A Constructive Proof of the Topological Kruskal Theorem

Jean Goubault-Larrecq

LSV, ENS Cachan, CNRS, INRIA.

Abstract. We give a constructive proof of Kruskal's Tree Theorem—precisely, of a topological extension of it. The proof is in the style of a constructive proof of Higman's Lemma due to Murthy and Russell (1990), and illuminates the role of regular expressions there. In the process, we discover an extension of Dershowitz' recursive path ordering to a form of cyclic terms which we call μ-terms. This all came from recent research on Noetherian spaces, and serves as a teaser for their theory.

1 Introduction

Kruskal's Theorem [33] states that the homeomorphic embedding ordering on finite trees is a a well quasi-ordering. This is a deep and fundamental theorem in the theory of well quasi-orderings. The aim of this paper is to give a *constructive*, that is, an intuitionistic proof of this fact[1].

I will explain what all that means in Section 2. I should probably admit right away that I have not *actively* looked for such a proof. It came to me in 2010 as a serendipitous by-product of research I was doing on Noetherian spaces, seen as a generalization of well quasi-ordered spaces. The result is, hopefully, a nice piece of mathematics. It is also an opportunity for me to explain various related developments which I would dare to say have independent interest.

I would like to issue a word of warning, though. The constructive proofs of the topological Higman and Kruskal theorems I am giving here were the first I found. The non-constructive proofs of [29, Section 9.7] came second. These are the ones I chose to publish, for good reason: once cast in formal language, the original constructive proofs are terribly heavy. I have therefore opted for a somewhat lighter presentation here, which stresses the beautiful *core* of the proof, at the cost at being somewhat sketchy in Sections 4 (Higman) and 5 (Kruskal). And this core is: these theorems reduce to questions of *termination* problems, which one can solve by using multiset orderings (Higman), resp. an extension of Dershowitz' multiset path ordering (Kruskal).

[1] I will avoid any debate of what intuitionism or constructivism is, and assume the logic of any of the modern proof assistants based on intuitionistic type theory, such as Coq [6]. The full calculus of inductive constructions with universes is definitely not needed, though. I only need first-order intuitionistic logic, plus a few inductively defined predicates and relations, and their associated induction principles.

Acknowledgments. I must thank David Baelde, who found a mistake in an early version of this paper, and Nachum Dershowitz, who gave me several additional pointers. I have had several interesting discussions with Sylvain Schmitz, Alain Finkel, and Jean-Pierre Jouannaud. All remaining errors are of course mine.

2 Well quasi-orderings, Noetherian spaces

A *quasi-ordering* on a set X is a reflexive and transitive binary relation $\leqslant$ on X. Given a subset A of X, we write $\uparrow A$ for its upward closure $\{y \in X \mid \exists x \in A \cdot x \leqslant y\}$, and call A *upward closed* if and only if $A = \uparrow A$. A *basis* of an upward closed subset E is any set A such that $E = \uparrow A$; E has a *finite basis* if and only if one can take A finite. We define the downward closure $\downarrow A$, and downward closed subsets, similarly. We also write $\geqslant$ for the converse of $\leqslant$, $<$ for the strict part of $\leqslant$ ($x < y$ iff $x \leqslant y$ and not $y \leqslant x$), $>$ for that of $\geqslant$.

There are many equivalent definitions of a *well* quasi-ordering (wqo for short), of which here are a few:

1. every infinite sequence $(x_n)_{n\in\mathbb{N}}$ in X is *good*, namely, there are two indices $m < n$ with $x_m \leqslant x_n$;
2. every infinite sequence $(x_n)_{n\in\mathbb{N}}$ in X is *perfect*, i.e., has an infinite ascending subsequence $x_{n_0} \leqslant x_{n_1} \leqslant \ldots \leqslant x_{n_i} \leqslant \ldots$ (with $n_0 < n_1 < \ldots < n_i < \ldots$);
3. $\leqslant$ is *well-founded* (there is no infinite descending sequence of elements $x_0 > x_1 > \ldots > x_n > \ldots$) and has *no infinite antichain* (an infinite sequence of pairwise incomparable elements);
4. every upward closed subset U has a finite basis;
5. every ascending chain $U_0 \subseteq U_1 \subseteq \ldots \subseteq U_n \subseteq \ldots$ of upward closed subsets is stationary (i.e., all U_ns are equal from some rank n onwards);
6. every descending chain $F_0 \supseteq F_1 \supseteq \ldots \supseteq F_n \supseteq \ldots$ of downward closed subsets is stationary;
7. the *strict inclusion* ordering $\subset$ is well-founded on downward closed subsets, i.e., there is no infinite descending chain $F_0 \supset F_1 \supset \ldots \supset F_n \supset \ldots$ of downward closed subsets.

The latter shows that being a wqo is merely a *termination* property, only one not on words, or on terms, as would be familiar in computer science [13], but rather on downward closed subsets.

There are many useful wqos in nature: $\mathbb{N}$ with its natural ordering $\leqslant$, any finite set, any finite product of wqos (in particular $\mathbb{N}^k$ with its componentwise ordering: this is Dickson's Lemma [18]), any finite coproduct of wqos, the set of finite words X^* over a well-quasi-ordered alphabet X (with the so-called word embedding quasi-ordering: this is Higman's Lemma [30]), the set of finite trees, a.k.a., first-order terms, $\mathcal{T}(X)$ over a well-quasi-ordered signature X (with the so-called tree embedding quasi-ordering: this is Kruskal's Theorem [33]), notably.

There are also more and more applications of wqo theory in computer science.

Termination. An early application is Nachum Dershowitz' discovery of the *multiset path ordering* on terms. This is a strict ordering $<^{\mathrm{mpo}}$ on terms that is well-founded, i.e., such that there is no infinite $>^{\mathrm{mpo}}$-chain $t_0 >^{\mathrm{mpo}} t_1 >^{\mathrm{mpo}} \ldots >^{\mathrm{mpo}} t_n >^{\mathrm{mpo}} \ldots$: to show that a rewrite system $\mathcal{R}$ terminates, it is enough to show that $\ell >^{\mathrm{mpo}} r$ for every rule $\ell \to r$ in $\mathcal{R}$. Dershowitz' initial proof ([12], see also [11]) rested on the remark that $>^{\mathrm{mpo}}$ is a simplification ordering: if t embeds into s, then $t \leqslant^{\mathrm{mpo}} s$. Given any infinite $>^{\mathrm{mpo}}$-descending chain as above, by Kruskal's Theorem one can find $i < j$ such that t_i embeds into t_j. It follows that $t_i \leqslant^{\mathrm{mpo}} t_j$, contradicting $t_i >^{\mathrm{mpo}} t_j$. This uses characterization 1 of wqos.

This simple argument definitely relies on Kruskal's deep result. The realization that Dershowitz' theorem required much less logical clout [26, 8] came to me as both a relief and a disappointment : I'll recapitulate the elementary argument in Section 3. I'll also give a slight extension of this elementary argument to a form of cyclic terms I have decided to call μ-terms. This will be instrumental in the rest of the paper, and may even be useful in the rewriting community.

Minimal patterns. A second application arises from characterization 4. Given an upward closed language L of elements in a wqo X, one can test whether $x \in L$ by just checking finitely many equalities $x_1 \leqslant x$, $\ldots$, $x_n \leqslant x$. Indeed, property 4 states that one can write L as $\uparrow\{x_1, \ldots, x_n\}$. For example, this is how van der Meyden shows that fixed monadic queries to indefinite databases can be evaluated in linear time in the size of the database [44], where x, x_1, $\ldots$, x_n are (encodings of models as) finite sequences of finite sets of logical atoms. The query L defines the minimal patterns $x_1, \ldots, x_n$ to be checked, in the embedding quasi-ordering on words. That the latter is a wqo is Higman's Lemma, and the fact that its standard proofs are non-constructive implies the curious fact that one cannot *a priori* compute $x_1, \ldots, x_n$ from L. That is, a linear time algorithm exists for each $L\ldots$ but what is it? Ogawa [40] solves the issue by extracting the computational content of Murthy and Russell's constructive proof of Higman's Lemma [37]. This computes the values x_1, $\ldots$, x_n, hence derives a linear-time algorithm for the query L, from L given as input.

WSTS. Another application is in verification of *well-structured transition systems* (WSTS) [1, 25]. A WSTS is a (possible infinite-state) transition system $(X, \to)$, with a *wqo* $\leqslant$ of the set of states X, satisfying a monotonicity property. For simplicity, we shall only consider strong monotonicity: if $s \to s'$ and $s \leqslant t$, then there is a state t' such that $t \to t'$ and $s' \leqslant t'$.

Examples of WSTS abound. Petri nets are WSTS whose state space is $\mathbb{N}^k$, where k is the number of places. Affine nets [24] generalize these and many other variants, and are still WSTS on $\mathbb{N}^k$. Lossy channel systems [3] are networks of finite-state automata that communicate over FIFO queues. They are WSTS whose state space is $\prod_{i=1}^{m} Q_i \times \prod_{j=1}^{n} \Sigma_j^*$, where Q_i is the finite state space of the ith automaton, and Σ_j is the finite alphabet of the jth queue. Let us also cite data nets [34], BVASS [46, 10], and recent developments in the analysis of processes [36, 4, 47, 42], which require tree representations of state.

The simple structure of a WSTS implies that *coverability* is decidable in every effective WSTS. This is the following question: given a state $s \in X$ and an upward closed subset U of X, is there a state $t \in U$ that is reachable from s, i.e., such that $s \to^* t$, where $\to^*$ is the reflexive-transitive closure of $\to$? By *effective* WSTS, we mean that we can represent states on a computer (which implies that every upward closed subset U is representable as well, as a finite set E, by property 4), that $\leqslant$ is decidable, and that the set of one-step predecessors $\mathrm{Pre}(U) = \{s \in X \mid \exists t \in U \cdot s \to t\}$ of a state t is computable. This is the case of all WSTS mentioned above. Inclusion of upward closed subsets is decidable, since $\uparrow E_1 \subseteq \uparrow E_2$ if and only if for every $x \in E_1$, there is a $y \in E_2$ with $y \leqslant x$. That coverability is decidable is almost trivial: using a while loop, compute the successive sets $U_0 = U$, $U_{n+1} = U_n \cup \mathrm{Pre}(U_n)$, and stop when $U_{n+1} \subseteq U_n$; this must eventually happen by property 5. Then there is a state in U that is reachable from s if and only if $s \in U_n$.

