A Formal Proof of Kruskal's Tree Theorem in Isabelle/HOL

Christian Sternagel*

JAIST, Japan
c-sterna@jaist.ac.jp

Abstract. We give the first formalization of Kruskal's tree theorem in a proof assistant – the closure of a long-standing challenge. More concretely, we present our Isabelle/HOL development of Nash-Williams' minimal bad sequence argument for proving the tree theorem. Along the way, we discuss proofs of Dickson's lemma and Higman's lemma; and close all gaps in the original proofs.

Keywords: Well-Quasi-Orders, Dickson's Lemma, Minimal Bad Sequences, Higman's Lemma, Kruskal's Tree Theorem

1 Introduction

Kruskal's tree theorem [1] (sometimes called *the* tree theorem in the following) is a famous result in combinatorics, more precisely well-quasi-order (wqo) theory. Neglecting details, it states that:

When a set A is wqo'd, then so is the set of finite trees over A. (†)

In [2], Nash-Williams gave a short and elegant (according to several authors citing his work) proof of the tree theorem, where he first established what is now known as the *minimal bad sequence argument*: assume the existence of a minimal "bad" infinite sequence of elements, construct an even smaller "bad" infinite sequence, thus contradicting minimality and proving wqo'dness (since the definition of wqo requires all infinite sequences of elements to be "good").

Besides the minimal bad sequence argument, [2] contains proofs of Dickson's lemma [3] (*if A and B are wqo'd, then so is the Cartesian product $A \times B$*) and a variant of Higman's lemma [4] (*if A is wqo'd, then so is the set of finite subsets of A*), where the latter also incorporates an instance of the minimal bad sequence argument. We will see later that some more facts are used implicitly.

We formalized Nash-Williams' proofs in the proof assistant Isabelle [5].[1] Mostly we agree with other authors that his argumentation is short (in fact,

* Supported by the Austrian Science Fund (FWF): J3202.

[1] Available from http://isabelle.in.tum.de/website-Isabelle2013-RC2 (we recommend Isabelle/jEdit for browsing).

[2] consists of only two and a half pages in total) and elegant, which was also the main reason why we chose his work for our formalization in the first place. However, formalizations (using proof assistants) typically require us to be more rigorous than with pen and paper. Thus, we had to close some gaps, which results in a somewhat longer (about three and a half thousand lines of Isabelle/HOL theories) and slightly less elegant proof. Fortunately the biggest of those gaps could be localized (pun intended; see also [6]), thus not derogating the elegance of the remaining proof.

We stress that everything we present in the following, was formalized using the proof assistant Isabelle. Here, we give a high-level overview of our formalization. The full development is part of the *Archive of Formal Proofs* [7].

Contributions. To the best of our knowledge, our work constitutes the first unrestricted (though non-constructive) formalization of Higman's lemma in Isabelle/HOL as well as the first formalization of Kruskal's tree theorem ever.

We think that formalizations are valuable since they (have to) contain *all* non-trivial (where the definition of "trivial" is directly related to the power of the used proof assistant) steps of a proof. No doubt, more often than not, those steps where already conducted in the minds of the original proof authors (but then again, maybe not). The point is that when the original author writes down his proof in a condensed form for publishing, some of the steps may get lost (be it that they are considered trivial or that explaining them would be too much trouble). If, much later, another person tries to understand the proof, there may be some mental gaps (or in the worst case even errors). Thus, formalizations are of archival and educational value – in addition to the more obvious fact that they are highly trustworthy. On the other hand, formalizations are often hard to read for non-experts (note that the Isar language for Isabelle was a huge improvement in that respect). Thus we hope that this high-level overview makes our formalization more accessible.

Differences to [2]. Our formalization of Nash-Williams' arguments differs from [2] in several details: As stated above, Nash-Williams proved a variant of Higman's lemma. The actual statement of Higman's lemma is usually along the lines: *If A is wqo'd, then so is A^*, the set of finite words (or lists) over A*; which is also the version that we formalized. Why do both of these variants lead to the tree theorem? In fact, it depends on the exact formulation of the tree theorem whether they do. From (†), we just know that there *exists* some wqo on the set of finite trees over A. Typically, we are interested in a specific order, namely homeomorphic embedding. However, this is not what Nash-Williams used (neither for his variant of Higman's lemma, nor in his proof of the tree theorem).

In the following, whenever we refer to Higman's lemma, we mean "*If A is wqo'd, then A^* is wqo'd by homeomorphic embedding on lists*," and when we refer to the tree theorem we mean "*If A is wqo'd, then the set of finite trees over A is wqo'd by homeomorphic embedding on trees.*"

Overview. The remainder is structured as follows. In Section 2, we cover preliminaries. Then, in Section 3, we review the structure of Nash-Williams' original proofs from [2]. The next four sections present our formalization of Dickson's lemma (featuring a proof of a variant of Dickson's lemma for almost-full relations, i.e., not relying on transitivity), in Section 4; our general construction of minimal bad sequences, in Section 5; our formalization of Higman's lemma, in Section 6; and ultimately, our formalization of Kruskal's tree theorem, in Section 7. Finally, we conclude in Section 8, where we also sketch applications, discuss future work, and refer to related work.

2 Preliminaries

Throughout our exposition, we use standard mathematical notation as far as possible. However, we also employ some Isabelle specific notation, since we used Isabelle's document preparation facilities for typesetting all lemmas and theorems (in the words of [8]: *no typos, no omissions, no sweat*; alas, this does not extend to the regular text). Thus, some explanation might be in order.

Isabelle/HOL, is a higher-order logic based on the simply-typed lambda calculus. Thus, every term has a type, where we use Greek letters α, β, γ, ... for *type variables*, and *type constructors* like *nat* for natural numbers, $\alpha \Rightarrow \beta$ for the function space, $\alpha \times \beta$ for ordered pairs, α *set* for sets, and α *list* for finite lists. *Type constraints* are written $t{::}\tau$ and denote that term t is of type τ. As usual for lambda calculi, function application is denoted by juxtaposition, i.e., $f\ x$ applies the function f to the argument x. We use the type $\alpha \Rightarrow \alpha \Rightarrow \mathit{bool}$ to encode binary relations. (An alternative would have been to use $(\alpha \times \alpha)$ *set*, however, the two representations are mostly equivalent and the former is used for many binary relations of Isabelle/HOL's standard library.)

We freely use the following constants from Isabelle/HOL's standard library: $\circ{::}(\alpha \Rightarrow \beta) \Rightarrow (\gamma \Rightarrow \alpha) \Rightarrow \gamma \Rightarrow \beta$, where $f \circ g$ denotes the functional composition of the two functions f and g, i.e., $f \circ g \stackrel{\text{def}}{=} \lambda x.\ f\ (g\ x)$; $\mathit{fst}{::}\alpha \times \beta \Rightarrow \alpha$ and $\mathit{snd}{::}\alpha \times \beta \Rightarrow \beta$ extract the first and second component of a pair, respectively; $\mathit{set}{::}\alpha\ \mathit{list} \Rightarrow \alpha\ \mathit{set}$, where *set xs* is the set of elements occurring in the list *xs*; $[]{::}\alpha\ \mathit{list}$, the empty list; $\cdot{::}\alpha \Rightarrow \alpha\ \mathit{list} \Rightarrow \alpha\ \mathit{list}$, where $x{\cdot}xs$ denotes "consing" the element x in front of the list *xs*; and $@{::}\alpha\ \mathit{list} \Rightarrow \alpha\ \mathit{list} \Rightarrow \alpha\ \mathit{list}$, where *xs* @ *ys* denotes the concatenation of the two lists *xs* and *ys*. Note that since $\cdot$ and @ are both right-associative and have the same priority, *xs* @ $y{\cdot}ys$ is the same as *xs* @ $(y{\cdot}ys)$ and denotes a list that is constructed by inserting the element y between the elements of *xs* and *ys*.

