Feasible proofs of Szpilrajn's Theorem
A proof-complexity framework for concurrent automata

Michael Soltys

McMaster University soltys@mcmaster.ca

Abstract. The aim of this paper is to propose a proof-complexity framework for concurrent automata. Since the behavior of concurrent processes can be described with partial orders, we start by formalizing proofs of Szpilrajn's theorem. This theorem says that any partial order may be extended to a total order. We give two feasible proofs of the finite case of Szpilrajn's theorem. The first proof is formalized in the logical theory **LA** extended to ordered rings; this yields a TC^0 Frege derivation. The second proof is formalized in the logical theory $\exists$**LA** and yields a $\mathsf{P/poly}$ Frege derivation. Although TC^0 is a much smaller complexity class than $\mathsf{P/poly}$, the trade-off is that the $\mathsf{P/poly}$ proof is algebraically simpler—it requires the algebraic theory **LA** over the simplest of rings: $\mathbb{Z}_2$.

Keywords: Proof complexity, concurrency, partial orders.

1 Introduction

The purpose of proof complexity is to study logical systems which use restricted reasoning based on concepts from computational complexity; see [CN10] for an introduction to the subject.

In this paper we propose to investigate the proof complexity of standard reasoning associated with finite orders. The aim is to establish a proof complexity framework for concurrent automata—a framework capable of formalizing, for example, the reasoning in [Jan08] and [Lê08]. See [JK93] for background related to the fundamentals of concurrency, such as order structures and traces, where Szpirlajn's theorem finds its important applications.

In particular, [Jan08] deals with the foundations of concurrency theory, and the author shows how structurally complex concurrent behaviors can be modeled by relational structures. We propose a proof-complexity framework (i.e., a logical theory) able to formalize this kind of reasoning. We start at the very beginning, and propose the logical theory **LA** (introduced in [SC04]) as an appropriate theory for formalizing reasoning about that most basic of relations: orders.

Szpilrajn's theorem[1], which says that every partial order can be extended to a total order, is a central result about orders. Here we give two proofs of

[1] As is pointed out in [Lê08, pg. 9], Szpilrajn's theorem is a fundamental theorem in the foundation of concurrency theory. The original proof, written in French, can be found in [Szp30].

the finite case of Szpilrajn's theorem. The first one follows an algorithm due to Ruzzo (cited as private communication in [Coo85]) for computing the simple extension of a finite order. The second follows the presentation in [Mon]. Both proofs are formalized in variants of the logical theory **LA** (see [SC04]), and in the first case the formalization yields a TC^0 proof, while in the second a P proof.

It is very interesting to observe that the two approaches trade conceptual complexity for computational complexity; we require ordered rings (and thus extra axioms) for a low-complexity proof (TC^0 Frege), while if we dispense with ordered rings we require a stronger induction ($\exists\mathbf{LA}$ induction) and obtain a higher complexity proof ($\mathsf{P/poly}$ Frege).

2 Background

2.1 Orders

Let A be a set, and $<$ a binary relation on A, i.e., $<\subseteq A \times A$. We say that $<$ is a *partial order* if it is irreflexive (i.e., for all $a \in A$ we have $\neg(a < a)$) and transitive (i.e., for all $a, b, c \in A$, if $a < b$ and $b < c$, then $a < c$). The pair $(A, <)$ is called a partially ordered set, or poset for short. In this paper we are interested in finite posets. If for all $a, b \in A$, it is the case that $a < b$ or $b < a$ or $a = b$ (i.e., any two elements in A are comparable), then $(A, <)$ is in fact a *simple* order (also known as a *total* order).

Szpilrajn's theorem says that if $(A, <)$ is a partially ordered set, then there always exists a simple order $(A, \prec)$ such that $<\subseteq\prec$, i.e., a simple order $\prec$ that extends $<$. Note that this is true whether A is finite or infinite, but when A is infinite the Axiom of Choice is required for the proof. Since we are interested in partial orders as a formalism for expressing the behavior of concurrent automata, we concentrate on finite orders.

2.2 LA

LA is a first order logical theory geared towards algebraic reasoning—we define it precisely in section 5, at the end of the paper. We present just enough background to formalize proofs of Szpilrajn's theorem in different extensions of **LA**.

By translating those proofs into propositional logic we obtain TC^0 and $\mathsf{P/poly}$ Frege proofs of Szpilrajn's theorem. Boolean circuits, or just circuits as well call them in this paper, are the archetypical example of computation: these are Boolean gates connected by directed edges; they can be seen as directed acyclic graphs, where the input "flows" from the input-gates to the output gates, modified by AND, OR, NOT gates on the way. $\mathsf{P/poly}$ simply denotes circuit families of polynomial size, while TC^0 is defined in section 3 below.

