# A polynomial procedure for trimming visibly pushdown automata 

Mathieu Caralp, Pierre-Alain Reynier, Jean-Marc Talbot

## To cite this version:

Mathieu Caralp, Pierre-Alain Reynier, Jean-Marc Talbot. A polynomial procedure for trimming visibly pushdown automata. 2011. hal-00606778v1

## HAL Id: hal-00606778 <br> https://hal.science/hal-00606778v1

Submitted on 7 Jul 2011 (v1), last revised 29 Mar 2012 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A polynomial procedure for trimming visibly pushdown automata 

Mathieu Caralp, Pierre-Alain Reynier, and Jean-Marc Talbot<br>Laboratoire d'Informatique Fondamentale de Marseille<br>Aix-Marseille Université \& CNRS, France


#### Abstract

We describe a polynomial procedure which, given a visibly pushdown automaton that accepts only well-nested words, returns an equivalent visibly pushdown automaton that is trimmed.


## 1 Introduction

Visibly pushdown automata (VPA) are a particular class of pushdown automata working over an alphabet splitted into call, internal and return symbols [1,2]. In VPA's, the stack behaviour is imposed by the input word: on a call symbol, the VPA pushes a symbol onto the stack, on a return symbol, it must pop the top symbol of the stack, and on an internal symbol, the stack remains unchanged.

Trimming a finite state automaton amounts to remove useless states, that is states that do not occur in some accepting computation of the automaton. This can be done easily in linear time simply by solving two reachability problems in the graph representing the automaton. However, the problem is more difficult for VPA's as the current state of a computation (called a configuration) is given by both a "control" state and a stack content.

To the best of our knowledge, the only trimming procedure for VPA can be deduced from the one on pushdown automata given in [3] yielding an exponential algorithm.

## 2 Definitions

Words and nested words Let $\Sigma$ be a finite alphabet partitioned into three disjoint sets $\Sigma_{c}, \Sigma_{r}$ and $\Sigma_{\iota}$, denoting respectively the call, return and internal alphabets. We denote by $\Sigma^{*}$ the set of (finite) words over $\Sigma$ and by $\epsilon$ the empty word. The length of a word $u$ is denoted by $|u|$. The set of well-nested words $\Sigma_{\mathrm{wn}}^{*}$ is the smallest subset of $\Sigma^{*}$ such that $\Sigma_{\iota}^{*} \subseteq \Sigma_{\mathrm{wn}}^{*}$ and for all $c \in \Sigma_{c}$, all $r \in \Sigma_{r}$, all $u, v \in \Sigma_{\mathrm{wn}}^{*}$, cur $\in \Sigma_{\mathrm{wn}}^{*}$ and $u v \in \Sigma_{\mathrm{wn}}^{*}$. We define the height $h(u)$ of some well-nested word $u$ by induction as follows: $h(u)=0$ if $u \in \Sigma_{\iota}^{*}, h(u v)=\max (h(u), h(v))$ for $u, v \in \Sigma_{\mathrm{wn}}^{*}$, and $h($ cur $)=1+h(u)$. Given a familly of words $w_{1}, w_{2}, \ldots, w_{n}$, we denote by $\Pi_{i=1}^{n} w_{i}$ the concatenation $w_{1} w_{2} \ldots w_{n}$.

Visibly pushdown automata (VPA) Visibly pushdown automata are a restriction of pushdown automata in which the stack behaviour is imposed by the input word. On a call symbol, the VPA pushes a symbol onto the stack, on a return symbol, it must pop the top symbol of the stack, and on an internal symbol, the stack remains unchanged. Formally:

Definition 1 (Visibly pushdown automata). A visibly pushdown automaton (VPA) on finite words over $\Sigma$ is a tuple $A=(Q, I, F, \Gamma, \delta)$ where $Q$ is a finite set of states, $I \subseteq Q$ is the set of initial states, $F \subseteq Q$ the set of final states, $\Gamma$ is a finite stack alphabet, $\delta=\delta_{c} \uplus \delta_{r} \uplus \delta_{\iota}$ the (finite) transition relation, with $\delta_{c} \subseteq Q \times \Sigma_{c} \times \Gamma \times Q$, $\delta_{r} \subseteq Q \times \Sigma_{r} \times \Gamma \times Q$, and $\delta_{\iota} \subseteq Q \times \Sigma_{\iota} \times Q$.