In 1969, Karp and Miller [32] devised another way (historically, the first one) of deciding coverability. They built a so-called coverability tree, and showed that it was finite and effectively constructible by resorting to Dickson's Lemma, plus a few additional tricks. One of the tricks they required was to extend the state space from $\mathbb{N}^k$ to $\mathbb{N}_\omega^k$, where $\mathbb{N}_\omega$ is $\mathbb{N}$ plus a fresh top element ω, the *limit* of any ever growing sequence. Although it would seem natural that the construction would generalize to every WSTS, progress was slow. One of the blocking factors was to define a completion $\widehat{X}$ of a well quasi-ordered state space X, so that Karp and Miller's construction would adapt.

By analogy with $\mathbb{N}^k$, $\widehat{X}$ should be X with some limit points added, and this naturally calls for topology. Alain Finkel once asked me whether there would be a notion of completion from topology that could serve this purpose. We realized that the *sobrification* of X (see [29, Section 8.2]) was the right candidate, and this led us to a satisfactory extension of Karp and Miller's procedure to all WSTS [20, 21, 23].

Noetherian spaces. In the process, going to topology begged the question whether there is a topological characterization of wqos. I realized in [27] that this would be the notion of Noetherian space, invented in algebraic geometry in the first half of the 20th century. A *Noetherian* space is a space where every ascending chain of opens is stationary: comparing this with property 5, we have merely replaced "upward closed" by "open".

Every quasi-ordered set can be equipped with the so-called *Alexandroff topology*, whose opens are just the upward closed subsets. Property 5 immediately implies that every wqo is Noetherian, once equipped with its Alexandroff topology. The framework of Noetherian spaces also allows us to extend the WSTS methodology to more kinds of transition systems. I have explained this in [28], applying this to two examples: a certain kind of multi-stack automata, and concurrent polynomial programs manipulating numerical values (in $\mathbb{R}$) that communicate through discrete signals over lossy channels. The decidability results that I'm stating in these settings are far from trivial, but are low-hanging fruit once we have the theory of Noetherian spaces available.

By "theory of Noetherian spaces", I do not mean the one we inherit from algebraic geometry, rather some natural results that arise from cross-fertilization with wqo theory. (See [29, Section 9.7] for a complete treatment.) Of interest to us are the following generalizations of Higman's Lemma and Kruskal's Theorem, respectively:

Topological Higman Lemma [29, Theorem 9.7.33]: if X is Noetherian, then the space of finite words X^* with the word topology is Noetherian, too.

Topological Kruskal Theorem [29, Theorem 9.7.46]: if X is Noetherian, then set space of finite trees $\mathcal{T}(X)$ with symbol functions taken from X is Noetherian under the tree topology.

We define the word and tree topologies as follows. Intuitively, think of an open set U as a test—namely, x passes the test if and only if $x \in U$. In the word topology, we wish the following to be a test: given tests $U_1, \ldots, U_n$ on letters (open subsets of X), the word w passes the test $X^*U_1X^* \ldots X^*U_nX^*$ if and only if w contains a (not necessarily contiguous) subword $a_1a_2 \ldots a_n$ with each a_i in U_i. In the tree topology, the basic tests are whether a given tree has an embedded subtree of a given shape, and where each function symbol is in a given open subset of X (possibly different at each node). In each case, these tests form bases for the required topologies, i.e., the opens are all unions of such tests.

The proofs I give of these theorems in [29, Section 9.7] are elegant, yet terribly topological, and rest on many results that require classical logic, and the Axiom of Choice. Instead, we shall use the following remark.

Call a closed subset F *irreducible* if and only if, for every finite family of closed subsets $F_1, \ldots, F_n$, if $F \subseteq F_1 \cup \ldots \cup F_n$, then $F \subseteq F_i$ for some i already. By [29, Theorem 9.7.12], a space X is Noetherian if and only if: ($\downarrow$) the strict inclusion relation $\subset$ is well-founded on the set $\mathcal{S}(X)$ of irreducible closed subsets of X ($\mathcal{S}(X)$ happens to be the sobrification of X we alluded to above), (T) the whole space X can be written as the union of finitely many irreducible closed subsets of X, and (W) given any two irreducible closed subsets F_1, F_2 of X, $F_1 \cap F_2$ can be written as the union of finitely many irreducible closed subsets of X. It follows that every closed subset will be a finite union of irreducible closed subsets, and that the strict inclusion ordering $\subset$ will be well-founded on closed subsets. The latter generalizes property 7, since in a quasi-ordered set, the (Alexandroff) closed sets are exactly the downward closed sets.

This leads us to the following proof plan:

(A) Find concrete representations of all irreducible closed subsets. This programme was initiated in [20] and carried out in [22], where we call the latter *S-representations*. In both the word and tree cases, our S-representations are certain forms of regular expressions, over words, or over trees. On words, this generalizes the products and the semi-linear regular expressions (SRE) of [2]; on trees, no prior work seems to have existed. These are effective representations: we can decide inclusion (in polynomial time, modulo an oracle deciding inclusion of irreducible closed subsets of letters, resp., of function symbols), and we can compute finite intersections of S-representations (in polynomial time again, provided the number of input representations is bounded).

(B) Show directly that strict inclusion is well-founded on S-representations. This will establish property $(\downarrow)$. Properties (T) and (W) are mostly obvious, since we even have *algorithms* to compute finite intersections.

In the case of the topological Higman Lemma (on words), we shall obtain a re-reading of Murthy and Russell's celebrated constructive proof of Higman's Lemma ([37]; see also [40], footnotes 6 and 7, for fixes to the definition of sequential regular expression). Our S-representations will be their *sequential regular expressions*, seen as the result of building SREs (originating in [2]) over a cotopology. Our constructive proof of Kruskal's Theorem, and indeed of its topological generalization, is in the same spirit, and we believe it provides a satisfactory answer to Murthy and Russell's final question [37].

Intuitionism. One difficulty with finding intuitionistic proofs in the theory of wqos is that properties 1–7 are *not* constructively equivalent. Notably, 2 is intuitionistically strictly stronger than 1, as Veldman notes [45, 1.3]. Indeed, 2 fails on $X = \mathbb{N}$ in an intuitionistic setting, while 1 is constructively valid. Similarly, 4 fails on $\mathbb{N}^2$, intuitionistically, even for decidable subsets of $\mathbb{N}^2$ [45, 1.2].

Following Murthy and Russell, a *constructive wqo* is defined by the following reformulation of property 1: (1') the opposite of the prefix ordering on bad finite sequences of words in X^* is well-founded. A finite sequence $x_0, x_1, \ldots, x_n$ is *bad* iff it is not good, that is, if $x_i \not\leqslant x_j$ for no $i < j$. The well-foundedness requirement means that one cannot extend finite sequences (adding x_{n+1}, x_{n+2}, etc.) indefinitely, keeping them all bad.

Murthy and Russell actually proved property 7. They derived (1') from 7, assuming $\leqslant$ decidable in the constructive sense that $\forall x, y \in X \cdot x \leqslant y \vee \neg(x \leqslant y)$ is provable. All the other constructive proofs I know of Higman's Lemma prove (1'), some of them directly [41, 7, 5]; the latter two do not require $\leqslant$ to be decidable. There are fewer intuitionistic proof of Kruskal's Theorem. One is due to Monika Seisenberger [43], who gives a direct proof of (1') on trees, based on a intuitionistic variant of Nash-Williams' minimal bad sequence argument [38]. She requires the quasi-ordering $\leqslant$ on function symbols to be decidable. Wim Veldman's proof [45] does not make this requirement, but models tree embedding with so-called at-most-ternary relations rather than using a binary relation $\leqslant$. He shows that Kruskal's original proof [33] can be made constructive, replaying the needed part of Ramsey theory in intuitionistic logic. Curiously, our proofs of the *topological* versions of Higman's Lemma and Kruskal's Theorem are entirely constructive, and we only need to assume $\leqslant$ decidable to deduce the *ordinary*, order-theoretic versions of these results from the topological versions.

3 Path Orderings

Path orderings (mpo, lpo, rpo) have been an essential ingredient of termination proofs for rewrite systems since their inception by Nachum Dershowitz in 1982 [12]. We shall concentrate on Dershowitz' original *multiset path ordering* (a.k.a., mpo). He proved that the mpo was well-founded as a consequence of Kruskal's

Theorem. We give an elementary, inductive, intuitionistic proof instead. This is based on a paper I wrote in 2001 [26]. Coupet-Grimal and Delobel [8] implemented a similar proof in Coq, with a proof of the Dershowitz-Manna Theorem (which I had not given, but Nipkow had [39]—see below). Dershowitz and Hoot's earlier proof that the general path ordering is well-founded [15] is non-constructive but elementary as well. Even earlier, Lescanne had already given an inductive proof that the mpo was well-founded [35, Theorem 5]; his proof relies on Zorn's Lemma (op.cit., Lemma 5), and ours will be simpler anyway, but his notion of decomposition ordering is illuminating.

Let X be a set with a binary relation $<$ on it. We again write $>$ for the converse of $<$. One thinks of $<$ as a strict ordering, but this is not needed. What will be important is that $<$ is *well-founded*: classically, this means that there is no infinite $>$-chain $x_1 > x_2 > \ldots > x_n > \ldots$ Constructively, it is better to say that $<$ is well-founded iff every element is $<$-*accessible*, where $<$-accessibility is the predicate defined inductively by (i.e., the least predicate such that):

$$\frac{\text{every } y < x \text{ is } <\text{-accessible}}{x \text{ is } <\text{-accessible}}$$

The set of $<$-accessible elements is traditionally called the *well-founded part* of $<$, i.e., the set of elements that cannot start an infinite $>$-chain. Since $<$-accessibility is defined inductively, we obtain the following useful principle of $<$-*induction*: to prove that a property P holds of every $<$-accessible element x, it is enough to show it under the additional assumption that P holds of every $y < x$ (the *induction hypothesis*). Another useful principle is $<$-*inversion*: if x is $<$-accessible, and $x > y$, then y is $<$-accessible as well.

Write $\{\!|x_1, \ldots, x_n|\!\}$ for the (finite) multiset consisting of the elements $x_1, \ldots, x_n \in X$. Let $\varnothing$ be the empty multiset, and $\uplus$ denote multiset union. We use the letters M, M', $\ldots$, for multisets. Intuitionistically, we assume an inductive definition of multisets, e.g., as finite lists, and we will reason up to permutation. (This actually incurs some practical difficulties in proof assistants such as Coq, which we shall merrily gloss over.) On the set $\mathcal{M}(X)$ of multisets of elements of X, we define the *multiset extension* $<_{\text{mul}}$ of $<$, inductively, by:

$$\frac{\text{for every } i \ (1 \leqslant i \leqslant n),\ x > x_i}{M \uplus \{\!|x|\!\} >_{\text{mul}} M \uplus \{\!|x_1, \ldots, x_n|\!\}}$$

That is, we replace some element x by arbitrarily many smaller elements $x_1, \ldots, x_n$. The following *Dershowitz-Manna Theorem* [17] is crucial.