We are going to encode finite posets as matrices over $\mathbb{Z}_2 = \{0, 1\}$, the ring of two elements. Given a finite set $A = \{a_1, a_2, \ldots, a_n\}$, and a poset $(A, <)$, we let $M_{(A,<)}$ be an $n \times n$ matrix where the entry (i, j) is 1 if and only if $a_i < a_j$ (and 0 otherwise).

Let Σ_0^B be the set of formulas over the language $\mathcal{L}_{\mathbf{LA}}$ *without* matrix quantifiers, *without* ring elements quantifiers, and where all the index quantifiers (which are allowed in any number and any alternation) are bounded. Free matrix and ring elements variables are allowed.

Let Σ_1^B be the set of formulas which contain all the Σ_0^B formulas, and also those formulas which in prenex form have a single bounded matrix quantifier in front, followed by a Σ_0^B formula, i.e., formulas which in prenex form is $(\exists M \leq t)\alpha$ where $\alpha \in \Sigma_0^B$ and t is term *without* the matrix variable M.

Given a square matrix M, we can state with a Σ_0^M-formula over $\mathcal{L}_{\mathbf{LA}}$ that M represents a poset:

$$\textsc{Poset}(M) := (\forall i \leq |M|) M_{ii} = 0 \wedge (\forall ijk \leq |M|)[(M_{ij} \wedge M_{jk}) \rightarrow M_{ik}], \quad (1)$$

that M is a simple order:

$$\textsc{Simple}(M) := \textsc{Poset}(M) \wedge (\forall ij \leq |M|)[M_{ij} = 1 \vee M_{ji} = 1 \vee i = j], \quad (2)$$

and that N extends M:

$$M \sqsubseteq N := |M| = |N| \wedge (\forall ij \leq |M|)[M_{ij} = 1 \rightarrow N_{ij} = 1]. \quad (3)$$

We can state Szpilrajn's theorem with a Σ_1^B-formula over $\mathcal{L}_{\mathbf{LA}}$:

$$\textsc{Szpilrajn}(M) := \textsc{Poset}(M) \rightarrow \exists N \leq |M|[\textsc{Simple}(N) \wedge M \sqsubseteq N]. \quad (4)$$

2.3 The circuit class TC^0

The (nonuniform) complexity class TC^0 consists of languages that are accepted by a family of polynomial-size constant-depth circuits whose gates can be Boolean gates or the majority gate. A *majority* gate has unbounded fan-in and outputs 1 if and only if the number of 1 inputs is more than the number of 0 inputs.

Instead of the majority gates, TC^0 can be equivalently defined using counting gates or threshold gates. A *counting* gate C_k (for $k \in \mathbb{N}$) has unbounded fan-in, and $\mathrm{C}_k(x_1, x_2, \ldots, x_n)$ is true if and only if there are exactly k inputs x_i that are true. Similarly, the *threshold* gate Th_k has unbounded fan-in, and $\mathrm{Th}_k(x_1, x_2, \ldots, x_n)$ is true if and only if there are at least k inputs that are true.

The class TC^0 is the smallest class known to contain problems such as sorting, integer multiplication and division—when the input integer arguments are presented in binary. See [CN10, §9C] for the logical theories related to TC^0.

3 A TC^0 proof

We show that TC^0-Frege can prove Szpilrajn's theorem in the finite case. To do so, we extend the theory **LA** to ordered rings, and formalize in this new theory the proof of correctness of an algorithm due to Ruzzo (cited as private communication in [Coo85]) for computing simple extensions.

Let our language be $\mathcal{L}_{\mathbf{LA}} \cup \{<_{\text{ring}}\}$, and we extend **LA** to ordered rings by adding the following two axioms defining $<_{\text{ring}}$ ([Mar02, Example 1.2.9]):

$$\begin{aligned}
&0 <_{\text{ring}} 1\\
&(\forall x \forall y \forall z)[x <_{\text{ring}} y \rightarrow (x + y <_{\text{ring}} y + z)]\\
&(\forall x \forall y \forall z)[(x <_{\text{ring}} y \wedge 0 <_{\text{ring}} z) \rightarrow x \cdot z <_{\text{ring}} y \cdot z],
\end{aligned} \tag{5}$$

we call the new theory $\mathbf{LA}_{\text{or}}$, and we write $<$ instead of $<_{\text{ring}}$ (and $\leq$ instead of $\leq_{\text{ring}}$) when it is clear from the context that we are comparing ring elements.