A configuration of a VPA is a pair $(q, \sigma) \in Q \times \Gamma^{*}$. A run of $A$ on a word $u=$ $a_{1} \ldots a_{l} \in \Sigma^{*}$ from a configuration $(q, \sigma)$ to a configuration $\left(q^{\prime}, \sigma^{\prime}\right)$ is a finite sequence of configurations $\rho=\left\{\left(q_{k}, \sigma_{k}\right)\right\}_{0 \leq k \leq l}$ such that $q_{0}=q, \sigma_{0}=\sigma, q_{l}=q^{\prime}, \sigma_{l}=\sigma^{\prime}$ and for each $1 \leq k \leq l$, there exists $\gamma_{k} \in \Gamma$ such that either $\left(q_{k-1}, a_{k}, \gamma_{k}, q_{k}\right) \in \delta_{c}$ and $\sigma_{k}=\sigma_{k-1} \gamma_{k}$ or $\left(q_{k-1}, a_{k}, \gamma_{k}, q_{k}\right) \in \delta_{r}$ and $\sigma_{k-1}=\sigma_{k} \gamma_{k}$, or $\left(q_{k-1}, a_{k}, q_{k}\right) \in \delta_{\iota}$ and $\sigma_{k}=\sigma_{k-1}$. We write $(q, \sigma) \xrightarrow{u}\left(q^{\prime}, \sigma^{\prime}\right)$ when there exists a run on $u$ from $(q, \sigma)$ to $\left(q^{\prime}, \sigma^{\prime}\right)$. We may omit the superscript $u$ when irrelevant. We denote by $\perp$ the empty word on $\Gamma$.

Initial (resp. final) configurations are configurations of the form $(q, \perp)$, with $q \in I$ (resp. $q \in F$ ). A configuration $(q, \sigma)$ is accessible (resp. is co-accessible) if there exist $u \in \Sigma^{*}$ and a configuration $c$ such that $c$ is initial and $c \xrightarrow{u}(q, \sigma)$ (resp. such that $c$ is final and $(q, \sigma) \xrightarrow{u} c)$.

Definition 2. An automaton $A$ is trimmed if every configuration of $A$ is accessible iff it is co-accessible.

We say that a run is accepting if it starts in an initial configuration and ends in a final configuration. A word is accepted by $A$ iff there exists an accepting run of $A$ on this word. The language of $A$, denoted by $L(A)$, it the set of words accepted by $A$. Note that we require here to end up with an empty stack, while the definition of [1] considers acceptance by final states only. This restriction implies that all accepted words are wellnested.

## 3 Trimming VPA

Let $A=(Q, I, F, \Gamma, \delta)$ be a VPA on the structured alphabet $\Sigma$. In this section, we define a new VPA $A^{\prime}=\left(Q^{\prime}, I^{\prime}, F^{\prime}, \Gamma^{\prime}, \delta^{\prime}\right)$ on $\Sigma$, denoted trim $(A)$, which recognizes the same language, and in addition is trimmed.

First, we define the following set:

$$
\mathrm{WN}=\left\{(p, q, r) \in Q^{3} \mid \exists(p, \perp) \rightarrow^{*}(q, \perp) \rightarrow^{*}(r, \perp)\right\}
$$

This set can be computed in polynomial time. More precisely, the following set can be computed first: $\mathrm{WN}_{2}=\left\{(p, q) \in Q^{2} \mid \exists(p, \perp) \rightarrow^{*}(q, \perp)\right\}$.
$W \mathrm{~N}_{2}$ can be defined as the least set such that
$-\{(q, q) \mid q \in Q\} \subseteq \mathrm{WN}_{2}$,

