Bisimilarity of Pushdown Systems is Nonelementary

Michael Benedikt¹, Stefan Göller², Stefan Kiefer¹, and Andrzej S. Murawski³

¹ University of Oxford, UK
² LIAFA Paris, France and Universität Bremen, Germany *
³ University of Leicester, UK

Abstract. Given two pushdown systems, the bisimilarity problem asks whether they are
bisimilar. While this problem is known to be decidable our main result states that it is
nonelementary, improving EXPTIME-hardness, which was the previously best known lower
bound for this problem. Our lower bound result holds for normed pushdown systems as well.

1 Introduction

A central problem in theoretical computer science is to decide whether two machines or systems
behave equivalently. While being generally undecidable for Turing machines, a lot of research has
been devoted to find subclasses of machine devices for which this problem becomes decidable.
Equivalence checking is the problem of determining whether two systems are semantically identical.

It is well-known that even language equivalence of pushdown automata is undecidable, in fact
already their universality is undecidable. On the positive side, a celebrated result due to Sénizergues
states that language equivalence of deterministic pushdown automata is decidable [13]. The best
known upper bound for the latter problem is a tower of exponentials [15] (see [4] for a more recent
proof), while only hardness of deterministic polynomial time is known to date.

Among the numerous notions of equivalence [18] in the realm of formal verification and con-
currency theory, the central one is bisimulation equivalence (bisimilarity for short), which enjoys
pleasant mathematical properties. It can be seen to take the king role: There are important char-
acterizations the bisimulation-invariant fragments of first-order logic and of monadic second-order
logic in terms of modal logic [17] and of the modal μ-calculus [5], respectively. In particular, bisim-
ilarity is a fundamental notion for process algebraic formalisms [11]. As a result, a great deal of
research in the analysis of infinite-state systems (such as pushdown systems or Petri nets) has been
devoted to deciding bisimilarity of two given processes, see e.g. [10] for a comprehensive overview.

A milestone result in this context has been proven by Sénizergues: Bisimilarity on pushdown
systems (i.e. transition systems induced by pushdown automata without ε-transitions) is decidable
[14] ; in fact, in [14] bisimilarity is proven to be decidable for the more general class of equational
graphs of finite out-degree. Since pushdown systems can be viewed as an abstraction of the call-
and-return behavior of a recursive program, the latter decidability result should be read as that
one can decide equivalence of recursive programs in terms of their visible behavior. Concerning
decidability the latter result can in some sense be considered as best possible since on the slightly
more general classes of type -1a and type -1b rewrite systems [7] and order-two pushdown graphs

Though being decidable, Sénizergues’ algorithm for deciding bisimilarity of pushdown systems
consists of two semi-decision procedures and in fact no complexity-theoretic upper bound is known
for this problem to date. On the other hand, the best known lower bound for this problem is
EXPTIME shown by Kučera and Mayr[9]. In [8] EXPTIME-hardness has been even be established
even for the subclass basic process algebras, for which a 2EXPTIME upper bound is known [3].
Such complexity gaps are typical in the context of infinite-state systems.

In fact, in case decidability is known, the precise computational complexity status of bisimilarity
on infinite-state systems is known only for few classes, let us mention basic parallel processes

* The research leading to these results has received funding from the European Union’s Seventh Frame-
work Programme (FP7/2007-2013) under grant agreement n° 259454.
(communication-free Petri nets) [6] and one-counter systems (the transition systems induced by pushdown automata over a singleton stack alphabet) [1].

Our contribution. The main result of this paper states that bisimilarity of (systems induced by) pushdown systems is nonelementary, even in the normed case. We give small descriptions of pushdown systems on which a bisimulation game is implemented that allows to push and verify encodings of nonelementarily big counters à la Stockmeyer [16]. As an important technical tool we realize deterministic verification phases in the bisimulation game by simulating non-erasing real-time transducers that are fed with the stack content. As basic gadgets, we use the well-established technique of Defender’s Forcing [7]. We are optimistic that our technique gives new insights for potential further lower bounds for bisimilarity of PA processes, regularity for pushdown systems, and weak bisimilarity of basic process algebras.

Organisation.
In Section 2 we introduce preliminaries. In Section 3 we recall basics on transductions, introduce useful abbreviations for pushdown rules and recall Defender’s forcing. Section 4 consists of our nonelementary lower bound proof for bisimilarity of pushdown systems.

2 Preliminaries
By \( \mathbb{N} \overset{\text{def}}{=} \{0, 1, \ldots \} \) we denote the set of non-negative integers. For \( n, m \in \mathbb{N} \) with we write \([n, m] \) for \{n, n + 1, \ldots , m\}; in particular note that \([n, m] = \emptyset \) if \( n > m \).

A labelled transition system (LTS) is a tuple \( S = (S, \mathcal{A},\{ \overset{a}{\rightarrow} \mid a \in \mathcal{A}\}) \), where \( S \) is a set of configurations, \( \mathcal{A} \) is a finite set of action labels, and \( \overset{a}{\rightarrow} \subseteq S \times S \) is a transition relation for each \( a \in \mathcal{A} \). We say that a state \( s \in S \) is a deadlock if there is no \( t \in S \) and no \( a \in \mathcal{A} \) such that \( s \overset{a}{\rightarrow} t \). A binary relation \( R \subseteq S \times S \) is a bisimulation if for each \( (s, s') \in R \) and each \( a \in \mathcal{A} \), we have: (1) if \( s \overset{a}{\rightarrow} t \), then there is some \( s' \overset{a}{\rightarrow} t' \) with \( (t, t') \in R \) and, conversely, (2) if \( s' \overset{a}{\rightarrow} t' \), then there is some \( s \overset{a}{\rightarrow} t \) with \( (t, t') \in R \). We write \( s \sim t \) is there is some bisimulation \( R \) with \( (s, t) \in R \).

Although not explicitly used in this paper, it is sometimes convenient to view bisimilarity as a game between Attacker and Defender. In every round of the game, there is a pebble placed on a unique state in each transition system. Attacker then chooses on one transition system and moves the pebble from the pebbled state to one of its successors by an action \( \overset{a}{\rightarrow} \), where \( a \) is some action label. Defender must imitate this by moving the pebbled state from the other system to one of its successors by the same action \( \overset{a}{\rightarrow} \). If one player cannot move, then the other player wins. Defender wins every infinite game. Two states \( s \) and \( t \) are bisimilar if and only if Defender has a winning strategy on the game with initial pebble configuration \((s, t)\).

A pushdown automaton (PDA) is a tuple \( \mathcal{P} = (Q, \Gamma, \mathcal{A}, \overset{a}{\rightarrow}) \), where \( Q \) is a finite set of control states, \( \Gamma \) is a finite set of stack symbols, \( \mathcal{A} \) is a finite set of actions, and \( \overset{a}{\rightarrow} \subseteq (Q \times \{ \varepsilon \} \times \mathcal{A} \times Q \times (Q \times \{ \varepsilon \}) \cup (Q \times \{ \varepsilon \} \times \mathcal{A} \times Q \times \Gamma) \cup (Q \times \mathcal{A} \times Q \times \{ \varepsilon \}) \) is a finite set of internal rules, push rules, and pop rules, respectively. The size of \( \mathcal{P} \) is defined as \( |\mathcal{P}| \overset{\text{def}}{=} |Q| + |\mathcal{A}| + |\overset{a}{\rightarrow}| \). We write \( qv \overset{a}{\rightarrow} q'w \) to mean \((q, v, a, q', w) \in \overset{a}{\rightarrow} \). Such a PDA \( \mathcal{P} \) induces an LTS \( S(\mathcal{P}) \overset{\text{def}}{=} (Q \times \Gamma^*, \mathcal{A}, \{ \overset{a}{\rightarrow} \mid a \in \mathcal{A}\}) \), where \( \overset{a}{\rightarrow} \overset{\text{def}}{=} \bigcup_{x \in \Gamma} \{(qx, qwx) \mid qv \overset{a}{\rightarrow} q'w\} \) for each \( a \in \mathcal{A} \). We will abbreviate each configuration \((q, w) \) in \( S(\mathcal{P}) \) by \( qw \); in particular the configuration \((q, \varepsilon) \) will be denoted by just \( q \).