Suppose that M is a poset, i.e., $\textsc{Poset}(M)$. We compute N, the simple extension of M, as follows: let $c_M(i)$ represent the number of ones in column i of M; it can be defined in **LA**:

$$c_M(i) := \Sigma \lambda pq \langle r(M), 1, e(M, p, i)\rangle.$$

Now we define the matrix N (in **LA**):

$$N := \lambda ij \langle r(M), c(M), \text{cond}(c_M(i) < c_M(j) \vee [c_M(i) = c_M(j) \wedge i < j], 1, 0)\rangle. \tag{6}$$

Note that the symbol "<" appears used with two different meanings. First, as $c_M(i) <_{\text{ring}} c_M(j)$, where it compares ring elements, and second as $i <_{\text{index}} j$, where it compares index elements.

In words, $N_{ij} = 1$ iff column j has more 1s than column i or if they have the same number of 1s, then $i < j$. Note that the number of 1s in column i is the number of predecessors of i; thus, we order all the elements of M by the number of predecessors they have (and if they have the same number of predecessors, by column number).

Lemma 1. $\mathbf{LA}_{\text{or}} \vdash M \sqsubseteq N$.

Proof. Suppose that $M_{ij} = 1$. To show that $N_{ij} = 1$ as well, we show that $c_M(i) < c_M(j)$. So consider column i of M; if the k-th entry of this column is 1, i.e., if $M_{ki} = 1$, then by transitivity $M_{kj} = 1$. Thus, whenever we have a 1 in position k of column i, we also have a 1 in position k of column j. We now prove the following claim.

Claim. $\mathbf{LA}_{\text{or}}$ proves that if A, B are matrices of size n, and it is the case that $(\forall r \forall s \leq n)[A_{rs} = 1 \rightarrow B_{rs} = 1]$, then $\Sigma A \leq \Sigma B$.

Proof. We show the claim by ($\mathbf{LA}_{\text{or}}$) induction on n. Clearly if $n = 1$ then $e(A, 1, 1) \leq e(B, 1, 1)$. Assume it holds for n; to show that it holds for $n + 1$ divide the matrices into 4 parts: the principal submatrix of A, the upper-left corner element, and the remaining row and column, i.e., $A[n], e(A, n+1, n+1)$, the row R, and the column S, respectively (see definitions given in equation (14), at the end of the paper).

The claim $(\forall r \forall s \leq n+1)[A_{rs} = 1 \rightarrow B_{rs} = 1]$ holds for all the submatrices, so we apply the induction hypothesis to them, and then use the fact that $a < b$ and $c < d$ implies that $a + c < b + d$.

Using the claim we conclude that $c_M(i) \leq c_M(j)$. But we need strict inequality, which follows from the fact that $M_{ii} = 0$ (irreflexivity) while $M_{ij} = 1$.

It follows therefore that $M \sqsubseteq N$.

Lemma 2. $\mathbf{LA}_{\mathrm{or}} \vdash \textsc{Simple}(N)$.

Proof. From the definition of N (i.e., from (6)) we see that $(\forall i \leq n) N_{ii} = 0$, i.e., N is irreflexive.

Now we show that N is transitive.

Suppose that $N_{ij} = 1 \wedge N_{jk} = 1$. We want to show that $N_{ik} = 1$; we consider the following four cases.

Case 1. $c(i) < c(j) \wedge c(j) < c(k)$
Then, by transitivity of $<$ we have that $c(i) < c(k)$.

Case 2. $c(i) < c(j) \wedge [c(j) = c(k) \wedge j < k]$
Then $c(i) < c(k)$ by virtue of equality.

Case 3. $[c(i) = c(j) \wedge i < k] \wedge c(j) < c(k)$
$c(i) < c(k)$.

Case 4. $c(i) = c(j) \wedge i < j \wedge c(j) = c(k) \wedge j < k$
Then $c(i) = c(k)$ by transitivity of equality, and $i < k$ by transitivity of $<$.

Finally, we can see that N is simple directly from its definition (6).

We therefore obtain the following theorem.

Theorem 1. $\mathbf{LA}_{\mathrm{or}} \vdash \textsc{Szpilrajn}(M)$.

Corollary 1. TC^0 *Frege proves Szpilrajn's Theorem.*

Proof. Following [SC04] we know that $\mathbf{LA}$ over $\mathbb{Z}$ can be translated into TC^0 Frege. But then so can $\mathbf{LA}_{\mathrm{or}}$ since the three axioms given by (5) can be translated into TC^0 formulas and proven in TC^0 Frege (they just require proving laws of addition and multiplication over integers).