- if $(p, r) \in \mathrm{WN}_{2}$ and $(r, q) \in \mathrm{WN}_{2}$, then $(p, q) \in \mathrm{WN}_{2}$
- if $(p, q) \in \mathrm{WN}_{2}$, and $\exists\left(q, i, q^{\prime}\right) \in \delta_{\iota}$, then $\left(p, q^{\prime}\right) \in \mathrm{WN}_{2}$
- if $(p, q) \in \mathrm{WN}_{2}$ and $\exists\left(p^{\prime}, c, \gamma, p\right) \in \delta_{c},\left(q, r, \gamma, q^{\prime}\right) \in \delta_{r}$, then $\left(p^{\prime}, q^{\prime}\right) \in \mathrm{WN}_{2}$

Then, WN is obtained from $\mathrm{WN}_{2}$ by the following property:

$$
(p, q, r) \in \mathrm{WN} \Longleftrightarrow(p, q) \in \mathrm{WN}_{2} \wedge(q, r) \in \mathrm{WN}_{2}
$$

We now define the four first components of the trimmed VPA $A^{\prime}$ as follows:

- $Q^{\prime}=\mathrm{WN}$
- $I^{\prime}=\mathrm{WN} \cap(I \times I \times F)$
- $F^{\prime}=\mathrm{WN} \cap(I \times F \times F)$
- $\Gamma^{\prime}=\Gamma \times Q \times Q$

Intuitively, the VPA $A^{\prime}$ simulates the VPA $A$ as follows: if a run of $A^{\prime}$ goes through a state $(p, q, r)$ with a stack $\sigma^{\prime}$ of height $n$, then the run of $A^{\prime}$ at this position mimicks a run of $A$ whose current configuration is $(q, \sigma)$, with $\sigma$ of height $n$, and such that the top symbol of $\sigma$ has been pushed when reaching the state $p$, and will be popped when leaving the state $r$. Moreover, from $p$ to $r$ in the run of $A$, the height of the stack is always larger or equal to $n$.

It remains to define the last component $\delta^{\prime}$. We define it by its restrictions on call, return and internal symbols respectively (namely $\delta_{c}^{\prime}, \delta_{r}^{\prime}$ and $\delta_{\iota}^{\prime}$ ).

Call symbols. Let $\left(p_{1}, p_{2}, p_{4}\right) \in \mathrm{WN}, c \in \Sigma_{c}$. Then $\left(\left(p_{1}, p_{2}, p_{4}\right), c,\left(\gamma, p_{1}, p_{4}\right),\left(q_{1}, q_{1}, q_{2}\right)\right) \in$ $\delta_{c}^{\prime}$ iff the following three conditions hold:
$-\left(p_{2}, c, \gamma, q_{1}\right) \in \delta_{c}$,
$-\left(q_{1}, q_{1}, q_{2}\right) \in \mathrm{WN}$,

- there exists a state $p_{3}$ such that $\left(q_{2}, r, \gamma, p_{3}\right) \in \delta_{r}$, and $\left(p_{3}, p_{3}, p_{4}\right) \in \mathrm{WN}$.

Return symbols. Let $\left(q_{1}, q_{2}, q_{2}\right) \in \mathrm{WN}, r \in \Sigma_{r}$. Then $\left(\left(q_{1}, q_{2}, q_{2}\right), r,\left(\gamma, p_{1}, p_{4}\right),\left(p_{1}, p_{3}, p_{4}\right)\right) \in$ $\delta_{r}^{\prime}$ iff the following three conditions hold:
$-\left(q_{2}, r, \gamma, p_{3}\right) \in \delta_{r}$,

- $\left(p_{1}, p_{3}, p_{4}\right) \in \mathrm{WN}$,
- there exists a state $p_{2}$ such that $\left(p_{2}, c, \gamma, q_{1}\right) \in \delta_{c}$, and $\left(p_{1}, p_{2}, p_{2}\right) \in \mathrm{WN}$.

Schematically, for $w_{1}, w_{2}, w_{3} \in \Sigma_{\mathrm{wn}}^{*}, c \in \Sigma_{c}$, and $r \in \Sigma_{r}$, we have:


Internal symbols. Let $\left(p_{1}, q, p_{2}\right),\left(p_{1}, q^{\prime}, p_{2}\right) \in \mathbf{W N}, a \in \Sigma_{\iota}$. Then $\left(\left(p_{1}, q, p_{2}\right), a,\left(p_{1}, q^{\prime}, p_{2}\right)\right) \in$ $\delta_{\iota}^{\prime}$ iff the transition $\left(q, a, q^{\prime}\right)$ belongs to $\delta_{\iota}$.