Moreover, we sometimes use Isabelle specific notation when stating formulas. Then, $\bigwedge$ is universal quantification, $\Longrightarrow$ is (right-associative) implication, and nested implication, like $A \Longrightarrow B \Longrightarrow C$, is abbreviated to $[\![A;\ B]\!] \Longrightarrow C$.

Let $\preceq$ be a binary relation and A a set. We say that $\preceq$ is *reflexive (on A)*, written $\mathit{refl}_A(\preceq)$, iff $\forall\, x{\in}A.\ x \preceq x$; and *transitive (on A)*, written $\mathit{trans}_A(\preceq)$, iff $\forall\, x{\in}A.\ \forall\, y{\in}A.\ \forall\, z{\in}A.\ x \preceq y \wedge y \preceq z \longrightarrow x \preceq z$.

Infinite sequences over elements of type α are represented by functions of type $nat \Rightarrow \alpha$. A binary relation $\preceq$ is transitive on a sequence f, written $trans_f(\preceq)$, iff $\forall\, i\ j.\ i < j \longrightarrow f\ i \preceq f\ j$. A sequence f is *good* (w.r.t. $\preceq$), written $good_{\preceq}(f)$, iff $\exists\, i\ j.\ i < j \wedge f\ i \preceq f\ j$. If a sequence is not good, it is called *bad*.

We follow Veldman [9] and Vytiniotis et al. [10] in basing wqos on *almost-full* relations (which are basically wqos without transitivity). The main reason for doing so, is that all the properties we are interested in also hold for almost-full relations and are easily extended to wqos.

The relation $\preceq$ is *almost-full* (on A), written $af_A(\preceq)$, iff all infinite sequences over elements of A are good, i.e., $af_A(\preceq) \stackrel{\text{def}}{=} \forall f.\ (\forall\, i.\ f\ i \in A) \longrightarrow good_{\preceq}(f)$. Note that every almost-full relation is necessarily reflexive (see, e.g., [6, Lemma 1]).

Let $\preceq$ be almost-full on A. If in addition $\preceq$ is *transitive* on A, then $\preceq$ is a wqo on A (equivalently, A is wqo'd by $\preceq$), written $wqo_A(\preceq)$.

3 Nash-Williams' Proof

Before we give a detailed account of our formalization (in the sections to come), let us review the structure of Nash-Williams' original proofs. This allows us to familiarize the reader with the overall structure, highlight differences to our approach, and indicate gaps. Since we do not reproduce the full proofs, a copy of [2] might be useful for reference.

Nash-Williams starts by giving a proof of Dickson's lemma: *If A and B are wqo'd, then so is $A \times B$* [2, Lemma 1]. The proof is presented rather sketchy, which makes it difficult to see what additional facts are actually required. In essence, we have two infinite sequences a (over elements of A) and b (over elements of B) that are both known to be good, and have to construct witnesses i and j such that $i < j$ and $(a\ i,\ b\ i) \preceq (a\ j,\ b\ j)$. It seems that the proof requires the axiom of choice[2] as well as transitivity of $\preceq$ (none of which is mentioned explicitly). In contrast, in our formalization we prove Dickson's lemma for almost-full relations based on an existing formalization of Ramsey's theorem (whose proof employs the axiom of dependent choice) and do not require transitivity of $\preceq$.

Next comes a proof of Higman's lemma: *If A is wqo'd, then A^* is wqo'd by homeomorphic embedding on lists* [2, Lemma 2]. Assume that the statement is false. Then construct a minimal bad sequence, i.e., a bad sequence such that replacing any given element by a smaller one, the resulting sequence would be good. The first gap is: we are to construct a minimal bad sequence, but

it is not shown how to construct a minimal bad sequence (G1)

or proved that such a sequence even exists. More specifically, the only thing Nash-Williams has to say about the construction of a minimal bad sequence is roughly (where we use A^{o} to denote the set of "objects" built over elements of A; which might refer to the set of finite subsets, the set of finite lists, the set of finite trees, ... in a concrete case):

[2] In Isabelle: $\forall x.\ \exists y.\ Q\ x\ y \Longrightarrow \exists f.\ \forall x.\ Q\ x\ (f\ x)$.

> *Select an $x_1 \in A^\circ$ such that x_1 is the first term of a bad sequence of members of A° and $|x_1|$ is as small as possible. Then select an x_2 such that x_1, x_2 (in that order) are the first two terms of a bad sequence of members of A° and $|x_2|$ is as small as possible [...]. Assuming the axiom of choice, this process yields a [minimal] bad sequence [...]*

This is not enough, since it is neither shown that the sketched process is well-defined, nor that it would indeed result in a minimal bad sequence. Interestingly, most non-formalized proofs of Higman's lemma and Kruskal's tree theorem in the literature (that the author is aware of) are similarly vague about the actual construction of a minimal bad sequence.

In our formalization, we found the construction of a minimal bad sequence (and the accompanying proof that the constructed sequence is indeed minimal and bad) to be the hardest and most technical part. For now, assume that we have a minimal bad sequence m (which is a sequence of finite lists). Let h be the sequence of heads of m and t the sequence of corresponding tails. It is then shown that there is no sequence of some special shape over $T = \bigcup_i \{t\ i\}$ (otherwise m would be good). Furthermore, it is claimed (without proof) that

> *a bad sequence over T indicates a sequence of this special shape* (G2)

which we consider the second gap. Thus T is wqo'd, since there are no bad sequences. Let $H = \bigcup_i \{h\ i\}$, which is wqo'd since A is. Then, by Dickson's lemma, $H \times T$ is wqo'd. Hence, there are i and j such that $i < j$ and $(h\ i,\ t\ i) \preceq (h\ j,\ t\ j)$, which implies $m\ i \preceq m\ j$ and thus contradicts the badness of m.

Finally, for the tree theorem the proof structure is very similar to the previous one (only using finite trees instead of finite lists and homeomorphic embedding on trees instead of homeomorphic embedding on lists). Assume that the statement is false. Again we have to construct a minimal bad sequence m (and thus we again have the same gap as for Higman's lemma). Instead of heads and tails of lists, we have roots and direct subtrees (which we also call successors) of trees. Let r and s denote the sequences of roots and successors of m. Let S' be the set of successors occurring in s, i.e., $S' = \{t \mid \exists\, i.\ t \in \mathit{set}\ (s\ i)\}$. This time it is claimed (again without proof) that

> *a bad sequence over S' indicates a sequence of a special shape* (G3)

(another special shape than above). This we consider the third gap.

4 Dickson's Lemma

Essentially, our formalization is about preservation of wqo'dness by certain set/type constructors (Dickson's lemma for pairs, Higman's lemma for lists, and the tree theorem for trees). For each of these set/type constructors, we need a construction that extends the order(s) on the base set/type(s) to an order on the newly constructed set/type. For Dickson's lemma we use the following construction: Given two orders $\preceq_1$ and $\preceq_2$, the pointwise order on pairs is defined by $(a_1,\ a_2) \preceq (b_1,\ b_2) \stackrel{\text{def}}{=} a_1 \preceq_1 b_1 \wedge a_2 \preceq_2 b_2$.