Lemma 1 (Dershowitz-Manna, Nipkow). *For all $<$-accessible elements $x_1, \ldots, x_n \in X$, $\{\!|x_1, \ldots, x_n|\!\}$ is $<_{mul}$-accessible. In particular, if $<$ is well-founded on X, then $<_{mul}$ is well-founded on $\mathcal{M}(X)$.*

Proof. We give Nipkow's intuitionistic proof [39]. Let *Acc* denote the set of $<_{\text{mul}}$-accessible multisets. We prove that $\{\!|x_1, \ldots, x_n|\!\} \in \mathit{Acc}$ by induction on n. The case $n = 0$ is obvious, while the induction step consists in showing that, for

every $<$-accessible x: $(*)$ for every $M \in Acc$, $M \uplus \{\!|x|\!\} \in Acc$. Fix an $<$-accessible x, and use $<$-induction. This provides us with the induction hypothesis: (a) for every $y < x$, for every $M \in Acc$, $M \uplus \{\!|y|\!\} \in Acc$. To prove $(*)$, we show by $<_{\mathrm{mul}}$-induction on $M \in Acc$ that: $(**)$ $M \uplus \{\!|x|\!\} \in Acc$. This gives us the extra induction hypothesis (b): for every $M' <_{\mathrm{mul}} M$, $M' \uplus \{\!|x|\!\} \in Acc$. It now remains to show that (a) and (b) imply $(**)$. By definition of $<_{\mathrm{mul}}$-accessibility, this means showing that every multiset $M_1 <_{\mathrm{mul}} M \uplus \{\!|x|\!\}$ is in Acc. There are two cases: either $M_1 = M' \uplus \{\!|x|\!\}$ for some $M' <_{\mathrm{mul}} M$, and the claim follows from (b); or $M_1 = M \uplus \{\!|x_1, \ldots, x_m|\!\}$ with $x > x_1, \ldots, x_m$, then the claim follows by induction on m, using (b) in the base case and (a) in the induction step. □

It follows that, under the same assumptions, the transitive closure $<^+_{\mathrm{mul}}$ of $<_{\mathrm{mul}}$ is well-founded: for any relation R, R-accessibility and R^+-accessibility coincide.

Let now Σ be a signature, i.e., just a set whose elements will be understood as function symbols, with arbitrary, finite arity. The *terms* s, t, u, v, ..., are inductively defined as tuples $f(t_1, \ldots, t_n)$ of an element f of Σ and of finitely many terms t_1, ..., t_n. The base case is obtained when $n = 0$. There are no variables here, so our terms are the *ground terms* considered in the literature [16]. This is no loss of generality, as one can encode general terms as ground terms over a signature that includes all variables, understanding the variable term x as the application $x()$ to no argument. However, please do not confuse the latter (free) variables with the (μ-bound) variables that we introduce later.

Let $\approx$ be the relation defined inductively by: $f(s_1, \ldots, s_m) \approx g(t_1, \ldots, t_n)$ if and only if $f = g$, $m = n$, and there is a permutation π of $\{1, \ldots, n\}$ such that $s_{\pi(i)} \approx t_i$ for each i, $1 \leqslant i \leqslant n$. This is an equivalence relation, and relates terms that are equal up to permutations of arguments, anywhere in the term.

Call *precedence* any binary relation $<$ on Σ. The *multiset path ordering*, or *mpo*, $<^{\mathrm{mpo}}$ is defined inductively (together with an auxiliary relation $\ll$) by:

$$\frac{\exists i \cdot s_i \gtrsim^{\mathrm{mpo}} t}{f(s_1, \ldots, s_m) >^{\mathrm{mpo}} t}\ (Sub) \qquad \frac{\begin{array}{c} f(s_1, \ldots, s_m) \gg g(t_1, \ldots, t_n) \\ \forall j \cdot f(s_1, \ldots, s_m) >^{\mathrm{mpo}} t_j \end{array}}{f(s_1, \ldots, s_m) >^{\mathrm{mpo}} g(t_1, \ldots, t_n)}\ (Gt)$$

where $s \gtrsim^{\mathrm{mpo}} t$ abbreviates $s >^{\mathrm{mpo}} t$ or $s \approx t$, and here are the clauses for $\ll$:

$$\frac{f > g}{f(s_1, \ldots, s_m) \gg g(t_1, \ldots, t_n)}\ (\gg Fun) \quad \frac{\{\!|s_1, \ldots, s_m|\!\}\ (>^{\mathrm{mpo}})^+_{\mathrm{mul}}\ \{\!|t_1, \ldots, t_n|\!\}}{f(s_1, \ldots, s_m) \gg f(t_1, \ldots, t_n)}\ (\gg Args)$$

In other words, $\ll$ is the lexicographic product of $<$ and of $(<^{\mathrm{mpo}})^+_{\mathrm{mul}}$. The relation $\ll$ is a *lifting* (a notion called as such in [19], and which one can trace back to [31]), meaning that it is well-founded on the set $\overline{Acc}$ of terms of the form $f(s_1, \ldots, s_m)$ with f $<$-accessible and s_1, ..., s_m $<^{\mathrm{mpo}}$-accessible. Beware that this does *not* mean that any $\gg$-chain starting from a term $f(s_1, \ldots, s_m)$ with f $<$-accessible and s_1, ..., s_m $<^{\mathrm{mpo}}$-accessible is finite. It only means that any infinite such chain must eventually exit $\overline{Acc}$, i.e., reach a term $g(t_1, \ldots, t_n)$ where g is not $<$-accessible, or where some t_j is not $<^{\mathrm{mpo}}$-accessible. Intuitionistically,

we define the restriction $\ll_{|\overline{Acc}}$ of $\ll$ to $\overline{Acc}$ by $t \ll_{|\overline{Acc}} s$ iff $t \in \overline{Acc}$ and $s \in \overline{Acc}$ and $t \ll s$; and we note that every term in $\overline{Acc}$ is $\ll_{|\overline{Acc}}$-accessible.

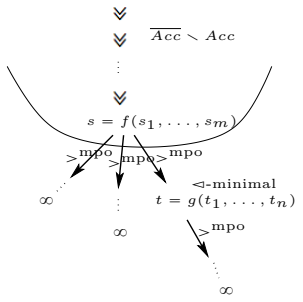
Replacing $\gg$ by other liftings would yield similar orderings: if we compare arguments lexicographically, for example, we would get the lexicographic path ordering (lpo), and mixing the two kinds yields the recursive path ordering (rpo) [13]. The following theorem is intuitionistic.

Proposition 1. *Every term whose function symbols are all $<$-accessible is $<^{mpo}$-accessible. In particular, if $<$ is well-founded, then $<^{mpo}$ is well-founded on terms.*

Proof. In the course of the proof, we shall need to observe that: $(*)$ for every $<^{\mathrm{mpo}}$-accessible term u, for every term t such that $u \approx t$, t is $<^{\mathrm{mpo}}$-accessible. This requires us to show first that if $u \approx t$ and $t >^{\mathrm{mpo}} s$, then $u >^{\mathrm{mpo}} s$, an easy induction on the definition of $<^{\mathrm{mpo}}$. We show $(*)$ by $<^{\mathrm{mpo}}$-induction on u, i.e., that for every t such that $u \approx t$, for every $s <^{\mathrm{mpo}} t$, s is $<^{\mathrm{mpo}}$-accessible; the assumptions imply $s <^{\mathrm{mpo}} u$, and the claim follows by induction hypothesis.

Let Acc be the set of $<^{\mathrm{mpo}}$-accessible terms, and W be the set of terms whose function symbols are all $<$-accessible. As above, we define $\overline{Acc}$ as the set of terms of the form $f(t_1, \ldots, t_n)$ such that f is $<$-accessible and whose arguments $t_1, \ldots, t_n$ are in Acc. We show that every $t \in W$ is in Acc, by structural induction on t. This means showing that for every $s \in \overline{Acc}$, s is in Acc.

We first give a classical argument, in the hope that it will be clearer. We shall need to use the immediate subterm relation $\lhd$, defined inductively by $g(t_1, \ldots, t_m) \rhd t_j$ for all $g, t_1, \ldots, t_m$ and j. This is a well-founded relation. Assume there is term $s \in \overline{Acc}$ that is not in Acc. In other words, the set $\overline{Acc} \setminus Acc$ is non-empty. Since $\ll$ is a lifting, it is well-founded on $\overline{Acc}$, hence on $\overline{Acc} \setminus Acc$: so there is a $\ll$-minimal element s in $\overline{Acc} \setminus Acc$. Since $s \notin Acc$, it starts an infinite $>^{\mathrm{mpo}}$-chain, so $s >^{\mathrm{mpo}} t$ for some $t \notin Acc$. Among these terms t we pick one that is $\lhd$-minimal: writing t as $g(t_1, \ldots, t_n)$, this assures us that for every j such that $s >^{\mathrm{mpo}} t_j$, $t_j \in Acc$.



The fact $s >^{\mathrm{mpo}} t$ is obtained by rule (Sub) or by rule (Gt). (Sub) is out of the question, though, since that would mean $s = f(s_1, \ldots, s_m)$ with some $s_i \gtrsim^{\mathrm{mpo}} t$; but $s \in \overline{Acc}$ implies $s_i \in Acc$, hence $t \in Acc$, either because $s_i \approx t$, using $(*)$, or because $s_i >^{\mathrm{mpo}} t$, using $<^{\mathrm{mpo}}$-inversion: contradiction. So rule (Gt) must have been used: $s \gg t = g(t_1, \ldots, t_n)$ with $s >^{\mathrm{mpo}} t_j$ for every j. Since s was chosen $\gg$-minimal, t cannot be in $\overline{Acc} \setminus Acc$, and since $t \notin Acc$, t is not in $\overline{Acc}$: so $t_j \notin Acc$ for some j. However, $s >^{\mathrm{mpo}} t_j$ together with the fact that t was $\lhd$-minimal implies $t_j \in Acc$, a contradiction.