4 A P proof

We now proceed to show that the theory $\exists\mathbf{LA}$ proves Szpilrajn's theorem in the finite case. By employing a stronger induction than the one in $\mathbf{LA}$ (or in $\mathbf{LA}_{\mathrm{or}}$ for that matter) we are able to prove Szpilrajn's theorem over the ring $\mathbb{Z}_2$ (rather than the ring $\mathbb{Z}$) without introducing new axioms (recall that we needed to bring in axioms for an ordered ring, namely the axioms given by (5), to give a TC^0 proof). We pay for this in our translations which now yield $\mathsf{P/poly}$ Frege proofs instead of TC^0 Frege.

Our goal now is to show that

$$\exists\mathbf{LA} \vdash \textsc{Szpilrajn}(M). \tag{7}$$

By a standard witnessing argument, it follows that the extension N can be computed in polynomial time in $|M|$ (which we already knew—what is of interest is the type of reasoning we now employ).

We prove $\textsc{Szpilrajn}(M)$ by induction on the principal submatrices of M (define the *principal submatrix* of a matrix A to be A with the last row and column removed); let $M[i]$ be the matrix consisting of the first i rows and first i columns on M. In particular, $M[1] = [M_{11}]$ and $M[|M|] = M$, and the principal submatrix of M is $M[|M| - 1]$. We want to show that

$$\exists\mathbf{LA} \vdash (\forall i)\textsc{Szpilrajn}(M[i]) \tag{8}$$

by induction on i (which is Σ_1^B-induction); since trivially

$$\exists\mathbf{LA} \vdash (\forall i)\textsc{Szpilrajn}(M[i]) \rightarrow \textsc{Szpilrajn}(M),$$

showing (8) gives us (7), i.e., Szpilrajn's theorem.

Thus, we want to show

$$\exists\mathbf{LA} \vdash \textsc{Szpilrajn}(M[1]) \wedge \forall i[\textsc{Szpilrajn}(M[i]) \rightarrow \textsc{Szpilrajn}(M[i+1])], \tag{9}$$

and conclude (8) from it by invoking Σ_1^B-induction.

The basis case, $\textsc{Szpilrajn}(M[1])$, is trivial since $M[1] = [M_{11}] = [0]$ and $[0]$ is its own extension, i.e., $N = M[1]$.

For the induction step, consider $M[i+1]$. By the induction hypothesis we have an extension N of $M[i]$; we show how to construct an extension N' of $M[i+1]$ from N. For the sake of clarity we unclutter our notation: set $M = M[i]$ and $M' = M[i+1]$.

To construct N' from M' and N we are going to use the λ-constructions of new matrices allowed in $\mathbf{LA}$. First,

$$X := \{u : u \leq i \wedge (\exists v \leq i)[M'_{v,i+1} = 1 \wedge [N_{u,v} = 1 \vee u = v]]\}. \tag{10}$$

Using the auxiliary matrix X we can now construct N':

$$N'_{pq} = 1 \iff \begin{cases} p, q \leq i \wedge N_{pq} = 1 \\ \text{or } p \in X \wedge q = i+1 \\ \text{or } p = i+1 \wedge q \in [i] - X \end{cases} \tag{11}$$

(All entries of N' not defined to be 1 are of course 0; in particular, $N'_{i+1,i+1} = 0$. The notation $[i]$ denotes the set $\{1, 2, \ldots, i\}$.)

We prove the correctness of this construction, i.e., we show that the output N' is a simple order which extends M'. But we must formalize this proof of correctness in $\exists\mathbf{LA}$, so we show:

$$\exists\mathbf{LA} \vdash \textsc{Simple}(N') \wedge M' \sqsubseteq N'. \tag{12}$$

The next two lemmas show (12).

Lemma 3. *$\exists\mathbf{LA} \vdash M' \sqsubseteq N'$.*

Proof. To show that $M' \sqsubseteq N'$ we must show that $M'_{p,q} = 1 \rightarrow N'_{p,q} = 1$ for $1 \leq p, q \leq i+1$. We consider three cases.

Case 1. $1 \leq p, q \leq i$:
Since for such p, q we have $M'_{p,q} = M_{p,q}$, and by the induction hypothesis $M \sqsubseteq N$, we know that $N_{p,q} = 1$. Since N is the principal submatrix of N' it follows that $N'_{p,q} = N_{p,q} = 1$.

Case 2. $p \leq i, q = i+1$:
Suppose that $M'_{p,i+1} = 1$. We want to show that $p \in X$ in order to conclude that $N'_{p,i+1} = 1$. By (10) we must show that $(\exists v \leq i)[M'_{v,i+1} = 1 \wedge [N_{p,v} = 1 \vee p = v]]$; taking $v = p$ does the job.