The following lemma shows that the pointwise combination of orders preserves wqo'dness, when forming the Cartesian product of two sets.

Lemma 1. $[\![af_{A_1}(\preceq_1);\ af_{A_2}(\preceq_2)]\!] \Longrightarrow af_{A_1 \times A_2}(\preceq)$

Before we consider the proof of Lemma 1, let us have a look at how Ramsey's theorem allows us to disregard transitivity (and hence prove the lemma already for almost-full relations rather than wqos).

We use the following variant of Ramsey's theorem (which was formalized as part of Isabelle/HOL's library; `~~/src/HOL/Library/Ramsey.thy`):

$$[\![\mathit{infinite}\ Z;\ \forall i{\in}Z.\ \forall j{\in}Z.\ i \neq j \longrightarrow h\ \{i, j\} < n]\!]$$
$$\Longrightarrow \exists I\ c.\ I \subseteq Z \wedge \mathit{infinite}\ I \wedge c < n \wedge (\forall i{\in}I.\ \forall j{\in}I.\ i \neq j \longrightarrow h\ \{i, j\} = c)$$

In words: Let Z be an infinite set and let h be a function that, given a two-element subset of Z, returns a natural number smaller than n. Then there is an infinite subset I of Z and a natural number c smaller than n such that h encodes all two-element subsets of I by c. More abstractly, assume we have an infinite graph with nodes from Z such that every edge has exactly one of n colors. Then there is an infinite subgraph with nodes from I and edges of color c.

Using Ramsey's theorem, we prove the auxiliary fact that whenever the union of two binary relations is transitive on an infinite sequence, then there is an infinite subsequence on which either the first or the second relation is transitive.

Lemma 2. $\mathit{trans}_f(\preceq_1 \cup \preceq_2) \Longrightarrow \exists \varphi.\ \mathit{trans}_\varphi(<) \wedge (\mathit{trans}_{f\varphi}(\preceq_1) \vee \mathit{trans}_{f\varphi}(\preceq_2))$

Here φ is a strictly monotone (since $<$ is transitive on it) mapping from natural numbers to natural numbers. Hence, $f \circ \varphi$ (which we write $f\varphi$, for brevity) is a subsequence of f whose elements are in the same relative order.

Proof (of Lemma 2). Assume $\mathit{trans}_f(\preceq_1 \cup \preceq_2)$, which means that

$$\text{for all } i < j, \text{ either } f\ i \preceq_1 f\ j \text{ or } f\ i \preceq_2 f\ j. \qquad (\star)$$

We colorize the set of two-element subsets $\{i, j\}$ of the natural numbers using h, defined by, if $i < j$ and $f\ i \preceq_1 f\ j$, then $h\ \{i, j\}$ is *0* (*white*), otherwise *1* (*black*). Now we can apply Ramsey's theorem from above (since the set of natural numbers is infinite and there are exactly two colors). Thus, we obtain an infinite set I of natural numbers and a color c such that for all $i \neq j$ in I, the corresponding color $h\ \{i, j\}$ is c. Since I is countably infinite, there is some function φ::*nat* $\Rightarrow$ *nat* that enumerates its elements in increasing order, i.e., $\mathit{trans}_\varphi(<)$. We consider two cases (for arbitrary but fixed $i < j$):

- **case** (c is *white*). Since φ is strictly monotone, we have $\varphi\ i < \varphi\ j$. Therefore $h\ \{\varphi\ i, \varphi\ j\} = 0$, and thus $f_\varphi\ i \preceq_1 f_\varphi\ j$.
- **case** (c is *black*). Again, we have $\varphi\ i < \varphi\ j$. Thus $h\ \{\varphi\ i, \varphi\ j\} = 1$ which, together with ($\star$), implies $f_\varphi\ i \preceq_2 f_\varphi\ j$. □

Proof (of Lemma 1). Assume $af_{A_1}(\preceq_1)$ and $af_{A_2}(\preceq_2)$. Moreover, to derive a contradiction, assume $\neg\, af_{A_1 \times A_2}(\preceq)$. Then there is some sequence f on $A_1 \times A_2$ which is bad. Let $x \lhd y$ and $x \blacktriangleleft y$ denote $fst\ x \not\preceq_1 fst\ y$ and $snd\ x \not\preceq_2 snd\ y$, respectively. Since f is bad, we have $\forall\, i\ j.\ i < j \longrightarrow f\ i \lhd f\ j \vee f\ i \blacktriangleleft f\ j$, i.e., $trans_f(\lhd \cup \blacktriangleleft)$. Then, by Lemma 2, we obtain a strictly monotone mapping φ such that $trans_{f\,\varphi}(\lhd)$ or $trans_{f\,\varphi}(\blacktriangleleft)$. In the first case $fst \circ f_\varphi$ is bad and in the second $snd \circ f_\varphi$ is bad, both contradicting our assumptions. □

Dickson's Lemma. $[\![wqo_{A_1}(\preceq_1);\ wqo_{A_2}(\preceq_2)]\!] \Longrightarrow wqo_{A_1 \times A_2}(\preceq)$

Proof. Assuming transitivity of $\preceq_1$ on A_1 and $\preceq_2$ on A_2, it is trivial to show transitivity of $\preceq$ on $A_1 \times A_2$. With Lemma 1, we obtain Dickson's lemma. □

5 Minimal Bad Sequences

Since the minimal bad sequence argument is needed for Higman's lemma as well as the tree theorem, we aim for a general construction that is applicable to both cases. To this end, we employ Isabelle/HOL's locale mechanism which allows us to define new constants and prove facts using an "interface" of hypothetical constants and assumptions. As long as the assumptions can be discharged, the new constants and proven facts can be instantiated to arbitrary special cases.

Below, we describe the locale *mbs* which captures the construction of a minimal bad sequence over elements from a given set (which we call *objects*). Such objects are built from elements of some other set (e.g., the elements of a list, or the nodes of a tree). The locale fixes the following constants:

- A function $\_^{\mathsf{o}}$, where A^{o} returns the set of all objects over elements of A,
- a relation $\preceq_A$ that is used to check whether an infinite sequence of objects is good (where A restricts the set of objects that may be compared),
- and a relation $\lhd$ that is used to compare the structural size of two objects.