We obtain an intuitionistic proof by replacing minimal counter-examples by induction principles. We wish to show that for every term $s \in \overline{Acc}$ then $s \in Acc$. Since $\ll$ is a lifting, $s \in \overline{Acc}$ is $\ll_{|\overline{Acc}}$-accessible, so $\ll_{|\overline{Acc}}$-induction applies and we obtain the following induction hypothesis: (a) for every $t \ll s$, if $t \in \overline{Acc}$ then

$t \in Acc$. Our goal is to prove that $s \in Acc$, i.e., that every $t <^{\mathrm{mpo}} s$ is in Acc. We show this by $\lhd$-induction on $t = g(t_1, \ldots, t_n)$, which means that we have the extra induction hypothesis: (b) for every j, if $t_j <^{\mathrm{mpo}} s$ then $t_j \in Acc$. If $t <^{\mathrm{mpo}} s$ was obtained by (Sub), then $s = f(s_1, \ldots, s_m)$ with $s_i \gtrsim^{\mathrm{mpo}} t$ for some i; since $s \in \overline{Acc}$, $s_i \in Acc$ hence $t \in Acc$, either by $(*)$ if $s_i \approx t$, or by $<^{\mathrm{mpo}}$-inversion if $s_i >^{\mathrm{mpo}} t$. If $t <^{\mathrm{mpo}} s$ was obtained by (Gt), then $s \gg t$ and $s >^{\mathrm{mpo}} t_j$ for every j. By (b), $t_j \in Acc$ for every j. Also, $s \gg t$ implies $f \gtrsim g$, and since $s \in \overline{Acc}$, f is $<$-accessible, hence also g: so $t = g(t_1, \ldots, t_n)$ is in $\overline{Acc}$. By (a), $t \in Acc$. □

I'm not claiming that the above proof is novel. This is the core of Theorem 1 of [26], later improved by Dawson and Goré [9]. Dershowitz [14] gives a broader perspective on this kind of results. That its proof is constructive is also one argument set forth in [26], and, as I've said already, this was made precise and implemented in Coq by Coupet-Grimal and Delobel [8].

I had also argued that the proof technique of [26] extended to prove abstract termination arguments, some of whose applied to graphs, for example. I'll develop this now for a new relation on a class of so-called *μ-terms*, defined by the following (pseudo-)grammar:

$$\begin{array}{rcll} s, t, u, v, \ldots & ::= & x & \text{variables} \\ & | & f(t_1, \ldots, t_n) & \text{applications, } f \in \Sigma, n \in \mathbb{N} \\ & | & \mu x = s \cdot t & \text{iterators.} \end{array}$$

The iterator $\mu x = s(x) \cdot t$ should be thought of as some kind of infinite term $\ldots s(\ldots s(s(t))\ldots)$. The variable x is bound in $\mu x = s \cdot t$, its scope is s. A term t is ground if and only if $\mathrm{fv}(t) = \varnothing$, where the set $\mathrm{fv}(t)$ of free variables of t is defined inductively by $\mathrm{fv}(x) = \{x\}$, $\mathrm{fv}(f(s_1, \ldots, s_m)) = \bigcup_{i=1}^{m} \mathrm{fv}(s_i)$, $\mathrm{fv}(\mu x = s \cdot s') = (\mathrm{fv}(s) \setminus \{x\}) \cup \mathrm{fv}(s')$. For instance, $\mu x = f(x) \cdot g(a)$ is a ground μ-term.

Again, we give ourselves a precedence $<$ on Σ. We extend the definition of $\approx$ by letting $\mu x = s \cdot s' \approx \mu x = t \cdot t'$ if and only if $s \approx t$ and $s' \approx t'$, and $x \approx x$ for every variable x. (We make an abuse of notation here and silently assume a form of α-renaming. A more correct definition would be: $\mu x = s \cdot s' \approx \mu y = t \cdot t'$ iff $s[x := z] \approx t[y := z]$ and $s' \approx t'$, for z a fresh variable. We shall make similar abuses of notation in rules $(\mu Gt\mu)$, $(\mu\ll)$ and $(\mu\ll\mu)$ below, to avoid clutter.) We take the same rules defining $<^{\mathrm{mpo}}$ and $\ll$ as above, and add the following to also compare variables and iterations, either together or with other terms:

$$\frac{(t \text{ ground } \mu\text{-term})}{x >^{\mathrm{mpo}} t}\,(Var) \qquad \frac{s' \gtrsim^{\mathrm{mpo}} t}{\mu x = s \cdot s' >^{\mathrm{mpo}} t}\,(\mu Sub)$$

$$\frac{\mu x = s \cdot s' \gg g(t_1, \ldots, t_n) \quad \forall j \cdot \mu x = s \cdot s' >^{\mathrm{mpo}} t_j}{\mu x = s \cdot s' >^{\mathrm{mpo}} g(t_1, \ldots, t_n)}\,(\mu Gt) \qquad \frac{\mu x = s \cdot s' \gg \mu x = t \cdot t' \quad \mu x = s \cdot s' >^{\mathrm{mpo}} t'}{\mu x = s \cdot s' >^{\mathrm{mpo}} \mu x = t \cdot t'}\,(\mu Gt\mu)$$

$$\frac{s >^{\mathrm{mpo}} g(t_1, \ldots, t_n)}{\mu x = s \cdot s' \gg g(t_1, \ldots, t_n)}\,(\mu\ll) \qquad \frac{s >^{\mathrm{mpo}} t}{\mu x = s \cdot s' \gg \mu x = t \cdot t'}\,(\mu\ll\mu)$$

The unusual rule (Var) states that *every* ground μ-term is strictly smaller than any variable. This allows us to check, for example, that $\mu x = f(x) \cdot g(a) >^{\mathrm{mpo}} f(f(f(f(g(a)))))$, where a is a constant: using (μGt) and $(\mu\ll)$, this requires us to check two premises, of which one is $f(x) >^{\mathrm{mpo}} f(f(f(f(g(a)))))$; the latter follows, using (Gt), from $x >^{\mathrm{mpo}} f(f(f(g(a))))$, and this, in turn, is an instance of (Var). We leave the rest of the verification to the reader.

The above rules are probably not the ones one would have imagined. In particular, it would seem natural to consider $\mu x = s(x) \cdot s'$ and $s(\mu x = s \cdot s')$ as equivalent. This would suggest the following alternative to (μGt): to prove $\mu x = s(x) \cdot s' >^{\mathrm{mpo}} t$ (where $t = g(t_1, \ldots, t_n)$, and for simplicity we assume both sides of the inequality to be ground), prove $s(\mu x = s(x) \cdot s') >^{\mathrm{mpo}} t$ and $\forall j \cdot \mu x = s \cdot s' >^{\mathrm{mpo}} t_j$. Instead of proving $s(\mu x = s(x) \cdot s') >^{\mathrm{mpo}} t$, (μGt) (together with $(\mu\ll)$) only requires us to prove $s(x) >^{\mathrm{mpo}} t$, a seemingly much weaker statement, since x is not just greater than or equal to $\mu x = s(x) \cdot s'$, but strictly greater than *any* ground term by (Var). Although they are not what we would imagined at first, these are the rules that arise from our study of the topological Kruskal Theorem (Section 5).

The following is new, and probably useful in other contexts. Our proof is intuitionistic. The proof is similar to Proposition 1, or to Theorem 1 of [26], but we need a few easy additional arguments near the end of the proof.

Theorem 1. *Every μ-term whose function symbols are all $<$-accessible is $<^{mpo}$-accessible. In particular, if $<$ is well-founded, then $<^{mpo}$ is well-founded on μ-terms.*

Proof. One might think that Theorem 1 is an easy consequence of Proposition 1: encode $\mu x = s \cdot s'$ as the ordinary term $\mu(s, s')$, and the variable x as $x()$, and extend the precedence appropriately. This strategy does not work, as for example (Var) requires $x >^{\mathrm{mpo}} \mu x = f(x) \cdot g(a)$. In the encoding, this would force $x >^{\mathrm{mpo}} \mu(f(x), g(a))$, which is plainly false, since $\mu(f(x), g(a)) >^{\mathrm{mpo}} x$.

We imitate the proof of Proposition 1. Again, we have: $(*)$ for every $<^{\mathrm{mpo}}$-accessible μ-term u, for every μ-term t such that $u \approx t$, t is $<^{\mathrm{mpo}}$-accessible. Define the *immediate subterms* of a μ-term in the expected way, as follows: the immediate subterms of $g(t_1, \ldots, t_m)$ are $t_1, \ldots, t_n$, the immediate subterms of $\mu x = s \cdot s'$ are s and s', and variables have no immediate subterms. We need to define $\lhd$ slightly differently, inductively, by: (i) $g(t_1, \ldots, t_m) \rhd t_j$ for all $g \in \Sigma$, μ-terms $t_1, \ldots, t_m$ and j; (ii) $x \rhd t$ for every variable x and ground μ-term t; (iii) $\mu x = s \cdot s' \rhd s'$ (not s!).

We first show that $\lhd$ is well-founded. This is done in several steps. We first show that every ground μ-term t is $\lhd$-accessible, by induction on t; crucially, if $\mu x = s \cdot s'$ is ground and $\mu x = s \cdot s' \rhd s'$, then s' is ground and the induction hypothesis applies. We then do a secondary induction to establish that every μ-term is $\lhd$-accessible, using the previous claim in the case of variables.

Let Acc be the set of $<^{\mathrm{mpo}}$-accessible μ-terms, and W be the set of μ-terms whose function symbols are all $<$-accessible. Say that a μ-term is *head accessible* if and only if it is a variable, an iterator $\mu x = s \cdot s'$ with s head accessible, or an

application $f(s_1, \ldots, s_m)$ with f $<$-accessible. The point is: (†) if $s \gg g(t_1, \ldots, t_n)$ and s is head accessible, then g is $<$-accessible. This is proved by induction on the proof of $s \gg g(t_1, \ldots, t_n)$; the base case is when s is of the form $f(s_1, \ldots, s_m)$, where necessarily $f \geqslant g$, and f is $<$-accessible since s is head accessible.

We also define $\overline{Acc}$ as the set of head accessible μ-terms s whose immediate subterms are all in Acc.

Again, $\ll$ is a lifting, namely, every term in $\overline{Acc}$ is $\ll_{|\overline{Acc}}$-accessible. This is proved in two steps. We first show that every variable x is $\ll_{|\overline{Acc}}$-accessible (vacuous: $x \gg t$ for no μ-term t), and that every application $f(s_1, \ldots, s_m)$ in $\overline{Acc}$ is $\ll_{|\overline{Acc}}$-accessible: this is by double induction ($<$-induction on f, then $(<^{\mathrm{mpo}})^+_{\mathrm{mul}}$-induction on $\{\!|s_1, \ldots, s_m|\!\}$), using the fact that $f(s_1, \ldots, s_m) \gg t$ implies that $t = g(t_1, \ldots, t_n)$ with $f > g$ or $[f = g$ and $\{\!|s_1, \ldots, s_m|\!\}$ $(>^{\mathrm{mpo}})^+_{\mathrm{mul}}$ $\{\!|t_1, \ldots, t_n|\!\}]$. We then show that every iterator $\mu x = s \cdot s'$ in $\overline{Acc}$ is $\ll_{|\overline{Acc}}$-accessible, by $<^{\mathrm{mpo}}$-induction on s. To do so, we consider the μ-terms $t \in \overline{Acc}$ such that $t \ll \mu x = s \cdot s'$. Those obtained by rule ($\mu\ll$) are $\ll_{|\overline{Acc}}$-accessible by the first step, and those obtained by rule ($\mu\ll\mu$) are $\ll_{|\overline{Acc}}$-accessible by the induction hypothesis.