Given a PDA \( \mathcal{P} = (Q, \Gamma, \mathcal{A}, \overset{a}{\rightarrow}) \), \( q_1, q_2 \in Q \) and \( w_1, w_2 \in \Gamma^* \) the PDA bisimilarity problem asks whether \( q_1w_1 \sim q_2w_2 \) holds in \( S(\mathcal{P}) \). In this paper we prove the following theorem:

**Theorem 1.** PDA bisimilarity is nonelementary.

3 Techniques

3.1 Large Counters

For each \( \ell, n \geq 0 \) we define \( \text{Tower}(\ell, n) \) inductively as \( \text{Tower}(0, n) \overset{\text{def}}{=} n \) and \( \text{Tower}(\ell + 1, n) \overset{\text{def}}{=} 2^{\text{Tower}(\ell, n)} \). Let \( \Omega_\ell \overset{\text{def}}{=} \{0, 1\} \) be alphabets whose letters have values: \( \text{val}(0_\ell) = 0 \) and \( \text{val}(1_\ell) = 1 \).
A \((0, n)\)-counter is a word from \(\Omega^n_0\). The value \(\text{val}(c)\) of a \((0, n)\)-counter \(c = \sigma_0 \cdots \sigma_{n-1}\) is defined as \(\text{val}(c) \stackrel{\text{def}}{=} \sum_{i=0}^{n-1} 2^i \cdot \text{val}(\sigma_i)\). So the set of values \(\text{val}(c)\) of \((0, n)\)-counters equals \([0, 2^n - 1]\) = \([0, \text{Tower}(1, n) - 1]\). An \((\ell, n)\)-counter with \(\ell \geq 1\) is a word \(c = c_0 \sigma_0 c_1 \cdots c_m \sigma_m\), where \(m = \text{Tower}(\ell, n) - 1\), each \(c_i\) is an \((\ell - 1, n)\)-counter with \(\text{val}(c_i) = i\), and \(\sigma_i \in \Omega_\ell\) for \(i \in [0, m]\). We define \(\text{val}(c) \stackrel{\text{def}}{=} \sum_{i=0}^{m} 2^i \cdot \text{val}(\sigma_i)\). Observe that \(\text{val}(c) \in [0, \text{Tower}(\ell + 1, n) - 1]\) and the length of each \((\ell, n)\)-counter is uniquely determined by \(\ell\) and \(n\). We call an \((\ell, n)\)-counter \(c\) zero if \(\text{val}(c) = 0\), and \(\text{ones}\) if \(\text{val}(c) = \text{Tower}(\ell + 1, n) - 1\). In the following we write \(\Omega_{\leq \ell}\) for \(\bigcup_{i=0}^{\ell} \Omega_i\). When \(n\) is clear from the context, we may speak of an \(\ell\)-counter to mean an \((\ell, n)\)-counter.

### 3.2 Transductions

A (real-time and non-erasing) transducer is a tuple \(T = (Q, q_0, \Sigma, \Upsilon, \delta)\), where \(Q\) is a finite set of states, \(q_0 \in Q\) is an initial state, \(\Sigma\) and \(\Upsilon\) are finite alphabets, and \(\delta : Q \times \Sigma \rightarrow Q \times \Upsilon^+\) is a transition function with output. We say that \(T\) is letter-to-letter if \(\delta(q, a) \in Q \times \Upsilon\) for each \(q \in Q\) and each \(a \in \Sigma\). We inductively extend \(\delta\) to the function \(\delta^*: Q \times \Sigma^* \rightarrow Q \times \Upsilon^*\) as follows: for each \(w \in \Sigma^*\) and \(a \in \Sigma\) we set \(\delta^*(q, \varepsilon) \stackrel{\text{def}}{=} (q, \varepsilon)\) and \(\delta^*(q, aw) \stackrel{\text{def}}{=} (q', w)\) if \(\delta(q, a) = (q', u)\) and on \(\delta^*(q, w) = (q', v)\). We define the transduction \(f_T : \Sigma^* \rightarrow \Upsilon^*\) of \(T\) as \(f_T(w) \stackrel{\text{def}}{=} v\), where \(\delta^*(q_0, w) = (q, \varepsilon)\) for some \(q \in Q\). A transduction \(f_T : \Sigma^* \rightarrow \Upsilon^*\) is said to be letter-to-letter if \(T\) is. We define the size of \(T\) as \(|T| \stackrel{\text{def}}{=} |Q| + |\Sigma| + |\Upsilon| + \sum |w| : q \in Q, a \in \Sigma, \delta(q, a) = (q', w)\)\).

Given two transductions \(f_1 : \Sigma_1^* \rightarrow \Upsilon^*\) and \(f_2 : \Sigma_2^* \rightarrow \Upsilon^*\) with \(\Sigma_1 \cap \Sigma_2 = \emptyset\), we define their shuffle as \(f_1 || f_2 : (\Sigma_1 \cup \Sigma_2)^* \rightarrow \Upsilon^*\) inductively for each \(w \in (\Sigma_1 \cup \Sigma_2)^*\) and each \(a \in (\Sigma_1 \cup \Sigma_2)\) as follows: \(f_1||f_2(\varepsilon) \stackrel{\text{def}}{=} \varepsilon\) and \(f_1||f_2(au) \stackrel{\text{def}}{=} f_1(a) \cdot (f_2(\varepsilon))\) if \(a \in \Sigma_i\), for each \(i \in \{1, 2\}\).

We note that from two given transducers \(T_1, T_2\) with transductions \(f_T : \Sigma_1^* \rightarrow \Upsilon^*\) and \(f_T : \Sigma_2^* \rightarrow \Upsilon^*\), one can compute in time \(O(|T_1| \cdot |T_2|)\) a transducer \(T\) such that \(f_T = f_{T_1} || f_{T_2}\).

We note that every non-erasing homomorphism is a transduction (witnessed by a single-state transducer). Having a transducer \(T\), we will often write \(T\) for \(f_T\) without risk of confusion. For \(w \in \Upsilon^*\), we denote by \(\Sigma \mapsto w\) the homomorphism \(h(a) \stackrel{\text{def}}{=} w\) for each \(a \in \Sigma\).

### 3.3 Defender's forcing and Attacker's forcing

![Fig. 1. (a) Or-gadget (Defender's forcing) (b) And-gadget (Attacker's forcing)](image)

For our reduction we will use Or-gadgets (“Defender’s forcing”) and And-gadgets (“Attacker’s forcing”) to express logical disjunction and conjunction with bisimulation. More precisely, we have the following lemma.

**Lemma 2** (see e.g. [7]). Consider the states and transitions in Figure 1 (a) or (b) as part of an LTS. The states \(s_s, s_s\) may have incoming transitions, the states \(t_t, t_t, t'_{t'}, t'_{t'}\) may have outgoing transitions (not shown). Then we have for the gadgets in Figure 1:

(a) for the Or-gadget: \(s_s \sim s_s\) if and only \(t_t \sim t_t\) or \(t'_{t'} \sim t'_{t'}\);
Let \( \Omega \) and the alphabets \( \Omega \) by \( \mathcal{S} \mathcal{T} \) where the path is deterministic, readability we use macro rules that compactly represent a set of PDA transition rules with a certain role. For the rest of the section, fix a PDA \( \mathcal{P} = (Q, \Gamma, \mathcal{A}ct, \rightarrow) \) with \( a, b \in \mathcal{A}ct \).

### 3.4 Macro Rules

We will construct a PDA with many control states and rules. In the interest of succinctness and readability we use macro rules that compactly represent a set of PDA transition rules with a certain role. For the rest of the section, fix a PDA \( \mathcal{P} = (Q, \Gamma, \mathcal{A}ct, \rightarrow) \) with \( a, b \in \mathcal{A}ct \).

**Macro rules with one state on the left-hand side.** For \( p, q \in Q, \sigma \in \Gamma \) and \( a_1 \cdots a_\ell \in \mathcal{A}ct^\ell \) with \( \ell \geq 1 \) we write

\[
p\sigma \leftarrow \rightarrow a_1 \cdots a_\ell \rightarrow q
\]

to denote that there are \( p_1, \ldots, p_\ell \in Q \) with \( p\sigma \rightarrow a_1 p_1 \rightarrow a_2 p_2 \rightarrow \cdots \rightarrow a_\ell p_\ell = q \), and there are no other rules with \( p\sigma \) on the left-hand side and no other rules involving \( p_1, \ldots, p_{\ell-1} \).

For \( p, q \in Q \) and a regular language \( L \subseteq \Gamma^* \) and a transduction \( T : \Gamma^* \rightarrow \mathcal{A}ct^* \) we write

\[
pL \xrightarrow{T} q
\]

to denote that \( \mathcal{P} \) contains control states and rules described below. These rules make sure that when \( \mathcal{P} \) is in a configuration \( py \) for \( y \in \Gamma^* \), then the shortest prefix \( w \) of \( y \) with \( w \in L \) will be popped, and \( \#T(w)\# \) will be read (where \( \# \in \mathcal{A}ct \) is a special action symbol), and the control state will be changed to \( q \); if \( y \) does not have a prefix \( w \) with \( w \in L \), then \( y \) will be popped, and \( \#T(y) \) will be output. This behaviour is the result of a product construction between the minimal deterministic finite automaton (DFA) accepting \( L \) and the transducer \( T \). More precisely, let \( A = (Q_A, q_0^A, \Gamma, F_A, \delta_A) \) be the minimal DFA that accepts \( L \), where \( Q_A \) is the finite set of states, \( q_0^A \in Q_A \) is the initial state, \( F_A \subseteq Q_A \) is the set of final states, \( \delta_A : Q_A \times \Gamma \rightarrow Q_A \) is the transition function. Assume \( T = (Q_T, q_0^T, \Gamma, \mathcal{A}ct, \delta_T) \). Then \( \mathcal{P} \) contains the control states \( Q_A \times Q_T \) and the following rules:

- \( p\# \rightarrow (q_0^A, q_0^T) \);
- \( (q^A, q^T) \# \rightarrow q \) for each \( q^A \in F_A \) and each \( q^T \in Q_T \);
- for each \( \sigma \in \Gamma \), each \( q^A \in Q_A \setminus F_A \) and each \( q^T \in Q_T \), where \( \delta_T(q^T, \sigma) = (r^T, w) \), we have the (macro) rule \( (q^A, q^T)\sigma \xrightarrow{w} (\delta_A(q^A, \sigma), r^T) \).