Furthermore, it has the assumptions:

$$wf_{A^{\mathsf{o}}}(\lhd) \tag{M1}$$

$$[\![x \lhd y;\ y \lhd z]\!] \Longrightarrow x \lhd z \tag{M2}$$

$$[\![x \lhd y;\ y \in A^{\mathsf{o}}]\!] \Longrightarrow x \in A^{\mathsf{o}} \tag{M3}$$

$$[\![z \in A^{\mathsf{o}};\ x \preceq_A y;\ y \lhd z]\!] \Longrightarrow x \preceq_A z \tag{M4}$$

That is, the structural comparison is well-founded on objects from A^{o} (M1) (thus, it makes sense to talk about a *minimal* object), transitive (M2), and preserves the property of being in A^{o} (M3). Moreover, it is *(right-)compatible* with $\preceq_A$ (M4). It turns out that these ingredients are enough to construct – under the assumption that there is a bad sequence – a minimal bad sequence. Well, to be precise, we first need to say explicitly what we mean by a *minimal* bad sequence. Informally, we mean of course that replacing any single object by

some smaller one, the result will no longer be bad. More formally, we say that an infinite sequence f is minimal at a position n, written $min^A_n(f)$, iff

$$\forall\, g.\ (\forall\, i{<}n.\ g\ i = f\ i) \wedge g\ n \lhd f\ n \wedge (\forall\, i{\geq}n.\ \exists\, j{\geq}n.\ g\ i \trianglelefteq f\ j) \longrightarrow good_{\preceq_A}(g)$$

In words: for every sequence g such that its initial part up to (but not including) position n coincides with f, the n-th object of g is strictly smaller than the n-th object of f, and every object of g after position n is smaller than or equal to some object of f at a later or equal position; g is good. This definition facilitates the construction of an overall minimal bad sequence from some given bad sequence by iterating over its positions: objects before the current position will never change, at the current position we insert the smallest possible object, and objects at later positions are required to be built from elements that have already been present in f (we will point out the place in our proofs where we require the last assumption, since it may not seem very natural).

As indicated above, we want to modify a given sequence iteratively. We will do so by splicing it with another sequence at a certain position. The way we splice two sequences at a position n is as follows: given the sequences f and g, we have $f\langle n\rangle g \stackrel{\text{def}}{=} \lambda j.\ \mathit{if}\ n \leq j\ \mathit{then}\ g\ j\ \mathit{else}\ f\ j$. Thus, all objects before position n are taken from f and all others from corresponding positions of g. This operation can be used to splice two bad sequences into a new one by the following lemma (which is the origin of the third assumption in the definition of *min*).

Lemma 3. $[\![\forall\, i.\ f\ i \in A^{\mathsf{o}};\ bad_{\preceq_A}(f);\ bad_{\preceq_A}(g);\ \forall\, i{\geq}n.\ \exists\, j{\geq}n.\ g\ i \trianglelefteq f\ j]\!] \Longrightarrow bad_{\preceq_A}(f\langle n\rangle g)$

Proof. Assume that $f\,i \in A^{\mathsf{o}}$ for all i, f and g are bad, and $\forall\, i{\geq}n.\ \exists\, j{\geq}n.\ g\ i \trianglelefteq f\ j$. Moreover, for the sake of a contradiction, assume that $f\langle n\rangle g$ is good. Thus, there are $i < j$ such that $(f\langle n\rangle g)\ i \preceq_A (f\langle n\rangle g)\ j$. We analyze three cases:

- **case** $(j < n)$. Hence $f\ i \preceq_A f\ j$, contradicting the badness of f.
- **case** $(n \leq i)$. Hence $g\ i \preceq_A g\ j$, contradicting the badness of g.
- **case** $(i < n$ and $n \leq j)$. Then there is some $k \geq n$ such that $g\ j \trianglelefteq f\ k$. Moreover, $f\ i \preceq_A g\ j$. By compatibility, we obtain $f\,i \preceq_A f\,k$ (since $f\ k \in A^{\mathsf{o}}$), contradicting the badness of f. □

For our iterative construction of a minimal bad sequence, we employ the following auxiliary lemma (which shows that from a sequence that is minimal at position n, we may obtain a sequence that is also minimal at the next position $n{+}1$):

Lemma 4. $[\![\forall\, i.\ f\ i \in A^{\mathsf{o}};\ min^A_n(f);\ bad_{\preceq_A}(f)]\!]$
$\Longrightarrow \exists\, g.\ (\forall\, i{\leq}n.\ g\ i = f\ i) \wedge$
$g\ (n{+}1) \trianglelefteq f\ (n{+}1) \wedge$
$(\forall\, i{\geq}n{+}1.\ \exists\, j{\geq}n{+}1.\ g\ i \trianglelefteq f\ j) \wedge$
$bad_{\preceq_A}(f\langle n{+}1\rangle g) \wedge min^A_{n+1}(f\langle n{+}1\rangle g)$

Proof. Since $\lhd$ is well-founded on A^{o}, we have the induction schema

$$[\![x \in A^{\mathsf{o}};\ \textstyle\bigwedge x.\ [\![x \in A^{\mathsf{o}};\ \bigwedge y.\ [\![y \in A^{\mathsf{o}};\ y \lhd x]\!] \Longrightarrow P\ y]\!] \Longrightarrow P\ x]\!] \Longrightarrow P\ x$$

Assume $f\ (n{+}1) \in A^{\mathsf{o}}$, $min^A_n(f)$, and $bad_{\preceq_A}(f)$. Let $\exists\, g.\ \mathfrak{C}\ g\ f\ (f\ (n{+}1))$ abbreviate the conclusion of Lemma 4 (parametrized over the sequences g and f and the element on which we apply induction). In order for the (later) induction to go through, we prove a slightly stronger statement than Lemma 4. To this end, let $\mathfrak{I}\ x$ abbreviate

$$\forall f.\ x = f\ (n{+}1) \wedge (\forall\, i.\ f\ i \in A^{\mathsf{o}}) \wedge min^A_n(f) \wedge bad_{\preceq_A}(f) \longrightarrow (\exists\, g.\ \mathfrak{C}\ g\ f\ x)$$

(i.e., we generalize over f and let x – on which we will apply well-founded induction – equal the $n{+}1$-th element of f).

For an arbitrary but fixed x, let $x = f\ (n{+}1)$. Hence, from the assumption $\forall\, i.\ f\ i \in A^{\mathsf{o}}$ we have that $x \in A^{\mathsf{o}}$. Now we use the above induction schema to prove (discharging its first assumption by $x \in A^{\mathsf{o}}$): $\bigwedge x.\ x = f\ (n{+}1) \Longrightarrow \mathfrak{I}\ x$.

Thus, we have $x \in A^{\mathsf{o}}$ for some arbitrary but fixed x, as well as the induction hypothesis (IH) $\bigwedge y.\ [\![y \in A^{\mathsf{o}};\ y \lhd x]\!] \Longrightarrow \mathfrak{I}\ y$. Then we prove $\mathfrak{I}\ x$.

Therefore, we assume $x = f\ (n{+}1)$, $\forall\, i.\ f\ i \in A^{\mathsf{o}}$, $min^A_n(f)$, and $bad_{\preceq_A}(f)$ for some arbitrary but fixed f. Now either we have $min^A_{n+1}(f)$ (and are done) or not. In the latter case we obtain a sequence h such that

$$h\ (n{+}1) \lhd f\ (n{+}1) \tag{1}$$
$$\forall\, i{\geq}n{+}1.\ \exists j{\geq}n{+}1.\ h\ i \trianglelefteq f\ j \tag{2}$$
$$bad_{\preceq_A}(h) \tag{3}$$

employing the definition of *min*.

Let g abbreviate $f\langle n{+}1\rangle h$. From (1), we have $g\ (n{+}1) \in A^{\mathsf{o}}$ and $g\ (n{+}1) \lhd x$ (by basic properties of $\lhd$ and $\langle\_\rangle$). Thus, by IH, we have $\mathfrak{I}\ (g\ (n{+}1))$.