Let us pause a minute, and observe the following, called 'Property 1' in [26]. For all μ-terms s, t, if $s >^{\mathrm{mpo}} t$ then either:

(i) $s \rhd u \gtrsim^{\mathrm{mpo}} t$ for some μ-term u, or:
(ii) $s \gg t$ and $s >^{\mathrm{mpo}} u$ for every $u \lhd t$.

Case (i) happens in case $s >^{\mathrm{mpo}} t$ was derived using (Sub), (μSub), or (Var). Case (ii) happens in case it was derived using (Gt), (μGt), or ($\mu Gt\mu$).

We now show that every $t \in W$ is in Acc, by structural induction on t. This means showing that for every $s \in \overline{Acc}$, s is in Acc. Since $\ll$ is a lifting, $s \in \overline{Acc}$ is $\ll_{|\overline{Acc}}$-accessible, so $\ll_{|\overline{Acc}}$-induction applies and we obtain the following induction hypothesis: (a) for every $t \ll s$, if $t \in \overline{Acc}$ then $t \in Acc$. Our goal is to prove that $s \in Acc$, i.e., that every $t <^{\mathrm{mpo}} s$ is in Acc. We show this by $\lhd$-induction on t, which means that we have the extra induction hypothesis: (b) for every $u \lhd t$, if $u <^{\mathrm{mpo}} s$ then $u \in Acc$. Since $t <^{\mathrm{mpo}} s$, either (i) or (ii) is true. If (i) holds, then $s \rhd u \gtrsim^{\mathrm{mpo}} t$, so $u \in Acc$ since $s \in \overline{Acc}$ and $u \lhd s$; therefore $t \in Acc$, by ($*$) if $u \approx t$, by $<^{\mathrm{mpo}}$-inversion if $u >^{\mathrm{mpo}} t$. So assume (ii). We claim that t is in $\overline{Acc}$. This is trivial if t is a variable. If t is an application $g(t_1, \ldots, t_n)$ then for each j, $t_j \lhd s$, so by taking $u = t_j$ in (b), we obtain that t_j is in Acc; g is $<$-accessible since $s \gg g(t_1, \ldots, t_n)$, using (†); so $t \in \overline{Acc}$. If t is an iterator $\mu x = t_1 \cdot t_2$, then (b) only implies that t_2 is in Acc. To obtain $t_1 \in Acc$, we realize that we can only have derived $s \gg t$ by rule ($\mu\ll\mu$), which implies that s is of the form $\mu x = s_1 \cdot s_2$ with $s_1 >^{\mathrm{mpo}} t_1$: since $s \in \overline{Acc}$, s_1 is in Acc hence t_1 is in Acc by $<^{\mathrm{mpo}}$-inversion. In any case, t is in $\overline{Acc}$. Since also $t \ll s$, (a) applies, so that t is in Acc, as desired. □

4 A constructive proof of Higman's Lemma

It is time to apply all this and prove the topological Higman Lemma. Given a set X with a quasi-ordering $\leqslant$, the *embedding* quasi-ordering $\leqslant^*$ on X^* is the small-

est relation such that $x_1 \leqslant y_1, \ldots, x_n \leqslant y_n$ imply $x_1 \ldots x_n \leqslant w_0 y_1 w_1 \ldots w_{n-1} y_n w_n$, where $w_0, w_1, \ldots, w_{n-1}, w_n$ are arbitrary words in X^*. In other words, to go down in $\leqslant^*$, remove some letters and replace the others by smaller ones. Higman's Lemma states that if $\leqslant$ is wqo, then so is $\leqslant^*$. The topological Higman Lemma states that if X is a Noetherian topological space, then X^* with the word topology is Noetherian, too. We have already discussed this in Section 2.

Step (A) of our proof plan consists in discovering an S-representation of X^*, for X Noetherian. (Step (A) is *not* constructive.) In [22], we defined an *S-representation* of a Noetherian space X as a tuple $(S, \mathcal{S}[\![\_]\!], \trianglelefteq, \tau, \wedge)$, where S is a set of elements, meant to denote the irreducible closed subsets of X, through the denotation map $[\![\_]\!]$, $\trianglelefteq$ denotes inclusion, τ represents the whole space, and $\wedge$ implements intersection. We change this slightly, and replace $\trianglelefteq$ by its strict part $\sqsubset$ [2]. Hence, call *S-representation* of a Noetherian space X any tuple $(S, \sqsubset, \tau, \wedge)$, where S is a set, $[\![\_]\!] : S \to \mathcal{S}(X)$ is a bijective denotation function, $\sqsubset$ is a binary relation on S denoting strict inclusion (i.e., $a \sqsubset b$ iff $[\![a]\!] \subset [\![b]\!]$), τ is a finite subset of S denoting the whole of X ($[\![\tau]\!] = X$, where we extend the notation $[\![a]\!]$ for $a \in S$ to $[\![A]\!]$ for $A \in \mathbb{P}(S)$, by letting $[\![A]\!] = \bigcup_{a \in A} [\![a]\!]$), and for all $a, b \in S$, $a \wedge b$ is a finite subset of S denoting their intersection ($[\![a \wedge b]\!] = [\![a]\!] \cap [\![b]\!]$). When X is Noetherian, $\sqsubset$ will be well-founded (property $(\downarrow)$), τ will exist by property (T), and $\wedge$ will make sense because of property (W).

Since $[\![a']\!]$ is irreducible for every $a' \in A'$, the inclusion $[\![A]\!] \subseteq [\![A']\!]$ is equivalent to $A \sqsubseteq^\flat A'$, where we write $\sqsubseteq$ for the union of $\sqsubset$ and $=$, and the Hoare quasi-ordering $\sqsubseteq^\flat$ is defined by: for every $a \in A$, there is an $a' \in A'$ such that $a \sqsubseteq a'$. Since A, A' are antichains, one can encode them as multisets. A moment's notice shows that the strict part of $\sqsubseteq^\flat$ is just ${\sqsubset_{\mathrm{mul}}}^+$. This will be used to compare antichains A, A' below.

$$\frac{e\boldsymbol{P} \sqsubseteq^{\mathrm{w}} \boldsymbol{P'}}{e\boldsymbol{P} \sqsubset^{\mathrm{w}} e'\boldsymbol{P'}}\ (\mathrm{w1}) \qquad \frac{a \sqsubset a' \quad \boldsymbol{P} \sqsubseteq^{\mathrm{w}} \boldsymbol{P'}}{a^?\boldsymbol{P} \sqsubset^{\mathrm{w}} a'^?\boldsymbol{P'}}\ (\mathrm{w2}) \qquad \frac{\boldsymbol{P} \sqsubset^{\mathrm{w}} \boldsymbol{P'}}{a^?\boldsymbol{P} \sqsubset^{\mathrm{w}} a^?\boldsymbol{P'}}\ (\mathrm{w3})$$

$$\frac{\forall i \cdot e_i \sqsubset^{\mathrm{e}} A'^* \quad \boldsymbol{P} \sqsubseteq^{\mathrm{w}} \boldsymbol{P'}}{e_1 \ldots e_k\boldsymbol{P} \sqsubset^{\mathrm{w}} A'^*\boldsymbol{P'}}\ (\mathrm{w4}) \qquad \frac{\boldsymbol{P} \sqsubset^{\mathrm{w}} \boldsymbol{P'}}{A^*\boldsymbol{P} \sqsubset^{\mathrm{w}} A^*\boldsymbol{P'}}\ (\mathrm{w5})$$

Fig. 1. Deciding strict inclusion between word-products

Given an S-representation $(S, \mathcal{S}[\![\_]\!], \sqsubset, \tau, \wedge)$ of X, Theorem 6.14 of [22] gives us an S-representation $(S^{\mathrm{w}}, \mathcal{S}[\![\_]\!]^{\mathrm{w}}, \sqsubset^{\mathrm{w}}, \tau^{\mathrm{w}}, \wedge^{\mathrm{w}})$ of X^*. S^{w} is a set of so-called

[2] In all rigor, we should also include the associated congruence $\equiv$, defined by $a \equiv b$ iff $a \trianglelefteq b$ and $b \trianglelefteq a$. We silently assume we are working in the quotient of the S-representation by $\equiv$. In proof assistants such as Coq, this is not an option, and the standard solution is to use setoid types. In any case, considering $\equiv$ explicitly would make our exposition too complex, and we shall therefore avoid it. We also change the notation from $\trianglelefteq$ to $\sqsubset$ to avoid a conflict with the relations $\rhd$ of Section 3

word-products, first invented in the setting of forward coverability procedures for lossy channel systems [2]. Define the *atomic expressions* as $a^?$ with $a \in S$ (denoting the set of words with at most one letter in $[\![a]\!]$), and A^* with A a non-empty finite antichain of S (denoting the set of words, of arbitrary length, whose letters are all in $[\![A]\!]$). The *word-products* $\boldsymbol{P}$, $\boldsymbol{P'}$, ..., are the finite sequences $e_1 e_2 \dots e_n$ of atomic expressions, denoting the concatenations of words in the denotations of $e_1, e_2, \dots, e_n$, and we define S^{w} as those that are *reduced*, namely those where $[\![e_i e_{i+1}]\!]^{\mathrm{w}}$ is included neither in $[\![e_i]\!]^{\mathrm{w}}$ nor in $[\![e_{i+1}]\!]^{\mathrm{w}}$ for every i. Inclusion between word-products is decidable, using simple formulae given for example in [22, Lemma 6.8, Lemma 6.9], and this allows us to give computable predicates that sieve out the non-reduced word-products. We are more interested in the relation $\sqsubset^{\mathrm{w}}$. Two reduced word-products, that is, two elements of S^{w}, have equal denotations iff they are equal. One can show that the strict inclusion relation $\sqsubset^{\mathrm{w}}$ on reduced word-products is defined inductively by the rules of Figure 1. We write $\boldsymbol{P} \sqsubseteq^{\mathrm{w}} \boldsymbol{P'}$ for $\boldsymbol{P} \sqsubset^{\mathrm{w}} \boldsymbol{P'}$ or $\boldsymbol{P} = \boldsymbol{P'}$. We also define the auxiliary relation $\sqsubset^{\mathrm{e}}$ (strict inclusion of atomic expressions) by: $a^? \sqsubset^{\mathrm{e}} a'^?$ iff $a \sqsubset a'$; $a^? \sqsubset^{\mathrm{e}} A'^*$ iff $a \sqsubseteq a'$ for some $a' \in A'$; $A^* \sqsubset^{\mathrm{e}} a'^?$ never; and $A^* \sqsubset^{\mathrm{e}} A'^*$ iff $A \sqsubset^+_{\mathrm{mul}} A'$. We define τ^{w} as the antichain $\{\tau^*\}$, and omit the definition of $\wedge^{\mathrm{w}}$ [22, Lemma 6.11].