There are no other rules with \( p \) on the left-hand side, and no other rules involving \( Q_A \times Q_T \).

If \( pL \xrightarrow{T} q \) and \( w \in L \) but no proper prefix of \( w \) is in \( L \), then we have in the LTS \( S(\mathcal{P}) \)

\[
\text{for all } x \in \Gamma^* : \quad pwx \# \xrightarrow{s_0 a_1 \cdots a_\ell \# qx,} (1)
\]

where the path is deterministic, \( T(w) = a_1 \cdots a_\ell \), and \( s_0, \ldots, s_\ell \) are configurations of \( \mathcal{P} \), i.e., states in \( S(\mathcal{P}) \).

We will need the following lemma, which shows how to compare two counters in terms of their images of two given transducers. For the statement of the lemma, recall the concept of counters and the alphabets \( \Omega \) from Section 3.1 and recall for each alphabet \( \Omega \) and each word \( w \) we denote by \( \Omega \rightarrow w \) the homomorphism that maps every element from \( \Omega \) to \( w \).

**Lemma 3.** Let \( T_1, T_2 : \Omega_{\leq \ell+1} \rightarrow \mathcal{A}ct^* \) be letter-to-letter transducers for some \( \ell \geq 0 \). Let \( \mathcal{P} = (Q, \Gamma, \mathcal{A}ct, \rightarrow) \) be a PDA with \( \{s, p, q, r\} \subseteq Q \), \( \Omega_{\leq \ell+1} \subseteq \Gamma \) and the following macro
rules:

\[
\begin{align*}
\bullet q (\Omega^*_{\leq t-1} \cdot \Omega_t)^* \cdot \Omega_{t+1} & \xrightarrow{T_1} \bullet r \\
q_\bullet (\Omega^*_{\leq t-1} \cdot \Omega_t)^* \cdot \Omega_{t+1} & \xrightarrow{T_2} r_\bullet \\
p_\bullet \Omega^*_{\leq t} \cdot \Omega_{t+1} & \xrightarrow{\Gamma \rightarrow a} \bullet q \\
p_\bullet \Omega^*_{\leq t} \cdot \Omega_{t+1} & \xrightarrow{\Gamma \rightarrow a a} q_\bullet
\end{align*}
\]

apply \( T_1 \)

apply \( T_2 \)

pop two \( \ell \)-counters

and two \( \Omega_{t+1} \)-symbols

pop one \( \ell \)-counter

and one \( \Omega_{t+1} \)-symbol

Let \( \sigma_1, \sigma_2, \sigma_3 \in \Omega_{t+1} \), and let \( w_1, w_2, w_3 \) be \( \ell \)-counters.

(a) Assume \( x_1, x_2 \in \Gamma^* \) such that \( \bullet r x_1 \sim \bullet r x_2 \). Then

\[
\bullet q w_1 \sigma_1 x_1 \sim \bullet q w_2 \sigma_2 x_2 \iff T_1(w_1 \sigma_1) = T_2(w_2 \sigma_2).
\]

(b) Assume \( x \in \Gamma^* \) such that \( \bullet r x \sim \bullet r w_1 \sigma_1 x \). Then

\[
\bullet p w_3 \sigma_3 w_2 \sigma_2 w_1 \sigma_1 x \sim \bullet p w_3 \sigma_3 w_2 \sigma_2 w_1 \sigma_1 x \iff T_1(w_1 \sigma_1) = T_2(w_2 \sigma_2).
\]

Proof. Part (a) is immediate from the definitions. For part (b) we have:

\[
\begin{align*}
\bullet p w_3 \sigma_3 w_2 \sigma_2 w_1 \sigma_1 x & \sim \bullet p w_3 \sigma_3 w_2 \sigma_2 w_1 \sigma_1 x \\
& \iff \bullet q w_1 \sigma_1 x \sim \bullet q w_2 \sigma_2 w_1 \sigma_1 x & \text{by the first two rules} \\
& \iff T_1(w_1 \sigma_1) = T_2(w_2 \sigma_2) & \text{by part (a)}
\end{align*}
\]

Micro rules with a state pair on the left-hand side. In the following we assume that control states, i.e., the elements of \( Q \), are of the form \( \bullet q \) and \( q_\bullet \). By \( q_\bullet \) we refer to the state pair \( (\bullet q, q_\bullet) \in Q^2 \). Given \( w \in \Gamma^* \), we write \( \sim q_\bullet w \) to denote that \( \bullet q w \sim q_\bullet w \).

For \( \sigma_1 \cdots \sigma_\ell \in \Gamma_\ell \) with \( \ell \geq 0 \) we write

\[
\overrightarrow{q} \rightsquigarrow \overrightarrow{r} \sigma_1 \sigma_2 \cdots \sigma_\ell
\]

to denote that there are state pairs \( \overrightarrow{q_1}, \ldots, \overrightarrow{q_\ell} \) with \( \overrightarrow{q_\ell} = \overrightarrow{q} \) and

\[
\bullet q_\ell \xrightarrow{a} \bullet q_{\ell-1} \sigma_\ell, \ldots, \bullet q_1 \xrightarrow{a} \bullet q_0 \sigma_1, \bullet q_0 \xrightarrow{a} r_\bullet
\]

and there are no other rules with \( \bullet q \) or \( q_\bullet \) on the left-hand side and no other rules involving \( \overrightarrow{q_0}, \ldots, \overrightarrow{q_{\ell-1}} \). With this macro rule we have

\[
\text{for all } x \in \Gamma^*: \sim \overrightarrow{q} x \iff \sim \overrightarrow{r} \sigma_1 \sigma_2 \cdots \sigma_\ell x. \tag{2}
\]

For \( \sigma_1, \sigma_2 \in \Gamma \cup \{ \varepsilon \} \) we write

\[
\overrightarrow{q} \overset{\text{Def}}{\rightsquigarrow} \{ \overrightarrow{r_1 \sigma_1}, \overrightarrow{r_2 \sigma_2} \}
\]

to denote that Defender’s forcing is implemented as described in Lemma 2 (a); i.e., in terms of Figure 1 (a) we have the state correspondences \( s = \bullet q \) and \( s_\bullet = q_\bullet \) and \( t = \bullet r_1 \sigma_1 \) and \( t_\bullet = r_\bullet \sigma_1 \) and \( t_\bullet^\prime = r_\bullet^\prime \sigma_2 \) and \( t_\bullet^\prime = r_\bullet^\prime \sigma_2 \), the internal rules \( q_\bullet \overset{a}{\rightarrow} u_1, \ldots, q_\bullet \overset{a}{\rightarrow} u_3 \) and finally the push rules \( u_1 \overset{a}{\rightarrow} \bullet r_1 \sigma_1, \ldots, u_3 \overset{b}{\rightarrow} r_\bullet \sigma_2 \), as prescribed by Figure 1 (a). Intuitively, in a Defender-Attacker game, when the play is in a configuration \((\bullet q x, q_\bullet x)\) for \( x \in \Gamma^* \), then Defender chooses whether the
game will be in \((\bullet r_1 \sigma_1 x, r_1 \bullet \sigma_1 x)\) or in \((\bullet r_2 \sigma_2 x, r_2 \bullet \sigma_2 x)\). In other words, we have \(\sim \overrightarrow{q} x\) iff \(\sim \overrightarrow{r_1} \sigma_1 x\) or \(\sim \overrightarrow{r_2} \sigma_2 x\). We generalise this notation to sets: for \(\{w_1, \ldots, w_k\} \subseteq \Gamma^*\) we also write

\[
\overrightarrow{q} \overset{\text{Def}}{\rightarrow} \{\overrightarrow{r_1} w_1, \ldots, \overrightarrow{r_k} w_k\}
\]

to denote that a sequence of Or-gadgets (Figure 1 (a)) is used to achieve

\[
\text{for all } x \in \Gamma^*: \quad \sim \overrightarrow{q} x \iff \bigvee_{i=1}^{\ell} \sim \overrightarrow{r_i} w_i x .
\]

Similarly we write

\[
\overrightarrow{q} \overset{\text{Att}}{\rightarrow} \{\overrightarrow{r_1} w_1, \ldots, \overrightarrow{r_k} w_k\}
\]

to denote that And-gadgets (Attacker’s forcing, Figure 1 (b)) are used to achieve

\[
\text{for all } x \in \Gamma^*: \quad \sim \overrightarrow{q} x \iff \bigwedge_{i=1}^{\ell} \sim \overrightarrow{r_i} w_i x .
\]

4 The Construction

We prove Theorem 1 by showing that PDA bisimilarity is \(k\)-EXPSPACE-hard for all \(k \geq 1\).