Moreover, we have $\forall\, i.\ g\ i \in A^{\mathsf{o}}$ (by (2) and some basic reasoning), $min^A_n(g)$ (from $min^A_n(f)$ with (2) and some further reasoning) as well as $bad_{\preceq_P}(g)$ by $bad_{\preceq_A}(f)$ and (3) with Lemma 3. Thus, by specializing $\mathfrak{I}\ (g\ (n{+}1))$ to g, we obtain a sequence m such that $\mathfrak{C}\ m\ g\ (g\ (n{+}1))$. Additionally, from (1) and transitivity of $\trianglelefteq$ we have $m\ (n{+}1) \trianglelefteq x$; and with (2) and $\mathfrak{C}\ m\ g\ (g\ (n{+}1))$ we have $\forall\, i{\geq}n{+}1.\ \exists j{\geq}n{+}1.\ m\ i \trianglelefteq f\ j$. Combining the previous facts, we finally have $\exists\, m.\ \mathfrak{C}\ m\ f\ x$, thus finishing our prove of $\bigwedge x.\ x = f\ (n{+}1) \Longrightarrow \mathfrak{I}\ x$. Choosing $x = f\ (n{+}1)$ and using our initial assumptions we have $\exists\, g.\ \mathfrak{C}\ g\ f\ (f\ (n{+}1))$. □

For a step-wise construction of a minimal bad sequence we still need to show that from an arbitrary bad sequence we can obtain one that is minimal at position *0*. This is taken care of by the next lemma.

Lemma 5. $[\![f\ 0 \in A^{\mathsf{o}};\ bad_{\preceq_A}(f)]\!]$
$\Longrightarrow \exists\, g.\ (\forall\, i.\ \exists j.\ g\ i \trianglelefteq f\ j) \wedge min^A_0(g) \wedge bad_{\preceq_A}(g)$

Proof. Similar structure to the proof of Lemma 4 (but much simpler). □

At this point we are ready to prove that if a relation is not almost-full, then there is a minimal bad sequence, thereby taking care of gap (G1).

Theorem 1. $\neg\ af_{A^{\mathsf{o}}}(\preceq_A) \Longrightarrow$
$\exists\, m.\ bad_{\preceq_A}(m) \wedge (\forall\, n.\ min^A_n(m)) \wedge (\forall\, i.\ m\ i \in A^{\mathsf{o}})$

Proof. Assume $\neg\ af_{A^\mathsf{o}}(\preceq_A)$. Then there is a bad sequence f, i.e., $\forall\, i.\ f\ i \in A^\mathsf{o}$ and $bad_{\preceq_A}(f)$. With Lemma 4, together with the axiom of choice, we obtain a choice function ν such that

$$\begin{array}{l}\forall f\ n.\\ \quad (\forall\, i.\ f\ i \in A^\mathsf{o}) \wedge min^A_n(f) \wedge bad_{\preceq_A}(f) \longrightarrow\\ \quad (\forall\, i{\leq}n.\ \nu\ f\ n\ i = f\ i) \wedge\\ \quad \nu\ f\ n\ (n{+}1) \trianglelefteq f\ (n{+}1) \wedge\\ \quad (\forall\, i{\geq}n{+}1.\ \exists j{\geq}n{+}1.\ \nu\ f\ n\ i \trianglelefteq f\ j) \wedge\\ \quad bad_{\preceq_A}(f\langle n{+}1\rangle(\nu\ f\ n)) \wedge min^A_{n+1}(f\langle n{+}1\rangle(\nu\ f\ n))\end{array}$$

That is, $\nu\ f\ n$ provides a witness to Lemma 4, provided that f and n satisfy its assumptions. Moreover, by Lemma 5, we obtain a sequence g such that $\forall\, i.\ \exists j.\ g\ i \trianglelefteq f\ j$ as well as $min^A_0(g)$ and $bad_{\preceq_A}(g)$.

Then we define an auxiliary sequence (of sequences) m' by $m'\ 0 = g$ and $m'\ (n{+}1) = (m'\ n)\langle n{+}1\rangle(\nu\ (m'\ n)\ n)$. We define the desired minimal bad sequence m, to be $\lambda i.\ m'\ i\ i$ (i.e., the "diagonal" of the auxiliary sequence m'). Of course, we have to *prove* that m actually is a minimal bad sequence. To this end, we simultaneously prove the following statements by induction on n:

$$\begin{array}{l}\forall\, i.\ m'\ n\ i \in A^\mathsf{o}\\ \forall\, i{\leq}n.\ min^A_i(m'\ n)\\ \forall\, i{\leq}n.\ m\ i = m'\ n\ i\\ bad_{\preceq_A}(m'\ n)\\ n = 0 \longrightarrow (\forall\, i{\geq}n.\ \exists j{\geq}n.\ m'\ n\ i \trianglelefteq g\ j)\\ 0 < n \longrightarrow (\forall\, i{\geq}n.\ \exists j{\geq}n.\ m'\ n\ i \trianglelefteq m'\ (n - 1)\ j)\end{array}$$

This is the most tedious part of our prove (which we spare the casual reader; all details are available from [7], lemma *mbs* in theory *Almost-Full-Relations*). Just note that the above statements are strengthened in order to make the induction go through. Afterwards, we are able to obtain $bad_{\preceq_A}(m)$, $\forall\, n.\ min^A_n(m)$, and $\forall\, i.\ m\ i \in A^\mathsf{o}$, concluding the proof. □

6 Higman's Lemma

Before we can formally state Higman's lemma for almost-full relations, we need to give a construction that extends a given order on elements to an order on lists: homeomorphic embedding. Furthermore, we need a kind of structural comparison between lists as well as the set of lists built over a given set of elements. The set of lists over elements from a set A, written A^*, is defined inductively:

$$\frac{}{[] \in A^*} \qquad \frac{x \in A \qquad xs \in A^*}{x{\cdot}xs \in A^*}$$

A list xs is a *proper suffix* of the list ys iff $\exists\, us.\ ys = us\ @\ xs \wedge us \neq []$ (we write $xs < ys$). Homeomorphic embedding on lists, for a given base order $\preceq$, is

defined inductively by the rules:

$$\frac{}{[] \preceq^* ys} \qquad \frac{xs \preceq^* ys}{xs \preceq^* y{\cdot}ys} \qquad \frac{x \preceq^= y \qquad xs \preceq^* ys}{x{\cdot}xs \preceq^* y{\cdot}ys}$$

Note that this definition makes $\preceq^*$ reflexive for arbitrary $\preceq$. For reflexive (and thus also for almost-full) $\preceq$, we can replace $\preceq^=$ by $\preceq$. Intuitively, it might be easier to think about homeomorphic embedding on lists as follows: a list xs is embedded in a list ys iff we can obtain xs from ys by dropping elements and replacing elements with arbitrary smaller ones (w.r.t. the base order). An important special case of embedding is $=^*$, which we call the *sublist relation*. In that case we have $xs =^* ys$ iff we can obtain xs from ys by dropping elements.

Using the definitions above, we can instantiate the *mbs* locale as follows (for some arbitrary relation $\preceq$): use $\_^*$ for $\_^\circ$, $(\lambda A\ xs\ ys.\ xs \preceq^* ys)$ for $\preceq$ (for this instance we just discard the parameter A, since we do not need it), and $<$ for $\lhd$. The assumptions of the *mbs* locale are discharged by the following facts (see [7] for the corresponding proofs):

$$wf_{A^*}(<) \qquad [\![xs < ys;\ ys \in A^*]\!] \Longrightarrow xs \in A^*$$
$$[\![xs < ys;\ ys < zs]\!] \Longrightarrow xs < zs \qquad [\![xs \preceq^* ys;\ ys < zs]\!] \Longrightarrow xs \preceq^* zs$$

Thus, we have

$$\neg\ af_{A^*}(\preceq^*) \Longrightarrow \exists\, m.\ bad_{\preceq^*}(m) \wedge (\forall\, n.\ min^A_n(m)) \wedge (\forall\, i.\ m\ i \in A^*)$$

At this point, we can state Higman's lemma for almost-full relations.