Case 3. $p = i+1, q \leq i$:
Suppose that $M'_{i+1,q} = 1$. We want to show that $q \notin X$ in order to conclude that $N'_{i+1,q} = 1$. So according to the definition of X given in (10) we must show that $\neg(\exists v \leq i)[M'_{v,i+1} = 1 \wedge [N_{q,v} = 1 \vee q = v]]$. So we must show that for all $v \leq i$, we have $M'_{v,i+1} \neq 1$ or $[N_{q,v} \neq 1 \wedge q \neq v]$. So suppose that $M'_{v,i+1} = 1$; if it were the case that $q = v$, then we would have $M'_{i+1,q}$ (by assumption) and $M'_{q,v+1}$, and by transitivity $M'_{i+1,i+1}$, which is not possible if M' represents a poset. If it were the case that $N_{q,v} = 1$, then (since we just established that $q \neq v$) it follows that $N_{v,q} = 0$, and so $M_{v,q} = 0$ and so $M'_{v,q} = 0$. On the other hand, $M'_{i+1,q} = 1$ and $M'_{v,i+1} = 1$ give us $M'_{v,q} = 1$ by transitivity, and hence a contradiction. Thus $N_{q,v} \neq 1$ as desired.

Since $M'_{i+1,i+1} = 0$, we are done with the proof that $\exists\mathbf{LA} \vdash M' \sqsubseteq N'$.

Lemma 4. $\exists\mathbf{LA} \vdash \textsc{Simple}(N')$.

Proof. According to (2) we must first show that $\textsc{Poset}(N')$, i.e., that N' is irreflexive and transitive. So following (1), we need to show that for $1 \leq p \leq i+1$, $N'_{pp} = 0$. For $p \leq i$ this follows since N is a poset, and $N'_{pp} = N_{pp}$. For $p = i+1$, $N'_{pp} = 0$ since it is not defined to be 1 by (11). To show that N' is transitive we need to show that if $N'_{pq} = 1 \wedge N'_{qr} = 1$ then $N'_{pr} = 1$. If $1 \leq p, q, r \leq i$, then this follows by the transitivity of N. Otherwise, we consider the following three cases.

Case 1. $p = i+1$:
If $N'_{pq} = 1$, it follows by (11) that $q \in [i] - X$. We also have $N'_{qr} = 1$. Suppose that $r = i+1$; by (11) it would follow that $q \in X$, which is not the case, and so $r \neq i+1$. Suppose that $r \in X$; by the definition of X, (10), we would have a $v \leq i$ such that $M'_{v,i+1} = 1 \wedge [N_{rv} = 1 \vee r = v]$. On the other hand, since we have already established that $q, r \leq i$, and $N'_{qr} = 1$, it follows that $N_{qr} = 1$, and so by transitivity of N, we have that $N_{qv} = 1$, and so $q \in X$, contradicting the first sentence of this paragraph. Thus $r \notin X$, i.e., $r \in [i] - X$, and so $N'_{pr} = 1$ by (11).

Case 2. $q = i+1$:
If $N'_{pq} = 1$, then by (11) $p \in X$ and $r \in [i] - X$. Since $p \in X$ we have a $v \leq i$ such that $N_{pv} = 1 \vee p = v$. Since N is a simple order, we have $N_{pr} = 1 \vee N_{rp} = 1 \vee p = r$. As $N'_{pr} = 1$ by assumption we know that $p \neq r$. Suppose that $N_{rp} = 1$; then $N_{rv} = 1$ by transitivity of N, and so $r \in X$, contradicting the first line of this

paragraph. The only possibility that remains is that $N_{pr} = 1$, and so it must be that $N'_{pr} = 1$.

Case 3. $r = i + 1$:
If $N'_{qr} = 1$, then by (11) $q \in X$. Since $N'_{pq} = 1$, if it were the case that $p = i+1$, then we would necessarily have that $q \in [i] - X$, giving us a contradiction. Thus $p \leq i$, and so $N_{pq} = 1$. Since $q \in X$, by (10) there exists a $v \leq i$ such that $M'_{v,i+1} \wedge [N_{qv} = 1 \vee q = v]$. Since we just showed that $N_{pq} = 1$, by transitivity of N we have that $N_{pv} = 1$, and so $p \in X$ as well. Using (11) we have that $N'_{pr} = 1$, and we are done.

Thus we have $\textsc{Poset}(N')$. It remains to show that N' is a simple order. Suppose that p, q are distinct elements; if $p, q \leq i$, then since N is a simple order, $N_{pq} = 1 \vee N_{qp} = 1$ and so $N'_{pq} = 1 \vee N'_{qp} = 1$. As they are not equal by assumption, the only other possibility is that one of them, say p, is $i+1$. There are two possibilities: $q \in X$ or $q \in [i] - X$. In the first case we have $N'_{qp} = 1$ (by (11) and in the second case we have $N'_{pq} = 1$ (also by (11)).