We now embark on step (B) of our proof plan. Contrarily to step (A), we must pay attention to only invoke *constructive* arguments. So forget everything we have done in step (A), except for the final result. Say that $(S, \sqsubset, \tau, \wedge)$ is a *constructive S-representation* (without reference to X) if and only if S is a set with a strict ordering $\sqsubset$, and where: $(\downarrow)$ $\sqsubset$ is well-founded; (T) $S = \downarrow \tau$; (W) for all $a, b \in S$, $\downarrow a \cap \downarrow b = \downarrow(a \wedge b)$; $\sqsubseteq$ stands for the union of $\sqsubset$ and $=$, $\downarrow A$ for the downward closure of a subset A of S with respect to $\sqsubseteq$, and $\downarrow a$ for $\downarrow\{a\}$.

We now *posit* $(S^{\mathrm{w}}, \sqsubset^{\mathrm{w}}, \tau^{\mathrm{w}}, \wedge^{\mathrm{w}})$ by the syntax given above, in step (A). S^{w} is the set of reduced word-products over S, $\sqsubset^{\mathrm{w}}$ is defined inductively by (w1)–(w5), $\tau^{\mathrm{w}} = \{\tau^*\}$, and we define $\wedge^{\mathrm{w}}$ by the recursive formula of [22, Lemma 6.11].

Theorem 2. *If $(S, \sqsubset, \tau, \wedge)$ is a constructive S-representation, then so is $(S^{\mathrm{w}}, \sqsubset^{\mathrm{w}}, \tau^{\mathrm{w}}, \wedge^{\mathrm{w}})$.*

Proof. (Sketch.) There is a boring part, consisting in checking that $\sqsubset^{\mathrm{w}}$ is a strict ordering, and that properties (T) and (W) hold. We omit it here. The interesting part is checking that $\sqsubset^{\mathrm{w}}$ is well-founded. Define a mapping μ from atomic expressions to pairs $(i, A) \in \{0, 1\} \times \mathcal{M}(S)$ by $\mu(a^?) = (0, \{\!|a|\!\})$, $\mu(A^*) = (1, A)$, and order them by the lexicographical product $\prec$ of the ordering $0 < 1$ and of $\sqsubset^+_{\mathrm{mul}}$. Extend μ to word-products by $\mu(e_1 \dots e_n) = \{\!|\mu(e_1), \dots, \mu(e_n)|\!\}$. In other words, we look at word-products as though they were multisets of atomic expressions, where the latter as read as multisets of letters from S, plus a tag, 0 or 1. It is fairly easy to show that for all reduced word-products $\boldsymbol{P}$, $\boldsymbol{P'}$, if $\boldsymbol{P} \sqsubset \boldsymbol{P'}$ then $\mu(\boldsymbol{P}) \prec^+_{\mathrm{mul}} \mu(\boldsymbol{P'})$, by induction on the structure of a proof of $\boldsymbol{P} \sqsubset \boldsymbol{P'}$. By Lemma 1, $\prec^+_{\mathrm{mul}}$ is well-founded. By $\prec^+_{\mathrm{mul}}$-induction on $\mu(\boldsymbol{P})$, $\boldsymbol{P}$ is then $\sqsubset$-accessible, for every $\boldsymbol{P} \in S^{\mathrm{w}}$. □

The statement of Theorem 2 seems very far from Higman's Lemma. Call *constructive Noetherian space* any tuple $(X, T, \prec, \varepsilon)$, where $\prec$ is a well-founded ordering on the set T (T is the *cotopology*) whose reflexive closure $\preceq$ makes T a distributive lattice (this much implies classically that $(T, \preceq)$ is the lattice of closed subsets of some Noetherian space, up to isomorphism), and $\varepsilon \subseteq X \times T$ (membership) is a binary relation such that for all $A, B \in T$, $A \preceq B$ iff for every $x \varepsilon A$, $x \varepsilon B$. We observe the following:

(a) Given a constructive S-representation $(S, \sqsubset, \tau, \wedge)$, we think of elements of S as irreducible closed subsets of some Noetherian space X, and we can build all closed sets as finite unions thereof. We encode the latter as finite antichains, hence as multisets. Letting $T = \mathcal{M}(S)$, $\prec = \sqsubset^+_{\mathrm{mul}}$ then defines the *canonical cotopology* on $(S, \sqsubset, \tau, \wedge)$. Any subset X of S then gives rise to a constructive Noetherian space $(X, \mathcal{M}(S), \sqsubset^+_{\mathrm{mul}}, \varepsilon)$, where $x \varepsilon M$ iff $\{\!|x|\!\}\ (\sqsubset_{\mathrm{mul}})^*\ M$.
(b) Conversely, every cotopology $(T, \prec)$ gives rise to a trivial constructive S-representation $(S, \sqsubset, \tau, \wedge)$ where $S = T$, $\sqsubset$ is $\prec$, $\tau = \{\top\}$ where $\top$ is the top element of T, and $A \wedge B = \{A \sqcap B\}$ where $\sqcap$ is meet in T.

Given (a) and (b), Theorem 2 and Lemma 1 then imply:

Corollary 1 (Topological Higman Lemma, Constructively). *For every constructive Noetherian space $(X, T, \prec, \varepsilon)$, $(X^*, \mathcal{M}(T^{\mathrm{w}}), (\prec^{\mathrm{w}}_{mul})^+, \varepsilon^{\mathrm{w}})$ is a constructive Noetherian space, with $w\ \varepsilon^{\mathrm{w}}\ M$ iff $\{\!|\eta^{\mathrm{w}}(w)|\!\}\ (\prec^{\mathrm{w}}_{mul})^*\ M$, where $\eta^{\mathrm{w}}(x_1\ x_2 \ldots x_m) = x_1^? x_2^? \ldots x_m^?$.*

This implies the usual form of Higman's Lemma, by similar arguments as in [37]. Assuming a decidable constructive wqo $\leqslant$ on a set X, one can show, constructively, that the antichains $E = \{x_1, \ldots, x_n\}$ (interpreted as the downward closed set $X \setminus \uparrow E$) are the elements of a cotopology, where $\prec$ is the strict part of $\preceq$; we let $E \preceq E'$ iff $X \setminus \uparrow E \subseteq X \setminus \uparrow E'$, iff for every $y \in X'$, there is an $x \in E$ such that $x \leqslant y$; and $x\ \varepsilon\ E$ iff $x \in X \setminus \uparrow E$, iff for every $y \in E$, $y \not\leqslant x$. Recall that a finite sequence $w_1, \ldots, w_n$ in X^* is *bad* iff $w_i \leqslant^* w_j$ for no $i < j$. Following Murthy and Russell, we show that the converse of the prefix ordering on bad sequences $w_1, \ldots, w_n$ is well-founded, by $(\prec^{\mathrm{w}}_{\mathrm{mul}})^+$-induction on the closed subset $X^* \setminus \uparrow\{w_1, \ldots, w_n\}$—this induction principle is given to us by Corollary 1. The set $X^* \setminus \uparrow\{w_1, \ldots, w_n\}$ is represented, constructively, as the finite intersection of the sets $X \setminus \uparrow w_i$, using the $\top$ and $\sqcap$ operations of the cotopology; writing w_i as the word $x_1 x_2 \ldots x_m$, $X \setminus \uparrow w_i$ is the word-product $(X \setminus \uparrow x_1)^* X^? (X \setminus \uparrow x_2)^* X^? \ldots X^? (X \setminus \uparrow x_m)^*$ if $m \geqslant 1$, the empty set otherwise [22, Lemma 6.1]. This is the core of Murthy and Russell's proof:

Theorem 3 (Murthy-Russell). *Let X be a set with a decidable constructive wqo $\leqslant$. Then $\leqslant^*$ is a (decidable) constructive wqo on X^*.*

5 A constructive proof of Kruskal's Theorem

We use the same strategy for trees, i.e., first-order terms. Given a set X with a quasi-ordering $\leqslant$, the (tree) *embedding* quasi-ordering $\preceq_\leqslant$ is inductively defined

by $s \preceq_\leqslant t$, where $s = f(s_1, \ldots, s_m)$ and $t = g(t_1, \ldots, t_n)$, iff $s \preceq_\leqslant t_j$ for some j, or $f \leqslant g$ and $s_1 \ldots s_m \preceq^*_\leqslant t_1 \ldots t_n$; note the use of the word embedding ordering $\preceq^*_\leqslant$ on lists of immediate subterms, considered as words.

Given a constructive S-representation $(S, \sqsubset, \tau, \wedge)$, we define a set S^{t} of regular expressions on trees (the *tree-products* P, Q, ...) inductively, as follows. Let $\Box$ be a fresh constant. The elements P of S^{t} are the *tree steps* $a^{[?]}(\boldsymbol{P})$, where $a \in S$ and $\boldsymbol{P}$ is a reduced word-product over S^{t}, and the *tree iterators* $(\bigcup_{i=1}^m a_i(\boldsymbol{Q}_i))^{[*]}.{}_\Box A$, where A is a finite set of elements of S^{t}, $a_i \in S$, and $\boldsymbol{Q}_i$ is a word-product over $S \cup \{\Box\}$, which is either equal to $\{\Box\}^*$ (which we shall simply write $\Box^*$), or of the form $\boldsymbol{Q}_{i1}\Box^?\boldsymbol{Q}_{i2}\Box^? \ldots \Box^?\boldsymbol{Q}_{ik_i}$ where all $\boldsymbol{Q}_{ij}$s are reduced word-products over S^{t} [22, Lemma 9.20].

Tree steps are the analogue of $a^?$ for words. Intuitively, $a^{[?]}(e_1e_2)$ will contain all the terms of the form $f(t_1, t_2)$ with f in (the denotation of) a, t_1 in e_1 and t_2 in e_2, plus all the terms from e_1 and from e_2. Of course, e_1e_2 is not a word-product, but, say, $e_1^?e_2^?$ is, and $a^{[?]}(e_1^?e_2^?)$ will contain not just the terms above, but also the terms of the form $f(t_1)$, $f(t_2)$ and $f()$, with $f \in a$, $t_1 \in e_1$, $t_2 \in e_2$.