As the first step of our reduction we consider a problem on letter-to-letter transducers. A transducer machine is a triple \(\mathcal{T} = (\ell, T_1, T_2)\), where \(\ell \geq 1\), and \(T_1, T_2 : \{0, 1\}^* \rightarrow \Gamma^*\) are letter-to-letter transducers. Given a transducer machine \(\mathcal{T}\) we call \(z \in \{0, 1\}^\ell\) a dead end if there is no \(z' \in \{0, 1\}^\ell\) with \(T_1(z) = T_2(z')\). We say that \(\mathcal{T}\) is deterministically terminating if there are \(t \in \mathbb{N}\) and words \(z_0, \ldots, z_t \in \{0, 1\}^\ell\) such that

- \(z_0 = 1^\ell\),
- for each \(z_i\) there is at most one \(z' \in \{0, 1\}^\ell\) with \(T_1(z_i) = T_2(z')\),
- \(T_1(z_i) = T_2(z_{i+1})\) holds for all \(i \in [0, t-1]\), and
- \(z_t\) is a dead end.

If \(\mathcal{T}\) is deterministically terminating we define \(\text{last}(\mathcal{T}) = z_t\). The first step of our reduction is applying the algorithm of the following proposition.

Proposition 4. For each \(k \geq 1\) there exists a \(k\)-EXPSPACE-complete language \(L \subseteq \Sigma^*\) such that the following is computable in polynomial time:

**INPUT:** \(w \in \Sigma^*\).

**OUTPUT:** Transducers \(T_1, T_2 : \{0, 1\}^* \rightarrow \Gamma^*\), where \(\mathcal{T} = (\text{Tower}(k, n), T_1, T_2)\) is a deterministically terminating transducer machine and \(0^{\text{Tower}(k, n)}\) is a dead end such that moreover

\[
w \in L \quad \text{if and only if} \quad \text{last}(\mathcal{T}) = 0^{\text{Tower}(k, n)} .
\]

Proof (sketch). Let us fix some \(k \geq 1\). Let us first mention how the language \(L\) can be chosen. The following claim is a simple adaption of the linear speedup theorem, we refer the reader to [12] for details.

**Claim:** For each \(k \geq 1\) there exists a deterministic Turing machine (DTM) \(\mathcal{M}\) and some \(m \in \mathbb{N}\) such that the following holds:

1. \(m\) is the sum of the number of states of \(\mathcal{M}\) plus the number of tape symbols of \(\mathcal{M}\).
2. the DTM \(\mathcal{M}\) is \(\left\lceil \frac{\text{Tower}(k, n)}{n + m} \right\rceil\) space bounded for all but finitely many \(n \in \mathbb{N}\).
3. For each \(n \geq 0\) we have
   
   (i) the DTM \(\mathcal{M}\) has a unique accepting configuration and a unique rejecting configuration that both do not have any successor configuration on each input of length \(n\) and
(ii) the unique computation of $M$ on $w$ either reaches the accepting or the rejecting configuration for each word $w$ of length $n$.

(4) The acceptance problem of $M$ is complete for $k$-EXPSPACE under polynomial time many-one reductions.

Let us fix some $k \geq 1$ and some Turing machine $M$ that satisfies points (1) to (4) of the above Claim for the rest of this proof. We define $L \overset{\text{def}}{=} L(M)$. Let us assume that the input alphabet of $M$ is $\Sigma$. Let us fix an input word $w \in \Sigma^n$. We can assume without loss of generality that $M$ is \[ \frac{\text{Tower}(k, n)}{n + m} \] - $n$ space bounded on input $w$ (the other finitely many cases can be dealt with explicitly in our reduction). Recall that by Point (3) of the above Claim we have that $M$ has a unique accepting configuration and a unique rejecting configuration, and that all computations of $M$ reach either the accepting or the rejecting configuration, and that the accepting and rejecting configurations have no successor configurations. In a first step, we modify $M$ to a DTM $M_w$ so that $M_w$ started on the empty tape first writes $w$ on the tape and then simulates $M$ on $w$. We note that we can construct $M_w$ in such a way that

- $M_w$ can be obtained from $M$ by adding at most $n$ additional states and corresponding transitions that allow us to initially copy $w$ onto the working tape,
- the sum of the number of states of $M_w$ plus the number of tape symbols of $M_w$ is $n + m$,
- $M_w$ is $\left\lfloor \frac{\text{Tower}(k, n)}{n + m} \right\rfloor - n + n = \left\lfloor \frac{\text{Tower}(k, n)}{n + m} \right\rfloor$ space bounded and
- $w \in L(M)$ if and only if $\varepsilon \in L(M_w)$.

Fix a binary encoding $\text{enc}$ of configurations so that each tape symbol and each pair consisting of a state and a tape symbol of $M_w$ can be (injectively) encoded by a binary string of length $n + m$. We extend this encoding to configurations of $M_w$ by mapping each configuration $c$ (injectively) to a string $\text{enc}(c) \in \{0, 1\}^\ell$ where $\ell = \text{Tower}(k, n)$. Moreover we assume that the initial configuration of $M_w$ (with empty tape) is encoded by $1^\ell$ and that the (unique) accepting configuration of $M_w$ is encoded by $0^\ell$. It remains to argue that one can construct transducers $T_1, T_2 : \{0, 1\}^* \rightarrow T^*$ so that

(*) for all configurations $c, c'$ of $M_w$ we have $T_1(\text{enc}(c)) = T_2(\text{enc}(c'))$ if and only if $c'$ is a successor configuration (i.e., the unique one) of $c$.

For establishing (*), the idea is to construct $T_1, T_2$ so that if $c$ is a configuration of $M_w$ then $T_1(\text{enc}(c))$ is an encoding of $c'$, where $c'$ is the successor configuration of $c$, and $T_2(\text{enc}(c))$ is an encoding of $c$. The most straightforward implementation of this idea would be to let $T_1(\text{enc}(c)) = \text{enc}(c')$ and $T_2(\text{enc}(c)) = \text{enc}(c)$. However, this cannot be easily done, if at all, loosely speaking because the read-write head of $M_w$ may move in the direction “opposite” to the transducers so that the transducer $T_1$ would have to “guess” where the read-write head is before it actually sees it. Therefore, we construct $T_1, T_2$ so that their output is “delayed” by a few steps: Transducer $T_1$ remembers in its finite control the last few bits of the encoded tape and outputs the bits of the encoding of the successor configuration only after $T_1$ can be sure about them. Transducer $T_2$ does not compute the successor configuration, but only re-encodes the encoded configuration, and outputs the bits of the new encoding in a similarly delayed way as $T_1$. Since transducers need to output a single symbol per step, the transducers $T_1, T_2$ output a dummy symbol in the first few steps. At the end they need to output a single symbol containing the last few bits of the new encoding. As a consequence the alphabet $T$ cannot be (easily) taken to be binary; thus we simply choose $T$ sufficiently large for this to work. \qed

Let us fix some $k \geq 1$ for the rest of this section and let us fix the $k$-EXPSPACE complete language $L \subseteq \Sigma^*$ that satisfies Proposition 4. Moreover let $w \in \Sigma^n$ be a word. Our overall goal is to compute from $w$ in polynomial time a PDA and two of its configurations that are bisimilar if and only if $w \in L$. As an intermediate step, let us fix for the rest of this section the output $(T_1, T_2)$ of the algorithm of Proposition 4 on input $w$, and let $T = (\text{Tower}(k, n), T_1, T_2)$. In the rest of the section we will show how to compute from $T_1, T_2$ and $n$ in time polynomial in $|T_1| + |T_2| + n$ a
PDA $P = (Q, \Omega, Act, \rightarrow)$ so that we have $\sim\vec{\text{start}}$ in $S(P)$ if and only if $\text{last}(T) = 0^{\text{Tower}(k,n)}$ holds in $T$, where $\bullet(\text{start})$ and $\langle \text{start} \rangle \bullet$ will be control states of $P$. We recall that $k$ is a fixed constant.

Let

$$B \equiv \{ \text{start}, \text{stop}_\ell, \text{testDec}_\ell, \text{testDec}_1, \text{ones}_\ell, \text{ones}_1, \text{dec}Ok_\ell, \text{zeroOk}_\ell, \text{zero}_1, \text{dec}_\ell, \text{dec}_1, \text{dec}_0^{(i)}, \text{fin}, \text{testFin}, \text{popAll}, \text{next}, \text{next}_1, \text{tran}, \text{testTran}, \text{testTran}^1 \mid 0 \leq \ell \leq k + 2, 1 \leq i \leq n \} \cup B_{\text{impl}}$$

be a set of “basic symbols” that we use to construct the control states $Q$. The set $B_{\text{impl}}$ contains further (implicit) symbols that are needed to implement macro rules. In the following we regard each element of $B$ as a single symbol (of length 1). Define $\Omega \equiv \Omega_{\leq k+1}, Act \equiv \{ 0, 1, \#, a, b \} \sqcup T$, $Q = \bullet B \cup B_\bullet$, where $\bullet B \equiv \{ \langle \alpha \rangle \mid \alpha \in B^*, 1 \leq |\alpha| \leq k + 2 \}$ and similarly $B_\bullet \equiv \{ \langle \alpha \rangle \bullet \mid \alpha \in B^*, 1 \leq |\alpha| \leq k + 2 \}$. For instance, we have $\langle \text{dec}_{k-1} \text{ones}_k \rangle \in Q$. We will use $\beta$ to indicate an arbitrary word $\alpha \in B^*$ with $1 \leq |\alpha| \leq k + 1$.