This gives us $\exists\mathbf{LA} \vdash \textsc{Simple}(N')$ and finishes the proof of lemma 4.

Lemma 3 and 4 give us (12), which finishes the proof of the induction step, and gives us (9). Finally, invoking Σ^B_1-induction we obtain (8), from which the main result follows:

Theorem 2. $\exists\mathbf{LA} \vdash \textsc{Szpilrajn}(M)$.

Corollary 2. $\mathsf{P/poly}$ *Frege proves Szpilrajn's Theorem.*

Proof. Following [SC04] we know that $\exists\mathbf{LA}$ over $\mathbb{Z}_2$ can be translated into $\mathsf{P/poly}$ Frege.

5 LA over rings

In this section we define **LA** in more detail, following [SC04]. We modify some definitions to suit the purpose of this paper.

The logical theory **LA** is strong enough to prove the ring properties of matrices such as $A(BC) = (AB)C, A + B = B + A$, but weak enough so that the theorems of **LA** translate into propositional tautologies with short Frege proofs. The nature of the translation depends on fixing an underlying ring. In this paper we are concerned with the ring $\mathbb{Z}$, for translating theorems of $\mathbf{LA}_{\mathrm{or}}$ (and obtaining TC^0 Frege proofs), and the ring $\mathbb{Z}_2$, for translating theorems of $\exists\mathbf{LA}$ (and obtaining $\mathsf{P/poly}$ Frege proofs).

Our theory has three sorts of object: *indices* (i.e., natural numbers), *ring elements*, and *matrices*, where the corresponding variables are denoted $i, j, k, ...$; $a, b, c, ...$; and $A, B, C, ...$, respectively. The semantic assumes that objects of type ring are from a fixed but arbitrary ring (for the purpose of this paper we are really only interested in $\mathbb{Z}_2$ and $\mathbb{Z}$), and objects of type matrix have entries from that ring.

Terms and formulas are built from the following function and predicate symbols, which together comprise the language $\mathcal{L}_{\mathbf{LA}}$:

$$\begin{aligned}&0_{\text{index}}, 1_{\text{index}}, +_{\text{index}}, *_{\text{index}}, -_{\text{index}}, \texttt{div}, \texttt{rem},\\ &0_{\text{ring}}, 1_{\text{ring}}, +_{\text{ring}}, *_{\text{ring}}, -_{\text{ring}}, {}^{-1}, \texttt{r}, \texttt{c}, \texttt{e}, \Sigma,\\ &\leq_{\text{index}}, =_{\text{index}}, =_{\text{ring}}, =_{\text{matrix}}, \text{cond}_{\text{index}}, \text{cond}_{\text{ring}}\end{aligned} \tag{13}$$

The intended meaning should be clear, except in the case of $-_{\text{index}}$, cut-off subtraction, defined as $i - j = 0$ if $i < j$. For a matrix A: $\texttt{r}(A), \texttt{c}(A)$ are the numbers of rows and columns in A, $\texttt{e}(A, i, j)$ is the ring element A_{ij} (where $A_{ij} = 0$ if $i = 0$ or $j = 0$ or $i > \texttt{r}(A)$ or $j > \texttt{c}(A)$), $\Sigma(A)$ is the sum of the elements in A. Also $\text{cond}(\alpha, t_1, t_2)$ is interpreted **if** α **then** t_1 **else** t_2, where α is a formula all of whose atomic sub-formulas have the form $m \leq n$ or $m = n$, where m, n are terms of type index, and t_1, t_2 are terms either both of type index or both of type ring. The subscripts $_{\text{index}}$, $_{\text{ring}}$, and $_{\text{matrix}}$ are usually omitted, since they ought to be clear from the context.

We use n, m for terms of type index, t, u for terms of type ring, and T, U for terms of type matrix. Terms of all three types are constructed from variables and the symbols above in the usual way, except that terms of type matrix are either variables $A, B, C, \ldots$ or λ-terms $\lambda ij\langle m, n, t\rangle$. Here i and j are variables of type index bound by the λ operator, intended to range over the rows and columns of the matrix. Also m, n are terms of type index *not* containing i, j (representing the numbers of rows and columns of the matrix) and t is a term of type ring (representing the matrix element in position (i, j)).