Proposition 3. For any VPA $A$, $\operatorname{trim}(A)$ can be computed in polynomial time.

This result easily follows from the definition of trim $(A)$.
Proposition 4. For any VPA $A, L(A)=L(\operatorname{trim}(A))$.
To prove this result, we will first prove different lemmas.
Lemma 5. Let $A$ be a VPA and $A^{\prime}=\operatorname{trim}(A)$. Let $\rho$ be a run of $A^{\prime}$ such that $\rho$ : $((p, q, r), \sigma) \xrightarrow{w}\left(\left(p^{\prime}, q^{\prime}, r^{\prime}\right), \sigma\right)$, with $\sigma \in \Gamma^{\prime *}$ and $w \in \Sigma_{\mathrm{wn}}^{*}$. Then we have $p=p^{\prime}$ and $r=r^{\prime}$.

Proof. The proof goes by induction on the structure of the word $w$. It holds trivially when $w$ is the empty word. Consider now a non-empty word $w$. There are two cases:

- First case: $w=a w^{\prime}$, with $a \in \Sigma_{\iota}, w^{\prime} \in \Sigma_{\mathrm{wn}}^{*}$ and $\left(\left(p_{1}, q_{1}, r_{1}\right), a,(p, q, r)\right) \in \delta_{\iota}^{\prime}$. By construction of $A^{\prime}, p_{1}=p$ and $r_{1}=r$. We conclude by using the induction hypothesis on $w^{\prime}$.
- Second case: $w=c w^{\prime} r w^{\prime \prime}$, with $c \in \Sigma_{c}, r \in \Sigma_{r}, w^{\prime}, w^{\prime \prime} \in \Sigma_{\mathrm{wn}}^{*}$. There exists a transition $\left(\left(p_{1}, p_{2}, p_{4}\right), c,\left(\gamma, p_{1}, p_{4}\right),\left(q_{1}, q_{1}, q_{2}\right)\right) \in \delta_{c}^{\prime}$ which is used when reading the first letter $c$. There exists a state $\left(r_{1}, r_{2}, r_{3}\right)$ such that the run of $A^{\prime}$ on $w^{\prime}$ goes from state $\left(q_{1}, q_{1}, q_{2}\right)$ to state $\left(r_{1}, r_{2}, r_{3}\right)$. As $w^{\prime}$ is well-nested, the induction hypothesis applied on $w^{\prime}$ entails that $r_{1}=q_{1}$ and $r_{3}=q_{2}$. In addition, as a return transition is used after reading $w^{\prime}$, the definition of $\delta_{r}^{\prime}$ implies that $r_{2}=r_{3}$. Thus by definition of $\delta_{c}^{\prime}$ and $\delta_{r}^{\prime}$, a transition of the form $\left(\left(q_{1}, q_{2}, q_{2}\right), r,\left(\gamma, p_{1}, p_{4}\right),\left(p_{1}, p_{3}, p_{4}\right)\right) \in$ $\delta_{r}^{\prime}$ for some state $p_{3}$ is used when reading the letter $r$. We conclude by induction hypothesis applied on $w^{\prime \prime}$.

We build a bijection between accepting runs of $A$ and $A^{\prime}$. First, we go from runs of $A^{\prime}$ to runs of $A$. We define the following standard projection mappings. Consider two integers $1 \leq j \leq k$. Given a tuple $s=\left(s_{1}, s_{2}, \ldots, s_{k}\right)$, we denote by $\pi_{j}(s)$ the element $s_{j}$. In addition, we extend this mapping over words on tuples, by letting $\pi_{j}(\sigma)=\pi_{j}\left(\gamma_{1}\right) \pi_{j}\left(\gamma_{2}\right) \ldots \pi_{j}\left(\gamma_{n}\right)$ where $\sigma=\gamma_{1} \gamma_{2} \ldots \gamma_{n}$.

