $A_{S(N^h)[\theta_1^c\tau]} >_\pi A_{S(N^{h'})[\delta_h'\tau]}$, by the 'stability under substitutions' property of $<_\pi$. Hence,

$$A_{S(N)[\theta^c\varepsilon]} >_\pi A_{(S(N^h)[\delta_h\varepsilon]} \equiv) A_{S(N^h)[\theta_1^c\tau]} >_\pi A_{S(N^{h'})[\delta_h'\tau]}$$

For similar reasons as given for $S(N^h)[\delta_h]$, we can show that $S(N^{h'})[\delta_h'\tau]$ is false, hence it can be treated similarly as $S(N^h)[\delta_h]$. And so on, the process can be repeated to produce an infinite strictly $<_\pi$-decreasing sequence s of measure values associated to instances of sequents labelling root nodes from π, of the form

$$A_{S(N)[\theta^c\varepsilon\cdots]} >_\pi A_{S(N^h)[\theta_1^c\tau\cdots]} >_\pi A_{S(N^{h'})[\cdots]} >_\pi \ldots$$

We can associate to s the infinite admissible list l_s of its sequents $[S(N)[\theta^c\varepsilon\cdots], S(N^h)[\theta_1^c\tau\cdots], S(N^{h'})[\cdots], \ldots]$ and define the path p *underlying* l_s as the concatenation of the rb-paths from π that built s, i.e., $[N,\ldots,B,N^h,\ldots,N^{f'},\ldots]$. By the construction of s, every successive $(Subst)$-, bud and root nodes in p are labelled by the same sequent instance in l_s, so the $(Subst)$-steps are stuttering in l_s. By Lemma 4, all (Gen)- can be replaced by (Wk)-steps. p is of the form $[N_\infty\ldots,N_1,\ldots,N_0]$ where N_0, N_1, ..., N_∞ are an infinite number of *all* the occurrences of N in p.

We will show that there is a trace following p that has an infinite number of progress points. As explained in [7], it means that there is an infinite strictly decreasing sequence of ordinals, hence contradiction. Since p is the concatenation of rb-paths in π and π discharges its IHs, for each such rb-path the bud sequent is $<_\pi$-derivable from the cumulative instance, along the rp-path, of the root sequent. By Lemma 5, there is an instance $S(N_\infty)[\theta_\infty]$ such that $S(N_0)$ is $<_\pi$-derivable from it along p, where θ_∞ is the composition of all cumulative substitutions of the rb-paths from l. For any two consecutive nodes N_i and N_{i-1} ($i \in [1..\infty]$), we have that $S(N_{i-1})[\theta_{i-1}]$ is $<_\pi$-derivable from $S(N_i)[\theta_i]$, where θ_i (resp., θ_{i-1}) are the compositions of all cumulative substitutions of the rb-paths along $[N_i,\ldots,N_0]$ (resp., $[N_{i-1},\ldots,N_0]$).

Let us denote by S (resp, S') the sequent $S(N_i)[\theta_i]$ (resp., $S(N_{i-1})[\theta_{i-1}]$), for some $i \in [1..\infty]$. By Definition 5 and the transitivity of the $<_\pi$-derivability relation, for each IAA l from A_S there is an IAA l' from $A_{S'}$ such that l derives from l'. Therefore, there are n traces along the path p' $[N_\infty,\ldots,N_i]$, where n is the number of IAAs from S.

We will show that the traces along p' have an infinite number of progress points. By contradiction, we assume that this number is finite. Therefore, there is a subpath p'' of p whose traces have no progress points and there exists $j \in [1..\infty]$ such that N_j and N_{j-1} belong to p''. Let us denote by S_j (resp, S_{j-1}) the sequent $S(N_j)[\theta_j]$ (resp., $S(N_{j-1})[\theta_{j-1}]$). Since S_{j-1} is $<_\pi$-derivable from S_j, we have that $A_{S_{j-1}} <_\pi A_{S_j}$. By the definition of $<_\pi$ as a multiset extension of the ordering $<_a$ over the instances of IAAs from the root sequents in π, there should be an IAA $l \in A_{S_{j-1}}$ for which there is another IAA $l \in A_{S_j}$ such that $l <_a l'$ and l derives from l', i.e., l and l' are from an infinite trace t following (a subpath of) p'' which has no progress points. According to the definition of a trace (see Definition 2) and the way l_s was built, $l <_a l'$ is possible only if the subtrace of t from l' to l has at least one progress point, so contradiction. Otherwise, $l \equiv l'$ since i) l_s is admissible, ii) the $(Subst)$-steps are stuttering, iii) the (Gen)-steps can be replaced by (Wk)-steps, and iv) the instantiation steps that built s preserve the equality relations. □