Lemma 6. $af_A(\preceq) \Longrightarrow af_{A^*}(\preceq^*)$

Proof. We assume $af_A(\preceq)$ but $\neg\ af_{A^*}(\preceq^*)$, for the sake of a contradiction. Then there is a bad sequence f. This, in turn, implies the existence of a minimal bad sequence m. All lists in m are non-empty (since otherwise m would be good). Hence, there are sequences h and t of heads and tails of m (i.e., $m\ i = h\ i{\cdot}t\ i$).

First we show that there is no index mapping φ such that $\varphi\ 0 \leq \varphi\ i$ for all i and the sequence t_φ is bad. Assume, to the contrary, that such a φ exists. Let n abbreviate $\varphi\ 0$ and c be the combination of m with t, defined by $c\ i \stackrel{\text{def}}{=} \mathit{if}\ i < n\ \mathit{then}\ m\ i\ \mathit{else}\ t\ (\varphi\ (i - n))$ (i.e., c is the same as t_φ, but prepended by the first n elements of m). Then c is bad, since otherwise we obtain a contradiction as follows. Assume c is good. Then we obtain $i < j$ such that $c\ i \preceq^* c\ j$. Now we analyze the following cases:

- **case** $(j < n)$. Then $m\ i \preceq^* m\ j$, contradicting badness of m.
- **case** $(n \leq i)$. Let $i' = i - n$ and $j' = j - n$. Then $i' < j'$ and $t_\varphi\ i' \preceq^* t_\varphi\ j'$, contradicting badness of t_φ.
- **case** ($i < n$ and $n \leq j$). Let $j' = j - n$. We have $t\ (\varphi\ j') \leq m\ (\varphi\ j')$ (since the tail of a non-empty list is obviously also a suffix) and $m\ i \preceq^* t\ (\varphi\ j')$ (from $c\ i \preceq^* c\ j$). Then $m\ i \preceq^* m\ (\varphi\ j')$ (since the suffix relation is a special case of embedding and embedding is transitive). Since we also have $i < \varphi\ j'$, this contradicts the badness of m.

Thus, c is bad. Furthermore, we have $\forall i{<}n.\ c\ i = m\ i$, $c\ n < m\ n$, and $\forall i{\geq}n.\ \exists j{\geq}n.\ c\ i \leq m\ j$, and thus c is good (since m is minimal): A contradiction, concluding the proof of

$$\nexists \varphi.\ (\forall i.\ \varphi\ 0 \leq \varphi\ i) \wedge \mathit{bad}_{\preceq^*}(t_\varphi). \tag{$\star$}$$

Let H and T denote the sets of heads and tails in m, respectively (i.e., $H = \bigcup_i\{h\ i\}$ and $T = \bigcup_i\{t\ i\}$). We obviously have that $\preceq$ is almost-full on H (since $H \subseteq A$ and $\preceq$ is almost-full on A). Moreover, since every bad sequence over T would admit a subsequence of the shape in ($\star$), we obtain that $\preceq^*$ is almost-full on T. With Lemma 1, we have that the pointwise combination of $\preceq$ and $\preceq^*$ is almost-full on $H \times T$. Thus, there are $i < j$ with $h\ i \preceq^= h\ j$ and $t\ i \preceq^*\ t\ j$. By the definition of $\preceq^*$, this implies $m\ i \preceq^*\ m\ j$, contradicting the badness of m. □

But wait a moment, "*since every bad sequence over* T *...*" above, is exactly gap (G2). To close it, we prove the claim by the lemma:

Lemma 7. $[\![\mathit{refl}_{\bigcup_i\{t\ i\}}(\preceq);\ \forall i.\ f\ i \in \bigcup_i\{t\ i\};\ \mathit{bad}_\preceq(f)]\!]$
$\Longrightarrow \exists \varphi.\ (\forall i.\ \varphi\ 0 \leq \varphi\ i) \wedge \mathit{bad}_\preceq(t_\varphi)$

Proof. Assume that $\preceq$ is reflexive (on $\bigcup_i\{t\ i\}$), and f is a bad sequence (over $\bigcup_i\{t\ i\}$). First note that for every i, there exists a j such that $f\ i = t\ j$. By the axiom of choice, we obtain an index mapping φ' with $f\ i = t_{\varphi'}\ i$ for all i. Since f is bad, also $t_{\varphi'}$ is bad. Next we prove that

$$\text{for every } i \text{ there is a } j > i \text{ such that } \varphi'\ 0 \leq \varphi'\ j. \tag{$\star$}$$

Assume otherwise, then there is some i such that for all $j > i$ we have $\varphi\ j < \varphi'\ 0$. Thus, the image of φ' under $\{j \mid i < j\}$ is finite, whereas $\{j \mid i < j\}$ itself is infinite. By the pigeonhole principle, we obtain a $k > i$ such that there are infinitely many $j > i$ with $\varphi'\ j = \varphi'\ k$. But then, there is some $l > k$ for which $\varphi'\ l = \varphi'\ k$. Since $\preceq$ is reflexive and $k < l$, this implies that $t_{\varphi'}$ is good; a contradiction. Using ($\star$) and the axiom of choice, we obtain an index mapping ψ' such that $i < \psi'\ i$ and $\varphi'\ 0 \leq \varphi'\ (\psi'\ i)$ for all i. Now, let ψ abbreviate $\lambda i.\ \psi'^{\,i}\ 0$ (the i-fold application of ψ' to 0) and φ abbreviate $\varphi' \circ \psi$. Then we have that ψ is strictly monotone and $\varphi\ 0 \leq \varphi\ i$ for all i. Moreover, since $t_{\varphi'}$ is bad and ψ is monotone, we also have that t_φ is bad. This concludes the proof. □

Higman's Lemma. $\mathit{wqo}_A(\preceq) \Longrightarrow \mathit{wqo}_{A^*}(\preceq^*)$

Proof. We refer to lemma *list-hembeq-trans* in theory *Sublist*, for transitivity of $\preceq^*$ (under the assumption that $\preceq$ is transitive). Together with Lemma 6, we obtain Higman's lemma. □

7 The Tree Theorem

The tree theorem is for finite trees, what Higman's lemma is for finite lists. However, whereas for finite lists, their representation inside Isabelle/HOL is

quite unambiguous and the existing datatype is generally applicable; this is not so much the case for finite trees. Consider the following two datatypes

datatype $\alpha\ t = \mathit{Node}\ \alpha\ (\alpha\ t\ \mathit{list})$
datatype $\alpha\ t' = \mathit{Empty} \mid \mathit{Node}\ \alpha\ (\alpha\ t'\ \mathit{list})$

or the type of first-order terms

datatype $(\alpha, \beta)\ \mathit{term} = \mathit{Var}\ \beta \mid \mathit{Fun}\ \alpha\ ((\alpha, \beta)\ \mathit{term}\ \mathit{list})$

also a kind of finite tree (and more importantly, one of the types to which we intend to apply the tree theorem, in order to formalize the fact that the Knuth-Bendix order is a simplification order [11]). Restricting our results to a specific datatype would strongly restrict their applicability. Therefore, we again employ Isabelle/HOL's locale mechanism. This time, for a locale *finite-tree* that fixes the following constants:

- A function $\mathit{mk}::\beta \Rightarrow \alpha\ \mathit{list} \Rightarrow \alpha$ that is used to construct a finite tree from a given node and a given list of finite trees.
- A function $\mathit{root}::\alpha \Rightarrow \beta$ that extracts the root node from a given tree.
- As well as a function $\mathit{succs}::\alpha \Rightarrow \alpha\ \mathit{list}$ that extracts the list of direct subtrees (successors) from a given tree.