Atomic formulas have the forms $m \leq n, m = n, t = u, T = U$, where the three occurrences of $=$ formally have subscripts $_{\text{index}}, _{\text{ring}}, _{\text{matrix}}$, respectively. General formulas are built from atomic formulas using the propositional connectives $\neg, \vee, \wedge$ and quantifiers $\forall, \exists$.

5.1 Axioms and rules of LA

For each axiom listed below, every legal substitution of terms for free variables is an axiom of **LA**. Note that in a λ term $\lambda ij\langle m, n, t\rangle$ the variables i, j are bound. Substitution instances must respect the usual rules which prevent free variables from being caught by the binding operator λij. The bound variables i, j may be renamed to any new distinct pair of variables.

Equality Axioms These are the usual equality axioms, generalized to apply to the three-sorted theory **LA**. Here $=$ can be any of the three equality symbols, x, y, z are variables of any of the three sorts (as long as the formulas are syntactically correct). In A4, the symbol f can be any of the non-constant function symbols of **LA**. However A5 applies only to $\leq$, since this in the only predicate symbol of **LA** other than $=$.

A1 $x = x$

A2 $x = y \rightarrow y = x$

A3 $(x = y \wedge y = z) \rightarrow x = z$

A4 $x_1 = y_1, ..., x_n = y_n \rightarrow fx_1...x_n = fy_1...y_n$

A5 $i_1 = j_1, i_2 = j_2, i_1 \leq i_2 \rightarrow j_1 \leq j_2$

Axioms for indices These are the axioms that govern the behavior of index elements. The index elements are used to access the entries of matrices, and so we need to define some basic number theoretic operations.

A6 $i + 1 \neq 0$

A7 $i * (j + 1) = (i * j) + i$

A8 $i + 1 = j + 1 \rightarrow i = j$

A9 $i \leq i + j$

A10 $i + 0 = i$

A11 $i \leq j \wedge j \leq i$

A12 $i + (j + 1) = (i + j) + 1$

A13 $[i \leq j \wedge j \leq i] \rightarrow i = j$

A14 $i * 0 = 0$

A15 $[i \leq j \wedge i + k = j] \rightarrow j - i = k$

A16 $\neg(i \leq j) \rightarrow j - i = 0$

A17 $[\alpha \rightarrow \text{cond}(\alpha, i, j) = i] \wedge [\neg\alpha \rightarrow \text{cond}(\alpha, i, j) = j]$

Axioms for a ring These are the axioms that govern the behavior for ring elements; addition and multiplication, as well as additive inverses. We do not need multiplicative inverses.

A18 $0 \neq 1 \wedge a + 0 = a$

A19 $a + (-a) = 0$

A20 $1 * a = a$

A21 $a + b = b + a$

A22 $a * b = b * a$

A23 $a + (b + c) = (a + b) + c$

A24 $a * (b * c) = (a * b) * c$

A25 $a * (b + c) = a * b + a * c$

A26 $[\alpha \rightarrow \text{cond}(\alpha, a, b) = a] \wedge [\neg\alpha \rightarrow \text{cond}(\alpha, a, b) = b]$

Axioms for matrices Axiom A27 states that $\mathtt{e}(A, i, j)$ is zero when i, j are outside the size of A. Axiom A28 defines the behavior of constructed matrices. Axioms A29-A32 define the function Σ recursively by first defining it for row vectors, then column vectors ($A^t := \lambda ij\langle \mathtt{c}(A), \mathtt{r}(A), A_{ji}\rangle$), and then in general using the decomposition (14). Finally, axiom A34 takes care of empty matrices.

A27 $(i = 0 \vee \mathtt{r}(A) < i \vee j = 0 \vee \mathtt{c}(A) < j) \rightarrow \mathtt{e}(A, i, j) = 0$

A28 $\mathtt{r}(\lambda ij\langle m, n, t\rangle) = m \wedge \mathtt{c}(\lambda ij\langle m, n, t\rangle) = n \wedge [1 \leq i \wedge i \leq m \wedge 1 \leq j \wedge j \leq n]$
$\rightarrow \mathtt{e}(\lambda ij\langle m, n, t\rangle, i, j) = t$

A29 $\mathtt{r}(A) = 1, \mathtt{c}(A) = 1 \rightarrow \Sigma(A) = \mathtt{e}(A, 1, 1)$
A30 $\mathtt{r}(A) = 1 \wedge 1 < \mathtt{c}(A) \rightarrow \Sigma(A) = \Sigma(\lambda ij\langle 1, \mathtt{c}(A) - 1, A_{ij}\rangle) + A_{1\mathtt{c}(A)}$
A31 $\mathtt{c}(A) = 1 \rightarrow \Sigma(A) = \Sigma(A^t)$
A32 $1 < \mathtt{r}(A) \wedge 1 < \mathtt{c}(A) \rightarrow \Sigma(A) = \mathtt{e}(A, 1, 1) + \Sigma(\mathtt{R}(A)) + \Sigma(\mathtt{S}(A)) + \Sigma(\mathtt{M}(A))$
A33 $\mathtt{r}(A) = 0 \vee \mathtt{c}(A) = 0 \rightarrow \Sigma A = 0$