Tree iterators $(\bigcup_{i=1}^m a_i(\boldsymbol{Q}_i))^{[*]}.{}_\Box A$ define the following language L, inductively, by the following two rules. First, A is a set $\{P_1, \ldots, P_n\}$ of tree-products, and every element of any P_i is in L. Second, given any term t in the set denoted by $a_i^{[?]}(\boldsymbol{Q}_i)$, the term obtained from t by replacing each occurrence of $\Box$ by a (possibly different) term from L is again in L. For example, if P contains terms t_1, t_2 and t_3, then $(a(\Box^?\Box^?))^{[*]}.{}_\Box\{P\}$ will contain $f(t_1, t_1)$, $f(t_2, t_2)$, but also $f(t_1, t_2)$, $f(t_1, f(t_2, t_2))$, $f(f(t_1, f(t_2, t_1)), f(t_1, t_3))$, for f in a, among other terms. As another example, $(a(\Box^*))^{[*]}.{}_\Box\varnothing$ is the set of terms all of whose function symbols are in a.

Much as we only considered reduced word-products in Section 4, we shall restrict to *canonical* tree-products here. The tree-products considered in [22, Section 9] are *normal* tree-products, a closely related notion. Normality requires, for example, that in a tree iterator, $(\bigcup_{i=1}^m a_i(\boldsymbol{Q}_i))^{[*]}.{}_\Box A$, (*i*) m is non-zero, (*ii*) $\Box$ occurs in every $\boldsymbol{Q}_i$, and (*iii*) A contains just one tree-product in case every $\boldsymbol{Q}_i$ is $\Box$-*linear*, i.e., does not contain $\Box^*$ and only one occurrence of $\Box^?$. Here, we need to require that the support $\operatorname{supp} \boldsymbol{Q}_i$ of every $\boldsymbol{Q}_i$, namely, the set of ($\Box$-free) terms t such that the one-element sequence t is in the denotation of $\boldsymbol{Q}_i$, is entirely contained in the denotation of A. This is easy to ensure, by adding the required tree-products from $\operatorname{supp} \boldsymbol{Q}_i$ to A... but breaks (*iii*). Instead, we define *canonical* tree iterators as those satisfying (*i*), (*ii*), (*iii′*): if every $\boldsymbol{Q}_i$ is $\Box$-linear, then A denotes the union of $\bigcup_{i=1}^m \operatorname{supp} \boldsymbol{Q}_i$ with at most one tree-product; we also require: (*iv*) the tree steps $a_i(\boldsymbol{Q}_i)$ are pairwise incomparable, (*v*) the elements of A are pairwise incomparable, and (*vi*) $(\bigcup_{i=1}^m a_i(\boldsymbol{Q}_i))^{[*]}.{}_\Box A$ must not be included in $\bigcup_{i=1}^m \operatorname{supp} \boldsymbol{Q}_i$. Similarly, we define *canonical* trees steps as those $a^?(\boldsymbol{P})$ that are not included in $\operatorname{supp} \boldsymbol{P}$. Every tree-product can be *canonicalized*, i.e., transformed to a canonical one with the same denotation.

One can decide inclusion of canonical tree-products, in polynomial time, and also compute finite intersections thereof ($\wedge^{\mathrm{t}}$, τ^{t}), using formulae given in [22, Section 9], plus canonicalization. From these formulae, we deduce the rules for

$$\frac{(P' \in S^{\mathrm{t}})\quad \exists P \in \mathrm{sub}(\boldsymbol{P}) \cdot P' \sqsubseteq^{\mathrm{t}} P}{P' \sqsubset^{\mathrm{t}} a^{[?]}(\boldsymbol{P})} \qquad \frac{(P' \in S^{\mathrm{t}})}{P' \sqsubset^{\mathrm{t}} \Box} \qquad \frac{a'^{[?]}(\boldsymbol{P}') \Subset^{\mathrm{t}} a^{[?]}(\boldsymbol{P})\quad \forall P' \in \mathrm{sub}(\boldsymbol{P}') \cdot P' \sqsubset^{\mathrm{t}} a^{[?]}(\boldsymbol{P})}{a'^{[?]}(\boldsymbol{P}') \sqsubset^{\mathrm{t}} a^{[?]}(\boldsymbol{P})}$$

$$\frac{a' \sqsubset a}{a'^{[?]}(\boldsymbol{P}) \Subset^{\mathrm{t}} a^{[?]}(\boldsymbol{P})} \qquad \frac{\boldsymbol{P}'\ (\sqsubset^{\mathrm{t}})^{\mathrm{w}}\ \boldsymbol{P}}{a^{[?]}(\boldsymbol{P}') \Subset^{\mathrm{t}} a^{[?]}(\boldsymbol{P})}$$

$$\frac{(P' \in S^{\mathrm{t}})\quad \exists P \in A \cdot P' \sqsubseteq^{\mathrm{t}} P}{P' \sqsubset^{\mathrm{t}} (\bigcup_{i=1}^{m} a_i(\boldsymbol{Q}_i))^{[*]}.\Box A} \qquad \frac{\exists i \cdot a'^{[?]}(\boldsymbol{P}') \Subset^{\mathrm{t}} a_i^{[?]}(\boldsymbol{Q}_i)\quad \forall P' \in \mathrm{sub}(\boldsymbol{P}') \cdot P' \sqsubset^{\mathrm{t}} (\bigcup_{i=1}^{m} a_i(\boldsymbol{Q}_i))^{[*]}.\Box A}{a'^{[?]}(\boldsymbol{P}') \sqsubset^{\mathrm{t}} (\bigcup_{i=1}^{m} a_i(\boldsymbol{Q}_i))^{[*]}.\Box A}$$

$$\frac{\{\!|a_j'^{[?]}(\boldsymbol{Q}'_j) \mid 1 \leqslant j \leqslant n|\!\} \Subset^{\mathrm{t}}_{\mathrm{mul}} \{\!|a_i^{[?]}(\boldsymbol{Q}_i) \mid 1 \leqslant i \leqslant n|\!\}\quad \forall P' \in A' \cdot P' \sqsubset^{\mathrm{t}} (\bigcup_{i=1}^{m} a_i(\boldsymbol{Q}_i))^{[*]}.\Box A}{(\bigcup_{j=1}^{n} a'_j(\boldsymbol{Q}'_j))^{[*]}.\Box A' \sqsubset^{\mathrm{t}} (\bigcup_{i=1}^{m} a_i(\boldsymbol{Q}_i))^{[*]}.\Box A}$$

Fig. 2. Deciding strict inclusion between tree-products

strict inclusion $\sqsubset^{\mathrm{t}}$ on S^{t}—to be precise, on S^{t} union $\{\Box\}$, where $\Box$ will be topmost—given in Figure 2. We again write $\sqsubseteq^{\mathrm{t}}$ for the reflexive closure of $\sqsubset^{\mathrm{t}}$. For a word-product $\boldsymbol{P}$ over $S^{\mathrm{t}} \cup \{\Box\}$, define $\mathrm{sub}(\boldsymbol{P})$ (denoting the support of $\boldsymbol{P}$) by: $\mathrm{sub}(e_1 \dots e_n) = \bigcup_{i=1}^{n} \mathrm{sub}(e_i)$, $\mathrm{sub}(P^?) = \{P\}$ for $P \in S^{\mathrm{t}}$, $\mathrm{sub}(\Box^?) = \varnothing$, $\mathrm{sub}(A^*) = A$ for A an antichain in S^{t}, $\mathrm{sub}(\Box^*) = \varnothing$. We also use an auxiliary relation $\Subset^{\mathrm{t}}$, which should be reminiscent of $\ll$. The whole definition should, in fact, remind you of the definition of $<^{\mathrm{mpo}}$ on μ-terms, and this is no accident.

Theorem 4. *If $(S, \sqsubset, \tau, \wedge)$ is a constructive S-representation, then so is $(S^{\mathrm{t}}, \sqsubset^{\mathrm{t}}, \tau^{\mathrm{t}}, \wedge^{\mathrm{t}})$.*

Proof. (Sketch.) Only property (↓) deserves attention. Define a syntactic translation from $P \in S^{\mathrm{t}}$ to μ-terms $\langle P \rangle$, as follows. Our signature consists of all elements of S, plus one fresh function symbol $\mathtt{u}$ (union). The following formulae also define $\langle \_ \rangle$ translations of various other syntactic categories, e.g., $\langle \boldsymbol{P} \rangle$ will be a list of μ-terms for every word-product $\boldsymbol{P}$ over S^{t}, so that $\langle a^?(\boldsymbol{P}) \rangle = a\langle \boldsymbol{P} \rangle$ will be the application of the function symbol a to the list of arguments $\langle \boldsymbol{P} \rangle$. We use only *one* μ-bound variable, which we call $\Box$: this serves for tree iterators, which are translated as iterators of the form $\mu\,\Box = s \cdot t$ (third row below).

$$\langle a^{[?]}(\boldsymbol{P}) \rangle = a\langle \boldsymbol{P} \rangle \qquad \langle e_1 e_2 \dots e_m \rangle = (\langle e_1 \rangle, \langle e_2 \rangle, \dots, \langle e_m \rangle) \qquad \langle \Box \rangle = \Box$$
$$\langle P^? \rangle = \langle P \rangle \qquad \langle A^* \rangle = \langle A \rangle = \mathtt{u}(\langle P_1 \rangle, \dots, \langle P_n \rangle) \text{ where } A = \{\!|P_1, \dots, P_n|\!\}$$
$$\langle (\textstyle\bigcup_{i=1}^{m} a_i(\boldsymbol{Q}_i))^{[*]}.\Box A \rangle \quad = \quad \mu\,\Box = \mathtt{u}(\langle a_1^{[?]}(\boldsymbol{Q}_1) \rangle, \dots, \langle a_m^{[?]}(\boldsymbol{Q}_m) \rangle) \cdot \langle A \rangle$$

Define the precedence $<$ by $a < b$ iff $a, b \in S$ and $a \sqsubset b$, or $a = \mathtt{u}$ and $b \in S$ ($\mathtt{u}$ is least). We check that $P \sqsubset^{\mathrm{t}} P'$ implies $\langle P \rangle <^{\mathrm{mpo}} \langle P' \rangle$. (This was how $<^{\mathrm{mpo}}$ was found on μ-terms!) Theorem 1 then implies that $\sqsubset^{\mathrm{t}}$ is well-founded on S^{t}. $\square$

Corollary 2 (Topological Kruskal Theorem, Constructively). *For every constructive Noetherian space $(X, T, \prec, \varepsilon)$, $(\mathcal{T}(X), \mathcal{M}(T^{\mathrm{t}}), (\prec^{\mathrm{t}}_{mul})^+, \varepsilon^{\mathrm{t}})$ is a constructive Noetherian space, with $t\ \varepsilon^{\mathrm{t}}\ M$ iff $\{\!|\eta^{\mathrm{t}}(t)|\!\}\ (\prec^{\mathrm{t}}_{mul})^*\ M$, where $\eta^{\mathrm{t}}(f(t_1, t_2, \ldots, t_n)) = f^{[?]}(\eta^{\mathrm{t}}(t_1)^? \eta^{\mathrm{t}}(t_2)^? \ldots \eta^{\mathrm{t}}(t_n)^?)$.*

As for Higman's Lemma, we obtain the ordinary form of Kruskal's Theorem by assuming a decidable constructive wqo $\leqslant$ on X, and proving that the complement $\complement t$ of the upward closure of a single tree t in $\preceq_\leqslant$ is defined as the following tree-product, built using tree iterators only: letting a abbreviate $X \smallsetminus \uparrow f$ and b abbreviate X itself, $\complement f() = (a(\Box^*))^{[*]}.{}_\Box\varnothing$, and $\complement f(t_1, \ldots, t_n)$ for $n \geqslant 1$ is equal to $(a(\Box^*) \cup b(\complement t_1 \Box^? \complement t_2 \Box^? \ldots \Box^? \complement t_n))^{[*]}.{}_\Box\varnothing$ (see [22, Lemma 9.8]; then use canonicalization). The following is then constructive.