These constants are required to satisfy the following assumptions (thereby turning *mk* into kind of a datatype constructor with extractors *root* and *succs*):

$$\mathit{root}\ (\mathit{mk}\ f\ \mathit{ts}) = f \tag{F1}$$
$$\mathit{succs}\ (\mathit{mk}\ f\ \mathit{ts}) = \mathit{ts} \tag{F2}$$
$$(\mathit{mk}\ f\ \mathit{ss} = \mathit{mk}\ g\ \mathit{ts}) = (f = g \wedge \mathit{ss} = \mathit{ts}) \tag{F3}$$

As opposed to a real datatype, the above assumptions do not guarantee that *all* finite trees are built from a finite number of applications of *mk*. Thus, we define the set of finite trees over nodes from A, written $\mathcal{T}(A)$, inductively by:

$$\frac{f \in A \qquad \forall\, t{\in}\mathit{set}\ \mathit{ts}.\ t \in \mathcal{T}(A)}{\mathit{mk}\ f\ \mathit{ts} \in \mathcal{T}(A)}$$

The notion of structural decrease, as needed to instantiate the *mbs* locale, is provided by the *subtree* relation:

$$\frac{t \in \mathit{set}\ \mathit{ts}}{t \lhd \mathit{mk}\ f\ \mathit{ts}} \qquad \frac{s \lhd t \qquad t \in \mathit{set}\ \mathit{ts}}{s \lhd \mathit{mk}\ f\ \mathit{ts}}$$

Where a tree s is a proper subtree of another tree t, if it is either a direct subtree of t itself or a proper subtree of one of the direct subtrees of t.

Homeomorphic embedding on finite trees is also defined inductively by:

$$\frac{t \in \mathit{set}\ \mathit{ts}}{t \preceq_{\mathsf{emb}} \mathit{mk}\ f\ \mathit{ts}} \qquad \frac{s \preceq_{\mathsf{emb}} t \qquad t \preceq_{\mathsf{emb}} u}{s \preceq_{\mathsf{emb}} u}$$

$$\frac{s \preceq_{\mathsf{emb}} t}{mk\ f\ (ss_1\ @\ s{\cdot}ss_2) \preceq_{\mathsf{emb}} mk\ f\ (ss_1\ @\ t{\cdot}ss_2)} \qquad \frac{f \preceq^{=} g \qquad ss =^{*} ts}{mk\ f\ ss \preceq_{\mathsf{emb}} mk\ g\ ts}$$

The first three rules are easy: homeomorphic embedding extends the subtree relation, is transitive, and is closed under contexts. The last rule states that we may replace the nodes of a tree by smaller ones (w.r.t. $\preceq$) as well as drop arbitrary successors. From this definition, we can prove the following property:

Lemma 8. $[\![f \preceq^{=} g;\ ss \preceq_{\mathsf{emb}}{}^{*}\ ts]\!] \Longrightarrow mk\ f\ ss \preceq_{\mathsf{emb}} mk\ g\ ts$

Proof. This property seems obvious, as $\preceq_{\mathsf{emb}}$ is reflexive, transitive, and closed under contexts. However, it turns out to be surprisingly tedious to formalize (or at least *we* did not find an elegant way). We spare the reader some tedium and refer to lemma *tree-hembeq-list-hembeq* in theory *Finite-Tree* for details. □

To instantiate the *mbs* locale, we proved the following facts (see [7] for proofs):

$$wf_{\mathcal{T}(A)}(\lhd) \qquad [\![s \lhd t;\ t \in \mathcal{T}(A)]\!] \Longrightarrow s \in \mathcal{T}(A)$$
$$[\![s \lhd t;\ t \lhd u]\!] \Longrightarrow s \lhd u \qquad [\![s \preceq_{\mathsf{emb}} t;\ t \lhd u]\!] \Longrightarrow s \preceq_{\mathsf{emb}} u$$

Thus, we have

$$\neg\ af_{\mathcal{T}(A)}(\preceq_{\mathsf{emb}}) \Longrightarrow \exists\, m.\ bad_{\preceq_{\mathsf{emb}}}(m) \wedge (\forall\, n.\ min^{A}_{n}(m)) \wedge (\forall\, i.\ m\ i \in \mathcal{T}(A))$$

Finally, we can state and prove the tree theorem for almost-full relations.

Theorem 2. $af_{A}(\preceq) \Longrightarrow af_{\mathcal{T}(A)}(\preceq_{\mathsf{emb}})$

Proof. We assume $af_{A}(\preceq)$ but $\neg\ af_{\mathcal{T}(A)}(\preceq_{\mathsf{emb}})$ for the sake of a contradiction. Then there is a bad sequence and thus a minimal bad sequence m. All trees in m are in the set $\mathcal{T}(A)$ (and thus non-empty). Hence, there are sequences r and s of roots and successor lists of m (i.e., $m\ i = mk\ (r\ i)\ (s\ i)$).

First we show that there is no sequence of trees t and index mapping φ such that $t\ i \in set\ (s_{\varphi}\ i)$ (i.e., the sequence t selects an arbitrary successor of $m_{\varphi}\ i$ as its i-th element) and $\varphi\ 0 \leq \varphi\ i$ for all i, and t is bad. Assume, to the contrary, that such t and φ exist. Let n abbreviate $\varphi\ 0$ and c be the sequence defined by $c\ i \stackrel{\text{def}}{=} \mathit{if}\ i < n\ \mathit{then}\ m\ i\ \mathit{else}\ t\ (i - n)$. Then c is bad, since assuming that it was good results in a contradiction by a similar case analysis as we conducted in the proof of Lemma 6 above. Furthermore, we have $\forall\, i{<}n.\ c\ i = m\ i$, $c\ n \lhd m\ n$, and $\forall\, i{\geq}n.\ \exists\, j{\geq}n.\ c\ i \trianglelefteq m\ j$, and thus c is good (since m is minimal). This contradiction concludes the proof of

$$\nexists\, t\ \varphi.\ (\forall\, i.\ t\ i \in set\ (s_{\varphi}\ i) \wedge \varphi\ 0 \leq \varphi\ i) \wedge bad_{\preceq_{\mathsf{emb}}}(t). \qquad (\star)$$

Let R and S denote the sets of roots and successor lists in m (i.e., $R = \bigcup_i \{r\ i\}$ and $S = \bigcup_i \{s\ i\}$). Clearly, $\preceq$ is almost-full on R (since $R \subseteq A$). Let S' abbreviate $\{t \mid \exists\, i.\ t \in set\ (s\ i)\}$. Every bad sequence over S' would admit a sequence of the shape in ($\star$), thus $\preceq_{\mathsf{emb}}$ is almost-full on S'. From Lemma 6,

together with $S \subseteq S'^*$, we have that $\preceq_{\mathsf{emb}}{}^*$ is almost-full on S. With Lemma 1, we have that the pointwise combination of $\preceq$ and $\preceq_{\mathsf{emb}}{}^*$ is almost-full on $R \times S$. Thus, there are $i < j$ such that $r\ i \preceq^{=} r\ j$ and $s\ i \preceq_{\mathsf{emb}}{}^*\ s\ j$, which, employing Lemma 8, implies that $m\ i \preceq_{\mathsf{emb}}\ m\ j$ and thus contradicts the badness of m. □