This section is organised as follows. In Section 4.1 we show how we can implement a bisimulation (sub-)game in $S(P)$ that allows us to test whether two $\ell$-counters have consecutive values for each $\ell \geq 0$. In Section 4.2 we show how we can implement a bisimulation (sub-)game in $S(P)$ that allows Defender to push an $\ell$-counter onto the stack for each $\ell \geq 0$. We conclude our reduction in Section 4.3.

### 4.1 Checking Counters for Consecutive Values

For each $\ell \in [0, k]$ we include control states $\bullet \langle \text{stop}_\ell \rangle, \langle \text{stop}_\ell \rangle \bullet$ such that for all $x \in \Omega^*$ and all $\sigma \in \Omega_{\ell+1}$ and all $\ell$-counters $w$ we have

$$\langle \text{stop}_\ell \rangle x \sim \langle \text{stop}_\ell \rangle \bullet w \sigma x$$

This is easily achieved, for instance by including no rules with $\bullet \langle \text{stop}_\ell \rangle$ or $\langle \text{stop}_\ell \rangle \bullet$ on the left-hand side.

We need to be able to verify whether two counters (at convenient positions) on the stack have consecutive values. To this end we include rules such that the following statement holds:

**Lemma 5.** Let $x \in \Omega^*$, and $\sigma_1, \sigma_2, \sigma_3 \in \Omega_{\ell+1}$, and let $w_1, w_2, w_3$ be $\ell$-counters. Then $\sim \langle \text{testDec}_\ell \rangle w_3 \sigma_3 w_2 \sigma_2 w_1 \sigma_1 x$ if and only if $\text{val}(w_1) = \text{val}(w_2) + 1$.

Let $T_{\ell}^{+0}, T_{\ell}^{+1} : (\Omega_\ell \cup \Omega_{\ell+1})^* \rightarrow \{ 0, 1, a, b \}^*$ be the transducers depicted in Figure 2. Transducers $T_{\ell}^{+0}, T_{\ell}^{+1}$ interpret the input word over $\Omega_\ell$ as a number in binary, with the least significant bit read first. Transducer $T_{\ell}^{+0}$ copies the number and outputs $a$ upon reading an $\Omega_{\ell+1}$-symbol. Transducer $T_{\ell}^{+1}$ attempts to increase the number by 1 and output $a$ upon reading an

![Fig. 2. Transducers $T_{\ell}^{+0}$ and $T_{\ell}^{+1}$](image-url)
Lemma 6. Let \( \ell \in [0, k] \), and \( x \in \Omega^* \), and \( \sigma, \tau \in \Omega_{\ell+1} \), and \( v, w \) be \( \ell \)-counters.

(a) Then \( \sim\langle \text{ones}_\ell \rangle \beta x \) iff \( \sim\langle \beta \rangle wx \), where \( w \) is the ones \( \ell \)-counter.

(b) Then \( \sim\langle \text{decOk}_\ell \rangle \beta \sigma \omega \tau x \) iff \( \text{val}(v) + 1 = \text{val}(w) \) and \( \sim\langle \beta \rangle \nu \sigma \omega \tau x \).

(c) Then \( \sim\langle \text{zero}_\ell \rangle \beta \sigma \omega \tau x \) iff \( \text{val}(w) = 1 \) and \( \sim\langle \beta \rangle \nu \sigma \omega \tau x \), where \( v \) is the \( \ell \)-counter with \( \text{val}(v) = 0 \).

(d) Then \( \sim\langle \text{dec}_\ell \rangle \beta \sigma \omega \tau x \) iff \( \text{val}(w) \neq 0 \) and \( \sim\langle \beta \rangle \nu \sigma \omega \tau x \), where \( v \) is the \( \ell \)-counter with \( \text{val}(v) + 1 = \text{val}(w) \).

The following rules for the special case \( \ell = 0 \) are included.
The following rules are included for $1 \leq \ell \leq k$.

\[
\begin{align*}
\langle \text{ones}_\ell \beta \rangle & \rightarrow \langle \text{ones}_{\ell-1} \, \text{ones}_1 \beta \rangle 1_\ell & \text{push } 1_\ell \text{ and a ones } (\ell - 1)-\text{counter} \\
\langle \text{ones}_\ell \beta \rangle & \rightarrow \left\{ \langle \text{dec}_{\ell-1} \, \text{ones}_1 \beta \rangle 1_\ell, \langle \text{zero}_{\ell-1} \beta \rangle 1_\ell \right\} & \text{push } 1_\ell \text{ and a decremented } (\ell - 1)-\text{counter} \\
\langle \text{decOk}_\ell \beta \rangle & \rightarrow \left\{ \langle \beta \rangle, \langle \text{testDec}_\ell \rangle 1_{\ell+1} \right\} & \text{push } 0_\ell \text{ and a ones } (\ell - 1)-\text{counter} \\
\langle \text{zero}_\ell \beta \rangle & \rightarrow \langle \text{ones}_{\ell-1} \, \text{zero}_1 \beta \rangle 0_\ell & \text{push } 0_\ell \text{ and a decr. } (\ell - 1)-\text{counter} \\
\langle \text{zero}_\ell \beta \rangle & \rightarrow \left\{ \langle \text{dec}_{\ell-1} \, \text{zero}_1 \beta \rangle 0_\ell, \langle \text{zero}_{\ell-1} \, \text{decOk}_\ell \beta \rangle 0_\ell \right\} & \text{push } 0_\ell \text{ and a zero } (\ell - 1)-\text{counter} \\
\langle \text{dec}_\ell \beta \rangle & \rightarrow \langle \text{ones}_{\ell-1} \, \text{dec}_1 \beta \sigma \mid \sigma \in \Omega_\ell \rangle & \text{push from } \Omega_\ell \text{ and a ones } (\ell - 1)-\text{counter} \\
\langle \text{dec}_\ell \beta \rangle & \rightarrow \langle \text{dec}_{\ell-1} \, \text{dec}_1 \beta \sigma, \langle \text{zero}_{\ell-1} \, \text{decOk}_\ell \beta \rangle \sigma \mid \sigma \in \Omega_\ell \rangle & \text{push from } \Omega_\ell \text{ and a decr. } (\ell - 1)-\text{counter} \\
\end{align*}
\]

\textit{Proof (of Lemma 6).} The proof is by induction on $\ell$.

\textit{Induction base.}

Case (a) immediately follows from the rule for $\langle \text{ones}_0 \beta \rangle$.

For part (b) we have:

\[
\forall \langle \beta \rangle \rightarrow \langle \text{decOk}_0 \beta \rangle \rightarrow \left\{ \langle \beta \rangle, \langle \text{testDec}_0 \rangle 0_0 0_1 \right\}
\]

\[
\langle \text{zero}_0 \beta \rangle \rightarrow \langle \text{decOk}_0 \beta \rangle 0_0
\]

\[
\langle \text{dec}_0 \beta \rangle \rightarrow \left\{ \langle \text{dec}_0(1) \beta \rangle 0_0, \langle \text{dec}_0(1) \beta \rangle 1_0 \right\}
\]

\[
\forall 1 \leq i < n : \langle \text{dec}_0(i) \beta \rangle \rightarrow \left\{ \langle \text{dec}_0(i+1) \beta \rangle 0_0, \langle \text{dec}_0(i+1) \beta \rangle 1_0 \right\}
\]

\[
\langle \text{dec}_0(n) \beta \rangle \rightarrow \langle \text{decOk}_0 \beta \rangle
\]

push a ones 0-counter

believe that the values of the top two 0-counters differ by 1

OR challenge that claim by invoking \text{testDec}_0

push a zero 0-counter and check if it is over a 0-counter with value 1

push the first bit of the decremented 0-counter

push the \((i + 1)^{\text{st}}\) bit of the decremented 0-counter

verify if the 0-counter has been correctly decremented

\[
\langle \text{ones}_\ell \beta \rangle \rightarrow \langle \text{ones}_{\ell-1} \, \text{ones}_1 \beta \rangle 1_\ell
\]

\[
\langle \text{ones}_\ell \beta \rangle \rightarrow \left\{ \langle \text{dec}_{\ell-1} \, \text{ones}_1 \beta \rangle 1_\ell, \langle \text{zero}_{\ell-1} \beta \rangle 1_\ell \right\}
\]

\[
\langle \text{decOk}_\ell \beta \rangle \rightarrow \left\{ \langle \beta \rangle, \langle \text{ones}_\ell \, \text{testDec}_\ell \rangle 0_{\ell+1} \right\}
\]

\[
\langle \text{zero}_\ell \beta \rangle \rightarrow \langle \text{ones}_{\ell-1} \, \text{zero}_1 \beta \rangle 0_\ell
\]

\[
\langle \text{zero}_\ell \beta \rangle \rightarrow \left\{ \langle \text{dec}_{\ell-1} \, \text{zero}_1 \beta \rangle 0_\ell, \langle \text{zero}_{\ell-1} \, \text{decOk}_\ell \beta \rangle 0_\ell \right\}
\]

\[
\langle \text{dec}_\ell \beta \rangle \rightarrow \langle \text{ones}_{\ell-1} \, \text{dec}_1 \beta \sigma \mid \sigma \in \Omega_\ell \rangle
\]

\[
\langle \text{dec}_\ell \beta \rangle \rightarrow \langle \text{dec}_{\ell-1} \, \text{dec}_1 \beta \sigma, \langle \text{zero}_{\ell-1} \, \text{decOk}_\ell \beta \rangle \sigma \mid \sigma \in \Omega_\ell \rangle
\]