Lemma 6. Let $A$ be a VPA and $A^{\prime}=\operatorname{trim}(A)$. For any accepting run $\rho^{\prime}=\left(q_{i}^{\prime}, \sigma_{i}^{\prime}\right)_{1 \leq i \leq k}$ of $A^{\prime}, \rho=\left(\pi_{2}\left(q_{i}^{\prime}\right), \pi_{1}\left(\sigma_{i}^{\prime}\right)\right)_{1 \leq i \leq k}$ is an accepting run of $A$.

Proof. We prove by induction on $\rho^{\prime}$ that the projection run $\rho$ defined above is a correct run of $A$. The fact that accepting runs of $A^{\prime}$ are projected on accepting runs of $A$ is trivial.

Conversely, considering translation of runs of $A$ to runs of $A^{\prime}$. States of $A^{\prime}$ extend states of $A$ by considering starting and ending states for the current stack level. Therefore we prove the following lemma in which we consider a word $w_{2} \in \Sigma_{\text {wn }}^{*}$ embedded into a context $\left(w_{1}, w_{3}\right) \in \Sigma_{\mathrm{wn}}^{*} \times \Sigma_{\mathrm{wn}}^{*}$ such that $w_{1} w_{2} w_{3}$ corresponds to the current stack level.

Lemma 7. Let $A$ be a VPA and $A^{\prime}=\operatorname{trim}(A)$. For all $w_{1}, w_{2}, w_{3} \in \Sigma_{w n}^{*}$, if there exists a run $\rho$ of the form $(p, \perp) \xrightarrow{w_{1}}{ }^{*}(q, \perp) \xrightarrow{w_{2}} *(r, \perp) \xrightarrow{w_{3}}{ }^{*}(s, \perp)$ in the VPA $A$, then there is a run $\rho^{\prime}$ of the form $((p, q, s), \perp) \xrightarrow{w_{2}}{ }^{*}((p, r, s), \perp)$ of $A^{\prime}$.

Proof. We prove this lemma by induction on the height of the word $w_{2}$. If $h\left(w_{2}\right)=0$, then $w_{2} \in \Sigma_{\iota}^{*}$, assuming $w_{2}=a_{1} a_{2} \ldots a_{n}$, the restriction of $\rho$ on $w_{2}$ is of the form

$$
(q, \perp) \xrightarrow{a_{1}}\left(q_{1}, \perp\right) \ldots \xrightarrow{a_{n}}\left(q_{n}, \perp\right) \text { with } r=q_{n} .
$$

Observe that for all $i,\left(p, q_{i}, s\right) \in \mathrm{WN}$. By definition of $A^{\prime}$ the following is a run of $A^{\prime}$ on $w_{2}$ :

$$
\rho^{\prime}=((p, q, s), \perp) \xrightarrow{a_{1}}\left(\left(p, q_{1}, s\right), \perp\right) \ldots \xrightarrow{a_{n}}\left(\left(p, q_{n}, s\right), \perp\right) \text { with } r=q_{n} \text {. }
$$

We now assume for the induction that the property holds when $h\left(w_{2}\right) \leq n$ and consider $w_{2}$ such that $h\left(w_{2}\right)=n+1$. There exists a unique decomposition of $w_{2}$ as follows:

$$
w_{2}=\left[\Pi_{i=1}^{k}\left(w_{i}^{\mathrm{int}} c_{i} w_{i}^{\mathrm{wn}} r_{i}\right)\right] w_{k+1}^{\mathrm{int}}
$$

with for all $i, w_{i}^{\mathrm{int}} \in \Sigma_{i}^{*}, w_{i}^{\mathrm{wn}} \in \Sigma_{w n}^{*}, c_{i} \in \Sigma_{c}$, and $r_{i} \in \Sigma_{r}$.
Let $w_{1}, w_{3} \in \Sigma_{w n}^{*}$ and $\rho$ be a run of $A$ on $w_{1} w_{2} w_{3}$. We decompose the run $\rho$ as follows:

- on $w_{1}:(p, \perp) \xrightarrow{w_{1}}{ }^{*}\left(p_{1}^{1}, \perp\right)$
- on each $w_{i}^{\text {int }} c_{i} w_{i}^{\text {wn }} r_{i}$ :

$$
\left(p_{i}^{1}, \perp\right) \xrightarrow{w_{i}^{\mathrm{intt}}} *\left(p_{i}^{2}, \perp\right) \xrightarrow{c_{i}}\left(q_{i}^{1}, \gamma_{i}\right) \xrightarrow{w_{i}^{\mathrm{wn}} *}\left(q_{i}^{2}, \gamma_{i}\right) \xrightarrow{r_{i}}\left(p_{i+1}^{1}, \perp\right)
$$

- on $w_{k+1}^{\mathrm{int}}:\left(p_{k+1}^{1}, \perp\right) \xrightarrow{w_{k+1}^{\mathrm{int}} *}\left(p_{k+1}^{2}, \perp\right)$
- on $w_{3}:\left(p_{k+1}^{2}, \perp\right) \xrightarrow{w_{3}}{ }^{*}(s, \perp)$

Note that $q=p_{1}^{1}$ and $r=p_{k+1}^{2}$.
As $w_{i}^{\mathrm{wn}}$ is well-nested, there exists a run in $A$ of the form $\left(q_{i}^{1}, \perp\right) \xrightarrow{w_{i}^{\mathrm{wn}} *}\left(q_{i}^{2}, \perp\right)$. By induction hypothesis, this implies that there exists a run $\rho_{i}^{\prime}=\left(\left(q_{i}^{1}, q_{i}^{1}, q_{i}^{2}\right), \perp\right) \xrightarrow{w_{i}^{\mathrm{wn}}} *$ $\left(\left(q_{i}^{1}, q_{i}^{2}, q_{i}^{2}\right), \perp\right)$ of $A^{\prime}$ on $w_{i}^{\mathrm{wn}}$. Again, using the fact that $w_{i}^{\mathrm{wn}}$ is well-nested, we have $\left(\left(q_{i}^{1}, q_{i}^{1}, q_{i}^{2}\right), \gamma_{i}^{\prime}\right) \xrightarrow{w_{i}^{w n}} *\left(\left(q_{i}^{1}, q_{i}^{2}, q_{i}^{2}\right), \gamma_{i}^{\prime}\right)$ in $A^{\prime}$ for any $\gamma_{i}^{\prime} \in \Gamma^{\prime}$.

We now describe the other parts of the run $\rho^{\prime}$ of $A^{\prime}$ on $w_{2}$.
Internal actions For all $i \in\{1, \ldots, k+1\},\left(\left(p, p_{i}^{1}, s\right), \perp\right) \xrightarrow{w_{i}^{\text {int }}} *\left(\left(p, p_{i}^{2}, s\right), \perp\right)$ on $w_{i}^{\text {int }}$, by the base induction with $w_{1}$ and $w_{3}$ well defined.
Calls For all $i \in\{1, \ldots, k\}$, we have $\left(p_{i}^{2}, c_{i}, \gamma_{i}, q_{i}^{1}\right) \in \delta_{c}$ in $A$. Thus by definition of $A^{\prime}$, there exists a transition $\left(\left(p, p_{i}^{2}, s\right), c_{i},\left(\gamma_{i}, p, s\right),\left(q_{i}^{1}, q_{i}^{1}, q_{i}^{2}\right)\right) \in \delta_{c}^{\prime}$ in $A^{\prime}$.
Returns For all $i \in\{1, \ldots, k\}$, we have $\left(q_{i}^{2}, r_{i}, \gamma_{i}, p_{i+1}^{1}\right) \in \delta_{r}$ in $A$. Thus by definition of $A^{\prime}$ there exists a transition $\left(\left(q_{i}^{1}, q_{i}^{2}, q_{i}^{2}\right), r_{i},\left(\gamma_{i}, p, s\right),\left(p, p_{i+1}^{1}, s\right)\right) \in \delta_{r}^{\prime}$ in $A^{\prime}$.

One can then easily check that all the above transitions and partial runs can be gathered to obtain a run of $A^{\prime}$ on $w_{2}$ with the expected form.