Where

$$\begin{aligned} \mathtt{R}(A) &:= \lambda ij\langle 1, \mathtt{c}(A) - 1, \mathtt{e}(A, 1, i + 1)\rangle, \\ \mathtt{S}(A) &:= \lambda ij\langle \mathtt{r}(A) - 1, 1, \mathtt{e}(A, i + 1, 1)\rangle, \\ \mathtt{M}(A) &:= \lambda ij\langle \mathtt{r}(A) - 1, \mathtt{c}(A) - 1, \mathtt{e}(A, i + 1, j + 1)\rangle. \end{aligned} \tag{14}$$

Rules for LA In addition to all the axioms just presented, **LA** has two rules: matrix equality and induction.

Matrix equality rule

From the three premises:

1. $\mathtt{e}(T, i, j) = \mathtt{e}(U, i, j)$
2. $\mathtt{r}(T) = \mathtt{r}(U)$
3. $\mathtt{c}(T) = \mathtt{c}(U)$

we conclude $T = U$.

The only restriction is that the variables i, j may not occur free in $T = U$; other than that, T and U can be arbitrary matrix terms. Our semantics implies that i and j are implicitly universally quantified in the top formula. The rule allows us to conclude $T = U$, provided that T and U have the same numbers of rows and columns, and corresponding entries are equal.

Induction rule

$$\frac{\alpha(i) \rightarrow \alpha(i + 1)}{\alpha(0) \rightarrow \alpha(n)}$$

Here $\alpha(i)$ is any formula, n is any term of type index, and $\alpha(n)$ indicates n is substituted for free occurrences of i in $\alpha(i)$. (Similarly for $\alpha(0)$.)

This completes the description of **LA**. We finish this section by observing the substitution property in the lemma below. We say that a formula S' of **LA** is a *substitution instance* of a formula S of **LA** provided that S' results by substituting terms for free variables of S. Of course each term must have the same sort as the variable it replaces, and bound variables must be renamed as appropriate.

Lemma 5. *Every substitution instance of a theorem of* **LA** *is a theorem of* **LA**.

This follows by straightforward induction on **LA** proofs. The base case follows from the fact that every substitution instance of an **LA** axiom is an **LA** axiom.

6 Conclusion and future work

We presented two different proofs of the finite version of Szpilrajn's theorem—a theorem at the foundation of concurrency theory. One proof uses concepts of a very low complexity, TC^0, and requires reasoning over ordered rings, $\mathbb{Z}$, while the second proof uses concepts of higher complexity, P, but over the simplest of rings: $\mathbb{Z}_2 = \{0, 1\}$.

A natural next step in this line of research is to formalize the reasoning in chapters 9 and 10 of [Lê08]. The ideas contained in those chapters, namely traces and comtraces, and the relational representation of generalized comtraces, are motivated by Szpilrajn's theorem, and more importantly by its proofs. It would be interesting to formalize those ideas along the lines presented in this paper. Hopefully, new algorithms of low complexity can be extracted from those proofs, using variants of the witnessing theorem.

References

[CN10] Stephen A. Cook and Phuong Nguyen. *Logical Foundations of Proof Complexity.* Cambridge Univeristy Press, 2010.

[Coo85] Stephen A. Cook. A taxonomy of problems with fast parallel algorithms. *Information and Computation*, 64(13):2–22, 1985.

[Jan08] Ryszard Janicki. Relational structures model of concurrency. *Acta Informatica*, 45(4):279–320, June 2008.

[JK93] R. Janicki and M. Koutny. Structure of concurrency. *Theoretical Computer Science*, 112:5–52, February 1993.

[Lê08] Dai Tri Man Lê. Studies in comtrace monoids. Master's thesis, McMaster University, 2008.

[Mar02] David Marker. *Model Theory: An Introduction.* Springer, 2002.

[Mon] Donald Monk. Fundamentals of the foundations of mathematics. Course notes for Math 5000 at the University of Colorado at Boulder in the Fall 2007.

[SC04] Michael Soltys and Stephen A. Cook. The complexity of derivations of matrix identities. *Annals of Pure and Applied Logic*, 130(1–3):277–323, December 2004.

[Szp30] E. Szpilrajn. Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*, 16:386–389, 1930.