OR push a \((\ell - 1)-\text{counter} \)

push a \((\ell - 1)-\text{counter} \)

push a \((\ell - 1)-\text{counter} \)

push a \((\ell - 1)-\text{counter} \)

push a \((\ell - 1)-\text{counter} \)

push a \((\ell - 1)-\text{counter} \)

OR push from \(\Omega_\ell \) and a ones \((\ell - 1)-\text{counter} \)

OR push from \(\Omega_\ell \) and a \((\ell - 1)-\text{counter} \)

OR push from \(\Omega_\ell \) and a \((\ell - 1)-\text{counter} \)
\( \sim \langle \text{dec} 0 \beta \rangle \sigma \omega \tau x \iff \sim \langle \text{testDec} \rangle 0^n_0 \sigma \omega \tau x \text{ and } \sim \langle \beta \rangle \sigma \omega \tau x \) (rule for \( \langle \text{dec} 0 \beta \rangle \))

\( \iff \text{val}(v) + 1 = \text{val}(w) \text{ and } \sim \langle \beta \rangle \sigma \omega \tau x \) (Lemma 5)

For part (c) we have:

\( \sim \langle \text{zero} 0 \beta \rangle \sigma \omega \tau x \iff \sim \langle \text{dec} 0 \beta \rangle 0^n_0 \sigma \omega \tau x \) (rule for \( \langle \text{zero} 0 \beta \rangle \) and part (b))

\( \iff \text{val}(0^n_0) + 1 = \text{val}(w) \text{ and } \sim \langle \beta \rangle \sigma \omega \tau x \) (Lemma 5)

\( \iff \text{val}(w) = 0 \text{ and } \sim \langle \beta \rangle \sigma \omega \tau x, \text{ where } v \text{ is the 0-counter with } \text{val}(v) = 0 \)

For part (d) we have:

\( \sim \langle \text{dec} \beta \rangle \sigma \omega \tau x \iff \exists v \in \Omega^n_0 : \sim \langle \text{dec} 0 \beta \rangle \sigma \omega \tau x \) (rules for \( \langle \text{dec} \beta \rangle \) and \( \langle \text{dec} 0 \beta \rangle \))

\( \iff \exists v \in \Omega^n_0 : \text{val}(v) + 1 = \text{val}(w) \text{ and } \sim \langle \beta \rangle \sigma \omega \tau x \) (part (b))

\( \iff \text{val}(w) \neq 0 \text{ and } \sim \langle \beta \rangle \sigma \omega \tau x, \text{ where } v \text{ is the } \ell \text{-counter with } \text{val}(v) + 1 = \text{val}(w) \)

**Induction step.** In the following let \( \ell \in [0, k - 1] \), let \( x \in \Omega^* \), let \( \sigma, \tau \in \Omega_{\ell+2} \) and let \( v, w \) be \((\ell + 1)\)-counters. Let \( m \overset{\text{def}}{=} \text{Tower}(n, \ell + 1) - 1 \). We write \( c_i \) for the \( \ell \)-counter with \( \text{val}(c_i) = i \) for each \( i \in [0, m] \).

For part (a) we obtain the following equivalences:

\( \sim \langle \text{ones} \beta \rangle x \iff \sim \langle \text{ones} \beta \rangle 1_\ell x \) (rule for \( \langle \text{ones} \beta \rangle \))

\( \iff \sim \langle \text{ones} \beta \rangle c_m 1_\ell x \) (ind. hyp. on (a))

\( \iff \sim \langle \text{ones} \beta \rangle 1_\ell c_m 1_\ell x \) (rule for \( \langle \text{ones} \beta \rangle \))

or \( \sim \langle \text{zero} \beta \rangle 1_\ell c_m 1_\ell x \)

\( \iff \sim \langle \text{ones} \beta \rangle c_m 1_\ell c_m 1_\ell x \) (ind. hyp. on (d),(c),\( m > 1 \))

or \( \sim \langle \beta \rangle 1_\ell c_m 1_\ell x \) (ind. hyp. on (c))

\( \iff \ldots \)

\( \iff \sim \langle \text{ones} \beta \rangle 1_\ell c_1 \cdots c_m 1_\ell x \) (ind. hyp. on (d))

or \( \sim \langle \beta \rangle 1_\ell c_1 \cdots c_m 1_\ell x \) (ind. hyp. on (c))

Further we have

\( \sim \langle \text{ones} \beta \rangle c_0 1_\ell c_1 \cdots c_m 1_\ell x \)

\( \iff \sim \langle \text{dec} \beta \rangle 1_\ell c_0 1_\ell c_1 \cdots c_m 1_\ell x \) (rule for \( \langle \text{ones} \beta \rangle \))

or \( \sim \langle \text{zero} \beta \rangle 1_\ell c_0 1_\ell c_1 \cdots c_m 1_\ell x \)

\( \iff \text{false} \) (ind. hyp. on (d),(c))
Combining this with the equivalences above yields part (a) for \( \ell + 1 \).

For part (b) we obtain the following equivalences:

\[
\sim (\text{decOk}_{\ell+1})_\beta \sigma wtx \iff \sim (\beta) v\sigma wtx
\]

and \( \sim (\text{ones}_{\ell+1})_{\text{testDec}_{\ell+1}} 0_{\ell+2} v\sigma wtx \) (rule for \( (\text{decOk}_{\ell+1})_\beta \))

\[
\iff \sim (\beta) v\sigma wtx \text{ and } \sim (\text{testDec}_{\ell+1}) w' 0_{\ell+2} v\sigma wtx
\]

where \( w' \) is the ones \((\ell + 1)\)-counter (part (a))

\[
\iff \sim (\beta) v\sigma wtx \text{ and } \text{val}(w) = \text{val}(v) + 1 \quad \text{(Lemma 5)}
\]

For part (c) we obtain the following equivalences:

\[
\sim (\text{zero}_{\ell+1})_\beta \sigma wtx
\]

\[
\iff \sim (\text{ones}_{\ell+1})_\beta 0_{\ell+1} \sigma wtx \quad \text{(rule for } (\text{zero}_{\ell+1})_\beta )
\]

\[
\iff \sim (\text{zero}_{\ell+1})_\beta c_m 0_{\ell+1} \sigma wtx \quad \text{(ind. hyp. on (a))}
\]

\[
\iff \sim (\text{dec}, \text{zero}_{\ell+1})_\beta 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx
\]

\[
\text{or } \sim (\text{zero}, \text{decOk}_{\ell+1})_\beta 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx \quad \text{(rule for } (\text{zero}_{\ell+1})_\beta )
\]

\[
\iff \sim (\text{zero}_{\ell+1})_\beta c_{m-1} 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx \quad \text{(ind. hyp. on (d),(c) and } m > 1)
\]

\[
\iff \ldots
\]

\[
\iff \sim (\text{dec}, \text{zero}_{\ell+1})_\beta 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx
\]

\[
\text{or } \sim (\text{zero}, \text{decOk}_{\ell+1})_\beta 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx
\]

\[
\iff \sim (\text{zero}_{\ell+1})_\beta c_0 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx \quad \text{(ind. hyp. on (d),(c))}
\]

\[
\iff \sim (\text{zero}_{\ell+1})_\beta c_0 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx
\]

\[
\text{or } (\sim (\beta) c_0 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx \text{ and } \text{val}(w) = 1 \text{ (part (b))}
\]

Further we have

\[
\sim (\text{zero}_{\ell+1})_\beta c_0 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx
\]

\[
\iff \sim (\text{dec}, \text{zero}_{\ell+1})_\beta 0_{\ell+1} c_0 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx
\]

\[
\text{or } \sim (\text{zero}, \text{decOk}_{\ell+1})_\beta 0_{\ell+1} c_0 0_{\ell+1} c_1 \cdots 0_{\ell+1} c_m 0_{\ell+1} \sigma wtx \quad \text{(rule for } (\text{zero}_{\ell+1})_\beta )
\]

\[
\iff \text{false} \quad \text{(ind. hyp. on (d),(c))}
\]