Example 4. *For the sequents labelling the nodes from the digraph given in Example 2, we define the measure values $A_{Nt\vdash R(t,0)} = \{Nt\}$, $\forall t$, and $A_{Nt_1,Nt_2\vdash R(t_1,t_2)} = \{Nt_2\}$, $\forall t_1,t_2$. The IH $S(N^{13})$ is $<_\pi$-derivable from $S(N^{10})[\{x' \mapsto sx''\}]$, hence discharged by the SCC π using the trace $[Nx',Nx'',Nx'']$, if $\{Nx''\} <_\pi \{Nsx''\}$. Also, the IH $S(N^7)$ is $<_{\pi'}$-derivable from $S(N^1)[\{x \mapsto sx'; y \mapsto sy'\}]$ in the SCC π' using the trace $[Ny,Ny,Ny',Ny',Ny',Ny']$ if $\{Ny'\} <_{\pi'} \{Nsy'\}$. The ordering constraints hold if $<_\pi$ and $<_{\pi'}$ are the multiset extensions of a recursive path ordering [2] $<_{rpo}$ for which $z <_{rpo} sz$, for every variable z.*

By Theorem 1, the root sequents in the pre-proof tree-set, $S(N^1)$ and $S(N^{10})$, are true.

Comparison with the soundness checking criterion from [11]. In [11], the ordering-derivability constraints issued when analysing if a pre-proof tree-set is a proof are defined at the level of the minimal cycles of its digraph, referred to as n-cycles. A *n-cycle* is defined as a finite circular list $[N_1^1,\ldots,N_1^{p_1}],\ldots,[N_n^1,\ldots,N_n^{p_n}]$ of n (>0) paths leading root nodes to buds such that $N^1_{next(i)} = \mathscr{MR}(N_i^{p_i})$, for any $i \in [1..n]$, where $next(i) = 1 + (i \bmod n)$.

Let π be a non-singleton SCC and C an *n-cycle* $[N_1^1,\ldots,N_1^{p_1}],\ldots,[N_n^1,\ldots,N_n^{p_n}]$ from π. The induction hypotheses are defined at the n-cycle level. For all $i \in [1..n]$, let θ_i^c be the cumulative substitution annotating $[N_i^1,\ldots,N_i^f]$, where the *IH-node* N_i^f is either i) $N_i^{p_i}$ if $(Subst)$ is not applied along $[N_i^1,\ldots,N_i^{p_i}]$, or ii) $N_i^{p_i-1}$, otherwise. The sequents labelling the IH-nodes correspond exactly to the induction hypotheses used in the paper. We say that the IHs $S(N_j^f)$ ($j \in [1..n]$) are *discharged* by C if, $\forall i \in [1..n]$, $S(N_i^f)$ is $<_\pi$-derivable from $S(N_i^1)[\theta_i^c]$ along $[N_i^1,\ldots,N_i^{p_i}]$. In [11], a *proof* is every pre-proof tree-set whose digraph has only n-cycles that discharge their IHs and it has been shown that its root sequents are true.

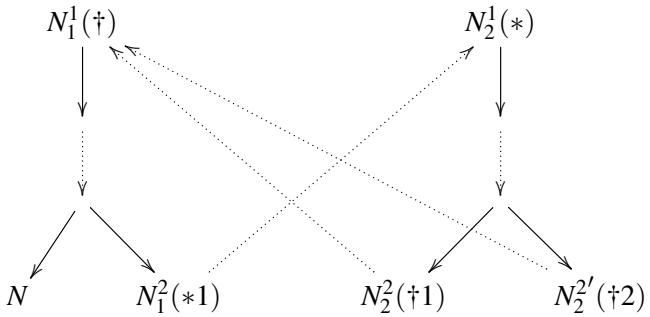


Figure 4: Two 2-cycles sharing the same path.