Theorem 5 (Kruskal, Constructively). *Let X be a set with a decidable constructive wqo $\leqslant$. Then $\preceq_\leqslant$ is a (decidable) constructive wqo on $\mathcal{T}(X)$.*

6 Conclusion

The main thing one should remember is that proving that a given quasi-ordering is well is just a matter of proving termination—not of the ordering itself, but of strict inclusion between downward-closed subsets. In and of itself, this would be no breakthrough. However, in applying this to Higman's and Kruskal's classical theorems, this exposed a tight coupling between the word embedding ordering and the multiset ordering (on word-products), and between the tree embedding quasi-ordering and Dershowitz' multiset path ordering (on tree-products).

While I have given relatively exhaustive proofs of the termination results of Section 3, I have barely sketched the constructive proofs of the (topological) Higman and Kruskal theorems in Sections 4 and 5. Playing these proofs in a proof assistant such as Coq is in order, but certainly somewhat of an endeavor.

Finally, I would like to stress that although our proof techniques establish the classical, order-theoretic versions of Higman's and Kruskal's theorems under a decidability assumption, the topological versions are *entirely* constructive. It therefore seems that the constructive contents of the order-theoretic and the topological theorems are different—something that should be explored, by investigating into the computational contents of the relevant constructive proofs. We also believe that the notion of constructive Noetherian space, and the related notion of S-representation, should be of some importance, in intuitionistic logic (where it sheds some light on the precise role of the sequential regular expressions of Murthy and Russell, notably), as well as in the field of WSTS model-checking.

References

1. P. A. Abdulla, K. Čerāns, B. Jonsson, and Y.-K. Tsay. Algorithmic analysis of programs with well quasi-ordered domains. *Inf. Comput.*, 160(1-2):109–127, 2000.

2. P. A. Abdulla, A. Collomb-Annichini, A. Bouajjani, and B. Jonsson. Using forward reachability analysis for verification of lossy channel systems. *Formal Methods in System Design*, 25(1):39–65, 2004.
3. P. A. Abdulla and B. Jonsson. Verifying programs with unreliable channels. In *LICS'93*, pages 160–170. IEEE Computer Society Press, 1993.
4. L. Acciai and M. Boreale. Deciding safety properties in infinite-state pi-calculus via behavioural types. In *ICALP'09*, pages 31–42. Springer Verlag LNCS 5556, 2009.
5. S. Berghofer. A constructive proof of Higman's lemma in Isabelle. In *TYPES'03*, pages 66–82. Springer Verlag LNCS 3085, 2004.
6. Y. Bertot and P. Castéran. *Interactive Theorem Proving and Program Development—Coq'Art: The Calculus of Inductive Constructions*, volume XXV of *Texts in Theor. Comp. Sci., an EATCS Series*. Springer Verlag, 2004.
7. T. Coquand and D. Fridlender. A proof of Higman's lemma by structural induction. `http://www.cse.chalmers.se/~coquand/open1.ps`. Unpublished note.
8. S. Coupet-Grimal and W. Delobel. An effective proof of the well-foundedness of the multiset path ordering. *Appl. Algebra Eng., Commun. Comput.*, 17(6):453–469, 2006.
9. J. Dawson and R. Goré. A general theorem on termination of rewriting. In *CSL'04*, pages 100–114. Springer Verlag LNCS 3210, 2004.
10. P. de Groote, B. Guillaume, and S. Salvati. Vector addition tree automata. In *LICS'04*, pages 64–73. IEEE Computer Society Press, 2004.
11. N. Dershowitz. A note on simplification orderings. *Inf. Proc. Letters*, 9(5):212–215, 1979.
12. N. Dershowitz. Orderings for term-rewriting systems. *Theor. Comp. Sci.*, 17(3):279–301, Mar. 1982.
13. N. Dershowitz. Termination of rewriting. *J. Symb. Comp.*, 3:69–116, 1987.
14. N. Dershowitz. Jumping and escaping: Modular termination and the abstract path ordering. *Theor. Comp. Sci.*, 464:35–47, 2012.
15. N. Dershowitz and C. Hoot. Natural termination. *Theor. Comp. Sci.*, 142(2):179–207, 1995.
16. N. Dershowitz and J.-P. Jouannaud. Rewrite systems. In J. van Leeuwen, editor, *Handbook of Theor. Comp. Sci.*, chapter 6, pages 243–320. Elsevier, 1990.
17. N. Dershowitz and Z. Manna. Proving termination with multiset orderings. *Comm. ACM*, 22(8):465–476, 1979.
18. L. E. Dickson. Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. *Amer. J. Math.*, 35(4):413–422, 1913.
19. M. C. F. Ferreira and H. Zantema. Well-foundedness of term orderings. In *CTRS'94*, pages 106–123. Springer Verlag LNCS 968, 1995.
20. A. Finkel and J. Goubault-Larrecq. Forward analysis for WSTS, part I: Completions. In *STACS'09*, pages 433–444, Freiburg, Germany, 2009. Leibniz-Zentrum für Informatik, Intl. Proc. in Informatics 3.
21. A. Finkel and J. Goubault-Larrecq. Forward analysis for WSTS, part II: Complete WSTS. In *ICALP'09*, pages 188–199. Springer LNCS 5556, 2009.
22. A. Finkel and J. Goubault-Larrecq. Forward analysis for WSTS, part I: Completions. *Logical Methods in Computer Science*, 2012. Submitted.
23. A. Finkel and J. Goubault-Larrecq. Forward analysis for WSTS, part II: Complete WSTS. *Logical Methods in Computer Science*, 8(3:28), 2012.
24. A. Finkel, P. McKenzie, and C. Picaronny. A well-structured framework for analysing Petri net extensions. *Inf. Comput.*, 195(1-2):1–29, 2004.

25. A. Finkel and P. Schnoebelen. Well-structured transition systems everywhere! *Theor. Comp. Sci.*, 256(1–2):63–92, 2001.
26. J. Goubault-Larrecq. Well-founded recursive relations. In L. Fribourg, editor, *CSL'01*, pages 484–497. Springer LNCS 2142, 2001.
27. J. Goubault-Larrecq. On Noetherian spaces. In *LICS'07*, pages 453–462. IEEE Computer Society Press, 2007.
28. J. Goubault-Larrecq. Noetherian spaces in verification. In *ICALP'10*, pages 2–21. Springer Verlag LNCS 6199, 2010.
29. J. Goubault-Larrecq. *Non-Hausdorff Topology and Domain Theory—Selected Topics in Point-Set Topology*, volume 22 of *New Mathematical Monographs*. Cambridge University Press, 2013.
30. G. Higman. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc.*, 2(7):326–336, 1952.
31. S. Kamin and J.-J. Lévy. Attempts for generalizing the recursive path orderings. Unpublished letter to N. Dershowitz, 1980. `http://nachum.org/term/kamin-levy80spo.pdf`.
32. R. M. Karp and R. E. Miller. Parallel program schemata. *J. Comp. Sys. Sci.*, 3(2):147–195, 1969.
33. J. B. Kruskal. Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Trans. AMS*, 95(2):210–225, 1960.
34. R. Lazič, T. Newcomb, J. Ouaknine, A. W. Roscoe, and J. Worrell. Nets with tokens which carry data. *Fund. Informaticae*, 88(3):251–274, 2008.
35. P. Lescanne. Some properties of decomposition ordering, a simplification ordering to prove termination of rewriting systems. *RAIRO Theor. Inform.*, 16(4):331–347, 1982.
36. R. Meyer. On boundedness in depth in the π-calculus. In *Proc. 5th IFIP Intl. Conf. Theor. Comp. Sci.*, volume 273 of *IFIP*. Springer-Verlag, 2008.
37. C. R. Murthy and J. R. Russell. A constructive proof of Higman's lemma. In *LICS'90*, pages 257–267. IEEE Computer Society Press, 1990.
38. C. S.-J. A. Nash-Williams. On well-quasi-ordering infinite trees. *Proc. Cambridge Phil. Soc.*, 61:697–720, 1965.
39. T. Nipkow. An inductive proof of the wellfoundedness of the multiset order. Technical report, Technische Universität München, Oct. 1998.
40. M. Ogawa. A linear time algorithm for monadic querying of indefinite data over linearly ordered domains. *Inf. Comput.*, 186(2):236–259, 2003.
41. F. Richman and G. Stolzenberg. Well quasi-ordered sets. Technical report, Northeastern University, Boston, MA and Harvard University, Cambridge, MA, 1990.
42. F. Rosa-Velardo and M. Martos-Salgado. Multiset rewriting for the verification of depth-bounded processes with name binding. *Inf. Comput.*, 215:68–87, June 2012.
43. M. Seisenberger. Kruskal's tree theorem in a constructive theory of inductive definitions. In *Proc. Symp. Reuniting the Antipodes—Constructive and Nonstandard Views of the Continuum*. Synthese Library 306, Kluwer Academic Publishers, Dordrecht, 2001, 1999.
44. R. van der Meyden. The complexity of querying indefinite data about linearly ordered domains. *J. Comp. Sys. Sci.*, 54(1):113–135, 1997.
45. W. Veldman. An intuitionistic proof of Kruskal's theorem. Report 0017, Dept. of Mathematics, U. Nijmegen, 2000.
46. K. N. Verma and J. Goubault-Larrecq. Karp-Miller trees for a branching extension of VASS. *Discr. Math. & Theor. Comp. Sci*, 7(1):217–230, 2005.
47. T. Wies, D. Zufferey, and T. A. Henzinger. Forward analysis of depth-bounded processes. In *FoSSaCS'10*, pages 94–108. Springer Verlag LNCS 6014, 2010.