Combining this with the equivalences above yields part (c) for \( \ell + 1 \).
For part (d) we obtain the following equivalences:

\[ \sim(\text{dec}_{t+1} \beta) \sigma w \tau x \iff \exists m \in \Omega_{t+1} : \sim(\text{ones}_{t+1} \beta) \sigma m \sigma w \tau x \quad \text{(rule for } \sim(\text{dec}_{t+1} \beta)) \]

\[ \exists m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c m \sigma m \sigma w \tau x \quad \text{(ind. hyp. on (a))} \]

\[ \exists m, \sigma m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_{m-1} \sigma m \sigma w \tau x \quad \text{(rule for } \sim(\text{dec}_{t+1} \beta)) \]

\[ \exists \sigma m, \sigma m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_{m-1} \sigma m \sigma w \tau x \quad \text{(ind. hyp. on (d), (c), } m > 1) \]

\[ \ldots \]

\[ \exists \sigma_1, \ldots, \sigma m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \]

\[ \exists \sigma_0, \ldots, \sigma m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \quad \text{(rule for } \sim(\text{dec}_{t+1} \beta)) \]

\[ \exists \sigma_0, \ldots, \sigma m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \quad \text{(ind. hyp. on (c))} \]

\[ \exists \sigma_0, \ldots, \sigma m \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \quad \text{(rule for } \sim(\text{dec}_{t+1} \beta)) \]

\[ \text{and } \text{val}(w) = \text{val}(c_0 \sigma_0 \cdots c m \sigma m) + 1 \quad \text{(}\star\text{)} \]

\[ \text{(ind. hyp. on (b))} \]

Further we have for all \( \sigma_0, \ldots, \sigma m \in \Omega_{t+1} : \)

\[ \sim(\text{dec}_{t+1} \beta) c_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \]

\[ \sim(\text{dec}_{t+1} \beta) c_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \quad \text{(ind. hyp. on (d))} \]

\[ \exists \sigma_{-1} \in \Omega_{t+1} : \sim(\text{dec}_{t+1} \beta) c_{-1} c_0 \sigma_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \quad \text{(ind. hyp. on (d), (c))} \]

\[ \text{false} \]

Hence \((\star)\) is equivalent to

\[ \exists \sigma_0, \ldots, \sigma m \in \Omega_{t+1} : \sim(\beta) c_0 \sigma_0 \sigma_1 \sigma_1 \cdots c m \sigma m \sigma w \tau x \quad \text{and } \text{val}(w) = \text{val}(c_0 \sigma_0 \cdots c m \sigma m) + 1 \]

which is in turn equivalent to

\[ \text{val}(w) \neq 0 \quad \text{and } \sim(\beta) v \sigma w \tau x, \text{ where } v \text{ is the } \ell\text{-counter with } \text{val}(v) + 1 = \text{val}(w), \]

which shows part (d) for \( \ell + 1. \) \( \square \)
4.3 Simulating a Transducer Machine

Returning to our overall reduction, let us recall we have fixed a deterministically terminating transducer machine \( \mathcal{T} = (\text{Tower}(k,n), T_1, T_2) \) with respect to which \( 0^{\text{Tower}(k,n)} \) is a dead end, i.e. we have the following in total:

- \( T_1, T_2 : \{0,1\}^* \rightarrow Y^* \) are letter-to-letter transducers,
- for each \( z \in \{0,1\}^{\text{Tower}(k,n)} \) there is at most one \( z' \) with \( T_1(z) = T_2(z') \),
- assume \( z_0, \ldots, z_t \in \{0,1\}^{\text{Tower}(k,n)} \) such that \( z_0 = 1^{\text{Tower}(k,n)} \) and \( T_1(z_i) = T_2(z_{i+1}) \) for each \( i \in [0, t-1] \), and \( z_t \) is a dead end (we defined \( \text{last}(\mathcal{T}) \equiv z_t \)), and
- \( 0^{\text{Tower}(k,n)} \) is a dead end with respect to \( \mathcal{T} \).

We include rules so that \( \text{match} \) holds if and only if \( \text{last}(\mathcal{T}) = 0^{\text{Tower}(k,n)} \), thus completing our reduction. The PDA \( \mathcal{P} \) will be able to push \( z_0, z_1, \ldots \) on the stack, where each word \( z_i \) is encoded as a \( k \)-counter, say \( d_i \), in the obvious way: \( z_i = \eta(d_i) \) where \( \eta : \Omega^* \rightarrow \Omega_k^* \) denotes the homomorphism with \( \eta(\sigma) = \sigma \) for \( \sigma \in \Omega_k \) and \( \eta(\sigma) = \varepsilon \) otherwise. We emphasize that in comparison to the counters that were present in the proof of Lemma 6 for each \( i \in [0, t] \) we do not generally have \( \text{val}(d_i) = i \); instead we have \( z_i = \eta(d_i) \), in particular the sequence \( z_0, \ldots, z_t \) and thus the sequence \( d_0, \ldots, d_t \) is determined. The \( d_i \) will be separated on the stack by the symbol \$ \equiv 0_{k+1} \). We include rules such that the following holds:

**Lemma 7.** Let \( x \in \Omega^* \) and let \( w_1, w_2, w_3 \) be \( k \)-counters.

(a) Then \( \text{match}^{\text{testFin}} w_1 x \) iff \( \eta(w_1) = 0^{\text{Tower}(k,n)} \).
(b) Then \( \text{match}^{\text{testTran}} w_3 w_2 x \) iff \( T_1(\eta(w_1)) = T_2(\eta(w_2)) \).
(c) Then \( \text{match}^{\text{start}} \) iff \( z_t = 0^{\text{Tower}(k,n)} \).

For Lemma 7 we include the following rules:

\[
\begin{align*}
\text{(start)} & \Rightarrow \text{(ones}_k\text{fin)}\, \$ \\
\text{(fin)} & \overset{\text{Def}}{\Rightarrow} \{\text{(testFin)}, \text{(next)}\, \$\}
\end{align*}
\]

\[
\begin{align*}
\text{(testFin)} & \left(\Omega_{<k-1}^* \Omega_k^*\right) \$ \overset{\{1_1\} \rightarrow b \quad \Omega(\{1\}) \rightarrow a}{\Rightarrow} \text{(popAll)}.
\text{(testFin)} & \left(\Omega_{<k-1}^* \Omega_k^*\right) \$ \overset{\Omega \rightarrow a}{\Rightarrow} \text{(popAll)}.
\text{(popAll)} & \omega \overset{a}{\Rightarrow} \text{(popAll)}.
\text{(popAll)} & \omega \overset{a}{\Rightarrow} \text{(popAll)}.
\text{(next)} & \overset{\text{Def}}{\Rightarrow} \left\{\text{ones}_{k-1} \text{next}^1 \sigma \mid \sigma \in \Omega_k\right\}
\end{align*}
\]

\[
\begin{align*}
\text{(next)} & \overset{\text{Def}}{\Rightarrow} \left\{\text{dec}_{k-1} \text{next}^1 \sigma \right\},
\text{(zero}_{k-1} \text{tran} \text{sigma} \mid \sigma \in \Omega_k\right\}
\end{align*}
\]

\[
\begin{align*}
\text{(tran)} & \overset{\text{Alt}}{\Rightarrow} \left\{\text{ones}_{k} \text{testTran} \text{sigma} \right\},
\text{(fin)} & \overset{\text{Alt}}{\Rightarrow} \left\{\text{ones}_{k} \text{testTran} \text{sigma} \right\},
\end{align*}
\]

\[
\begin{align*}
\text{(testTran)} & \Omega_{<k}^* \Omega_{<k}^* \$ \overset{\Omega \rightarrow a}{\Rightarrow} \text{(testTran)}.
\text{(testTran)} & \Omega_{<k}^* \$ \overset{\Omega \rightarrow a a}{\Rightarrow} \text{(testTran)}.
\text{(testTran)} & \left(\Omega_{<k-1}^* \Omega_k^*\right) \$ \overset{T_1 \parallel ((\Omega \Omega_k) \rightarrow a)}{\Rightarrow} \text{(stop}_k\text{)}.
\text{(testTran)} & \left(\Omega_{<k-1}^* \Omega_k^*\right) \$ \overset{T_2 \parallel ((\Omega \Omega_k) \rightarrow a)}{\Rightarrow} \text{(stop}_k\text{)}.
\end{align*}
\]

push \$ and the encoded \( z_0 \)

rule à la Lemma 3 (a) to test \( 0^{\text{Tower}(k,n)} \)

rule à la Lemma 3 (a) to test \( 0^{\text{Tower}(k,n)} \)

for all \( \omega \in \Omega \): Erase stack content

for all \( \omega \in \Omega \): Erase stack content

choose the next \( z_i \)

choose the next \( z_i \)

rule à la Lemma 3

rule à la Lemma 3

rule à la Lemma 3 for \( T_1 \)

rule à la Lemma 3 for \( T_2 \)
Proof (of Lemma 7). Parts (a) and (b) are simple consequences of the above rules and Lemma 3 (a) and (b), respectively. For part (c) we have:

\[ \sim \langle \text{start} \rangle \iff \sim \langle \text{fin} \rangle d_0 \quad \text{rule for } \sim \langle \text{start} \rangle \text{ and Lemma 6 (a)} \]

\[ \iff \eta(d_0) = 0^{\text{Tower}(k,n)} \text{ or } \sim \langle \text{next} \rangle d_0 \quad \text{rule for } \sim \langle \text{fin} \rangle \text{ and part (a)} \]