Since several n-cycles may share the *same* root-bud path, some ordering-derivability constraints may be duplicated when checking that a pre-proof is a proof. For example, the path $[N_1^1,\ldots,N_1^2]$ is shared between the two 2-cycles $[N_1^1,\ldots,N_1^2][N_2^1,\ldots,N_2^2]$ and $[N_1^1,\ldots,N_1^2][N_2^1,\ldots,N_2^{2'}]$ from the digraph given in Figure 4. Even if the number of n-cycles from a digraph can be large, as explained in the introduction, the number of *distinct* ordering-derivability constraints is always smaller or equal than the number of buds from the non-singleton SCCs. With the approach from [11], the duplicates of the constraints do not need to be again processed if the already processed constraints are recorded. It has been shown that the time complexity of the soundness checking procedure is polynomial if the number of the ordering-derivability constraints is that of the buds from the non-singleton SCCs. With our new approach, the number of operations for normalising a CLKID^ω_N pre-proof of n nodes is given by the sum of non-root companions, non-terminal $(Subst)$-nodes and nodes labelled by some sequent that is the premise of a rule r different from $(Subst)$. So, it is smaller than $3n$. Let c be the maximal cost of an operation, including the node duplication and the creation of a $(Subst)$-node or bud-companion relation. Their total cost is smaller than $4nc$ (the second operation duplicates it twice). The costs for annotating substitutions and for evaluating an ordering-derivability constraint are given in [11].

4 Conclusions and future work

We have defined a more efficient soundness criterion for a class of CLKID$^\omega$ pre-proofs considered in [11], by building a set of non-redundant ordering-derivability constraints. We have shown that these constraints can also be extracted from those that define the soundness criterion from [11], by deleting the duplicated values. The new normal forms and their digraphs allow to uniformly represent rb-paths and can be built in linear time. We conclude that the two soundness checking criteria have the same polynomial-time complexity if the time complexity for comparing two IAAs is at most polynomial.

In the future, we plan to adapt our approach to make more effective other soundness criteria based on minimal cycles, e.g., those involving cyclic formula-based Noetherian induction reasoning [10, 12], and other systems where the soundness can be checked by the global trace condition, as CLJID$^\omega$ [3].

Acknowledgements

The author thanks the anonymous reviewers for their comments that helped to improve the quality of the paper.

References

[1] P. Aczel. An introduction to inductive definitions. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 739–782. North Holland, 1977. https://doi.org/10.1016/S0049-237X(08)71120-0

[2] F. Baader and T. Nipkow. *Term Rewriting and All That*. Cambridge University Press, 1998.

[3] S. Berardi and M. Tatsuta. Intuitionistic Podelski-Rybalchenko theorem and equivalence between inductive definitions and cyclic proofs. In *CMCS'18*, LNCS, 2018. to appear.

[4] J. Brotherston. Cyclic proofs for first-order logic with inductive definitions. In *Proceedings of TABLEAUX-14*, volume 3702 of *LNAI*, pages 78–92. Springer-Verlag, 2005. https://doi.org/10.1007/11554554_8

[5] J. Brotherston. *Sequent Calculus Proof Systems for Inductive Definitions*. PhD thesis, University of Edinburgh, November 2006.

[6] J. Brotherston, N. Gorogiannis, and R. L. Petersen. A generic cyclic theorem prover. In *APLAS-10 (10th Asian Symposium on Programming Languages and Systems)*, volume 7705 of *LNCS*, pages 350–367. Springer, 2012. https://doi.org/10.1007/978-3-642-35182-2_25

[7] J. Brotherston and A. Simpson. Sequent calculi for induction and infinite descent. *Journal of Logic and Computation*, 21(6):1177–1216, 2011. https://doi.org/10.1093/logcom/exq052

[8] G. Gentzen. Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift*, 39:176–210, 1935.

[9] O. Kupferman and M. Vardi. Weak alternating automata are not that weak. *ACM Transactions on Computational Logic (TOCL)*, 2(3):408–429, 2001. https://doi.org/10.1145/377978.377993

[10] S. Stratulat. A unified view of induction reasoning for first-order logic. In A. Voronkov, editor, *Turing-100 (The Alan Turing Centenary Conference)*, volume 10 of *EPiC Series*, pages 326–352. EasyChair, 2012. https://doi.org/10.29007/nsx4

[11] S. Stratulat. Cyclic proofs with ordering constraints. In R. A. Schmidt and C. Nalon, editors, *TABLEAUX 2017*, volume 10501 of *LNAI*, pages 311–327. Springer, 2017. https://doi.org/10.1007/978-3-319-66902-1_19

[12] S. Stratulat. Mechanically certifying formula-based Noetherian induction reasoning. *Journal of Symbolic Computation*, 80, Part 1:209–249, 2017. https://doi.org/10.1016/j.jsc.2016.07.014