Note that "*Every bad sequence over* $S' \ldots$" above, corresponds to gap (G3). To close it, we prove the lemma:

Lemma 9. *Let* $\preceq$ *be a binary relation and* X *be the set* $\{t \mid \exists\, i.\ t \in \mathit{set}\ (s\ i)\}$ *for a sequence of lists* s*. Then we have*

$[\![\mathit{refl}_X(\preceq);\ \forall\, i.\ f\ i \in X;\ \mathit{bad}_{\preceq}(f)]\!]$
$\Longrightarrow \exists\, t\ \varphi.\ (\forall\, i.\ t\ i \in \mathit{set}\ (s_\varphi\ i) \wedge \varphi\ 0 \leq \varphi\ i) \wedge \mathit{bad}_{\preceq}(t)$

Proof. The proof is structured similar to the proof of Lemma 7 but slightly more involved, due to the extra indirection via list elements. For details, we refer to lemma *bad-of-special-shape′* in theory *Kruskal-Auxiliaries* of [7]. □

Kruskal's Tree Theorem. $\mathit{wqo}_A(\preceq) \Longrightarrow \mathit{wqo}_{\mathcal{T}(A)}(\preceq_{\mathsf{emb}})$

Proof. Theorem 2 and transitivity of $\preceq_{\mathsf{emb}}$ yield the tree theorem. □

8 Conclusion

We presented our Isabelle/HOL formalization of three important results from combinatorics: Dickson's lemma, Higman's lemma, and Kruskal's tree theorem.

Parts of our formalization were used by Wu et al. [12] to formalize a proof of: *For every language A, the languages of sub- and superstrings of A are regular.* (Details are presented in a submitted journal version of [13].)

Moreover, in [11], we employ the tree theorem for a proof that the Knuth-Bendix order is a simplification order. To this end, we actually need a variant of the tree theorem as presented here – which we call *the term theorem*. The reason is that in the above mentioned proof it is essential to consider arities of function symbols, whereas in Section 7, we allow a node in a tree to have an arbitrary (finite) number of successors. The term theorem is the reason for having the additional parameter A in $\preceq_A$ in Section 5 (which is neither used in Higman's lemma nor in the tree theorem), since in the presence of a signature we have to make sure that the terms we compare are well-formed w.r.t. this signature.

As future work, we will investigate whether the tedious induction in the proof of Theorem 1 can be replaced by an invocation of Zorn's lemma (and this in turn, by an application of open induction [14,15], thereby hopefully giving also insight in the computational content of the minimal bad sequence argument).

There are formalizations of Higman's lemma in Isabelle/HOL [16] and other proof assistants [17,18,19,20,21]. Those usually strive for constructive proofs, whereas our approach is purely classical. An intuitionistic proof of Kruskal's tree theorem is presented in [9]. However, to the best of our knowledge our work constitutes the first formalization of the tree theorem in a proof assistant ever.

Acknowledgments. We thank Mizuhito Ogawa (小川先生) for helpful discussions on everything related to the tree theorem, as well as enabling (together with the Austrian Science Fund) our stay in Japan.

References

1. Kruskal, J.B.: Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. **95**(2) (1960) 210–225 doi:10.2307/1993287.
2. Nash-Williams, C.S.J.A.: On well-quasi-ordering finite trees. Proc. Cambridge Philos. Soc. **59**(4) (1963) 833–835 doi:10.1017/S0305004100003844.
3. Dickson, L.E.: Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. Amer.J.Math. **35**(4) (1913) 413–422 doi:10.2307/2370405.
4. Higman, G.: Ordering by divisibility in abstract algebras. Proc. London Math. Soc. **s3-2**(1) (1952) 326–336 doi:10.1112/plms/s3-2.1.326.
5. Nipkow, T., et al.: Isabelle/HOL - A Proof Assistant for Higher-Order Logic. Volume 2283 of LNCS. (2002) doi:10.1007/3-540-45949-9.
6. Sternagel, C.: A locale for minimal bad sequences. In: IUW'12. arXiv:1208.1366.
7. Sternagel, C.: Well-Quasi-Orders. In Klein, G., et al., eds.: AFP. (2012) http://afp.sf.net/devel-entries/Well_Quasi_Orders.shtml.
8. Haftmann, F., et al.: LaTeX sugar for Isabelle documents (2012) http://isabelle.in.tum.de/dist/Isabelle2012/doc/sugar.pdf.
9. Veldman, W.: An intuitionistic proof of Kruskal's theorem. Arch. Math. Logic **43**(2) (2004) 215–264 doi:10.1007/s00153-003-0207-x.
10. Vytiniotis, D., et al.: Stop when you are almost-full - adventures in constructive termination. In Beringer, L., et al., eds.: ITP'12. Volume 7406 of LNCS. 250–265 doi:10.1007/978-3-642-32347-8_17.
11. Sternagel, C., Thiemann, R.: Formalizing Knuth-Bendix orders and Knuth-Bendix completion. In: RTA'13. submitted.
12. Wu, C., et al.: The Myhill-Nerode theorem based on regular expressions. In Klein, G., et al., eds.: AFP. (2011) http://afp.sf.net/entries/Myhill-Nerode.shtml.
13. Wu, C., et al.: A formalisation of the Myhill-Nerode theorem based on regular expressions (proof pearl). In van Eekelen, M., et al., eds.: ITP'11. Volume 6898 of LNCS. 341–356 doi:10.1007/978-3-642-22863-6_25.
14. Raoult, J.C.: Proving open properties by induction. Inform. Process. Lett. **29**(1) (1988) 19–23 doi:10.1016/0020-0190(88)90126-3.
15. Ogawa, M., Sternagel, C.: Open Induction. In Klein, G., et al., eds.: AFP. (2012) http://afp.sf.net/devel-entries/Open_Induction.shtml.
16. Berghofer, S.: A constructive proof of Higman's lemma in Isabelle. In Berardi, S., et al., eds.: TYPES'03. Volume 3085 of LNCS. 66–82 doi:10.1007/978-3-540-24849-1_5.
17. Murthy, C.R.: Extracting Constructive Content from Classical Proofs. PhD thesis, Cornell University (1990) http://hdl.handle.net/1813/6991.
18. Fridlender, D.: Higman's lemma in type theory. In Giménez, E., et al., eds.: TYPES'96. Volume 1512 of LNCS. 112–133 doi:10.1007/BFb0097789.
19. Herbelin, H.: A program from an A-translated impredicative proof of Higman's lemma (1994) http://coq.inria.fr/pylons/contribs/view/HigmanNW/v8.3.
20. Seisenberger, M.: On the Constructive Content of Proofs. PhD thesis, LMU Munich (2003) http://nbn-resolving.de/urn:nbn:de:bvb:19-16190.
21. Martín-Mateos, F.J., et al.: A formal proof of Higman's lemma in ACL2. J. Autom. Reason. **47**(3) (2011) 229–250 doi:10.1007/s10817-010-9178-x.