By the rules for \( \sim \langle \text{next} \rangle \) and \( \sim \langle \text{next}^1 \rangle \) and reasoning as in Lemma 6 we have

\[ \sim \langle \text{next} \rangle d_0 \iff \text{there is a } k\text{-counter } w_1 \text{ with } \sim \langle \text{tran} \rangle w_1 d_0. \]

By the rules for \( \sim \langle \text{tran} \rangle \) and part (b) we have

\[ \sim \langle \text{tran} \rangle w_1 d_0 \iff w_1 = d_1 \text{ and } \sim \langle \text{fin} \rangle w_1 d_0. \]

It follows that:

\[ \sim \langle \text{next} \rangle d_0 \iff \sim \langle \text{fin} \rangle d_1 d_0 \iff \eta(d_1) = 0^{\text{Tower}(k,n)} \text{ or } \sim \langle \text{next} \rangle d_1 d_0 \quad \text{rule for } \sim \langle \text{fin} \rangle \text{ and part (a)} \]

Combining this with the equivalences above we obtain:

\[ \sim \langle \text{start} \rangle \iff \bigvee_{i=0}^{1} \eta(d_i) = 0^{\text{Tower}(k,n)} \text{ or } \sim \langle \text{next} \rangle d_i d_0 \]

By iterating this reasoning, we obtain:

\[ \sim \langle \text{start} \rangle \iff \bigvee_{i=0}^{t} \eta(d_i) = 0^{\text{Tower}(k,n)} \text{ or } \sim \langle \text{next} \rangle d_i d_{i-1} \cdots d_0 \]

\[ \iff \bigvee_{i=0}^{t} z_i = 0^{\text{Tower}(k,n)} \text{ or } \sim \langle \text{next} \rangle d_i d_{i-1} \cdots d_0 \]

\[ \iff z_t = 0^{\text{Tower}(k,n)} \text{ or } \sim \langle \text{next} \rangle d_i d_{i-1} \cdots d_0, \]

where the last equivalence follows from the fact that \( 0^{\text{Tower}(k,n)} \) is a dead end. By reasoning as above we have:

\[ \sim \langle \text{next} \rangle d_i \cdots d_0 \implies \text{there is a } k\text{-counter } w \text{ with } \sim \langle \text{tran} \rangle w d_i \cdots d_0 \implies T_1(\eta(d_i)) = T_2(\eta(w)), \]

which is false, by definition of \( t \). We conclude that \( \sim \langle \text{start} \rangle \) holds iff \( z_t = 0^{\text{Tower}(k,n)} \).

By Lemma 7 (c) we have completed the reduction, and hence the proof of Theorem 1.

4.4 Normedness

We can strengthen Theorem 1. Given a PDA \( \mathcal{P} \) with control state set \( Q \), we say a state \( s \) of \( \mathcal{S}(\mathcal{P}) \) is normed if every state \( t \) that is reachable from \( s \) can reach some deadlock \( t' \), where \( t' = q \) for some \( q \in Q \) (i.e. the stack is empty).

**Theorem 8.** PDA bisimilarity for is nonelementary, even when the initial states are normed.
Proof. Recall that (5) requires for $x \in \Omega^*$ and $\sigma \in \Omega_{\ell+1}$ and $\ell$-counters $w$ that we have

\[ \ast (\text{stop}_\ell)x \sim (\text{stop}_\ell)w_\sigma x. \]

We satisfied this by not giving any rules for $\ast (\text{stop}_\ell)$ and $(\text{stop}_\ell)\ast$, thus creating deadlocks. Thus we allowed to reach states that are not normed. We show that the construction can be amended to avoid the latter. Assume that $x = w'\sigma'x'$ holds for an $\ell$-counter $w'$, $\sigma' \in \Omega_{\ell+1}$ and $x' \in \Omega^*$. Add the following rules:

\[ 
\ast (\text{stop}_\ell)\Omega^*_{\leq \ell}\Omega_{\ell+1} \xrightarrow{\Omega \rightarrow a_\ast} (\text{stop}_\ell)\ast \Omega^*_{\leq \ell}\Omega_{\ell+1} \\
\ast (\text{stop}_\ell)\ast \Omega^*_{\leq \ell}\Omega_{\ell+1} \xrightarrow{\Omega \rightarrow a} \ast (\text{popAll})
\]

and recall that rules for $\ast (\text{popAll})$ and $(\text{popAll})\ast$ were given previously. Intuitively, $w'\sigma'$ is popped off the left stack with “half” speed, while $w_\sigma w'\sigma'$ is popped off the right stack with “full” speed. Afterwards the stacks are of equal height and can be fully erased. Now (5) is satisfied, and these states are normed. For the latter fact, observe all reachable stacks are of the form $((\Omega^*_{\ell}\Omega_{\ell+1})^*\Omega_{\ell+1})^*$ and $(\text{popAll})\ast$ is always reachable. The only remaining problem is to guarantee that, whenever we require $\ast (\text{stop}_\ell)x \sim (\text{stop}_\ell)\ast w_\sigma x$, $x$ is indeed prefixed by $w'\sigma'$. Note that $\text{stop}_k$ is introduced in the rule for $\text{testTran}$ and other occurrences of $\text{stop}_\ell$ are used in the counter-related rules based on $\text{testDec}_\ell$. Consequently, if the above requirement is not already satisfied, $w$ must be the ones $\ell$-counter. To prevent this from happening, for $\text{testTran}$ we shall add an extra ones $k$-counter at the beginning of the simulations. For $\text{testDec}$, we shall eliminate the need for decrement tests involving ones counters by introducing a new symbol $\ast \text{ones}_\ell$ that will correspond to the predecessor of a ones counter.

This can be achieved by the following modifications:

- Replace $(\text{start}) \mapsto \langle \text{ones}_k \text{ fin} \rangle \delta$ with $(\text{start}) \mapsto \langle \text{ones}_k \text{ start} \rangle \delta$ and $(\text{start}) \mapsto \langle \text{ones}_k \text{ fin} \rangle \delta$; i.e., push one additional ones $k$-counter and $\delta$ in the beginning.
- Extend $B$ by symbols $\ast \text{ones}_\ell$ with $\ell \in [0, k + 2]$, whose role is to push an $\ell$-counter $w$ with $\text{val}(w) = \text{Tower}(\ell + 1, n) - 2$. (This value is one less than the value of a ones $\ell$-counter.)
- Replace $(\text{ones}_\ell \beta) \mapsto \langle \text{ones}_{\ell - 1} \text{ ones}_\ell \beta \rangle \ast$ with

\[
\langle \text{ones}_{\ell - 1} \text{ ones}_\ell \beta \rangle \mapsto \langle \text{ones}_{\ell - 1} \text{ ones}_\ell^0 \beta \rangle \ast \text{ and add }
\langle \text{ones}_\ell^0 \beta \rangle \mapsto \langle \text{ones}_\ell \ast \rangle \ast.
\]

Similarly, replace $(\text{dec}_\ell \beta) \mapsto \langle \text{ones}_{\ell - 1} \text{ dec}_\ell \beta \rangle \beta$ with

\[
\langle \text{ones}_{\ell - 1} \text{ dec}_\ell \beta \rangle \beta \mapsto \langle \text{ones}_{\ell - 1} \text{ dec}_\ell^0 \beta \rangle \beta \ast \text{ and add }
\langle \text{ones}_\ell^0 \beta \rangle \mapsto \langle \text{ones}_\ell \ast \rangle \ast.
\]

Similarly, replace $(\text{zero}_\ell \beta) \mapsto \langle \text{ones}_{\ell - 1} \text{ zero}_\ell \beta \rangle 0$ with

\[
\langle \text{ones}_{\ell - 1} \text{ zero}_\ell \beta \rangle 0 \mapsto \langle \text{ones}_{\ell - 1} \text{ zero}_\ell^0 \beta \rangle 0 \ast \text{ and add }
\langle \text{ones}_\ell^0 \beta \rangle \mapsto \langle \text{ones}_\ell \ast \rangle \ast.
\]

- For $\ell \geq 1$ add the following rules:

\[
\langle \text{ones}_\ell \ast \rangle \mapsto \langle \text{ones}_{\ell - 1} \ast \text{ ones}_\ell \beta \rangle \ast
\]

\[
\langle \text{ones}_\ell^0 \beta \rangle \mapsto \langle \text{ones}_{\ell - 1} \ast \text{ ones}_\ell^0 \beta \rangle \ast
\]

\[
\langle \text{ones}_\ell \beta \rangle \mapsto \langle \text{ones}_{\ell - 1} \ast \text{ ones}_\ell \beta \rangle \ast
\]

Finally, add \( \langle \text{ones}_0 \beta \rangle \mapsto \langle \beta \rangle 0_{\ast} \ast_{\ast} \).
By these modifications, whenever we require \( \cdot \langle \text{stop}_\ell \rangle x \sim \langle \text{stop}_\ell \rangle \cdot w \sigma x \) in our proofs, the word \( x \) will start with an \( \ell \)-counter and an \( \Omega_{\ell+1} \)-symbol.

\[ \square \]

References