To conclude the proof of Proposition 4, note that Lemma 6 implies that $L(\operatorname{trim}(A)) \subseteq$ $L(A)$. Conversely, using Lemma 7 for $w_{1}=w_{3}=\epsilon$ yields $p=q$ ans $r=s$ and if moreover $p$ is initial and $q$ is final then any accepting run in $A$ is associated with an accepting run in $\operatorname{trim}(A)$. Hence, $L(A) \subseteq L(\operatorname{trim}(A))$.

Proposition 8. For any VPA $A$, $\operatorname{trim}(A)$ is trimmed.
Proof. Let $A^{\prime}$ be trim $(A)$. Therefore, we first prove that any accessible configuration of $A^{\prime}$ is also co-accessible, and then the converse.

Let $c o=((p, q, r), \sigma)$ be an accessible configuration of $A^{\prime}$. There exists a run $\rho$ of $A^{\prime}$ of the form $((i, i, f), \perp) \rightarrow^{*}$ co between an initial configuration $((i, i, f), \perp)$ and $c o$. We show by induction on the height of the stack that we can reach a final configuration from $c o$.

If $|\sigma|=0$, by the lemma 5 , we obtain $p=i$ and $r=f$. In particular, this implies $p \in I$ and $r \in F$. Since $(p, q, r) \in \mathrm{WN}$, there is a run $(p, \perp) \rightarrow^{*}(q, \perp) \rightarrow^{*}(r, \perp)$ in $A$, thus by Lemma 7 we have a run $((p, q, r), \perp) \rightarrow^{*}((p, r, r), \perp)$ of $A^{\prime}$. This concludes this case as by the above observation, we have $(p, r, r) \in I \times F \times F=F^{\prime}$.

We now assume for the induction that the property holds when $|\sigma| \leq n$ and we consider a stack $\sigma$ such that $|\sigma|=n+1$. Let denote by $\left(\gamma, p^{\prime}, r^{\prime}\right)$ the top symbol of $\sigma$, write $\sigma=\sigma^{\prime} .\left(\gamma, p^{\prime}, r^{\prime}\right)$, and consider the first position in the run $\rho$ that pushes this symbol onto the stack. We denote by $c$ the associated call. More precisely, there exists a unique decomposition of $\rho$ as follows:

$$
((i, i, f), \perp) \rightarrow^{*}\left(\left(p^{\prime}, q^{\prime}, r^{\prime}\right), \sigma^{\prime}\right) \xrightarrow{c}\left(\left(p^{\prime \prime}, p^{\prime \prime}, r^{\prime \prime}\right), \sigma\right) \rightarrow^{*}((p, q, r), \sigma)
$$

such that the run from $\left(\left(p^{\prime \prime}, p^{\prime \prime}, r^{\prime \prime}\right), \sigma\right)$ to $((p, q, r), \sigma)$ is associated with a well-nested word. By Lemma 5, we obtain $p^{\prime \prime}=p$ and $r^{\prime \prime}=r$. Considering the call transition associated with $c$, and by definition of $\delta_{c}^{\prime}$ and $\delta_{r}^{\prime}$, there exists a return transition $\left((p, r, r), r,\left(\gamma, p^{\prime}, r^{\prime}\right),\left(p^{\prime}, q^{\prime \prime}, r^{\prime}\right)\right) \in \delta_{r}^{\prime}$ for some letter $r \in \Sigma_{r}$. In addition, as $(p, q, r) \in \mathrm{WN}$, there is a run $(p, \sigma) \rightarrow^{*}(q, \sigma) \rightarrow^{*}(r, \sigma)$ in $A$, and thus by Lemma 7 we have a run $((p, q, r), \sigma) \rightarrow^{*}((p, r, r), \sigma)$. As the top symbol of $\sigma$ is $\left(\gamma, p^{\prime}, r^{\prime}\right)$, the above return transition can be used to reach configuration $\left(\left(p^{\prime}, q^{\prime \prime}, r^{\prime}\right), \sigma^{\prime}\right)$ whose height is $n$. The result follows by induction hypothesis.

Conversely, let $c o=((p, q, r), \sigma)$ be a co-accessible configuration of $A^{\prime}$. There exists a run $\rho$ of $A^{\prime}$ of the form $c o \rightarrow^{*}((i, f, f), \perp)$ between co and a final configuration $((i, f, f), \perp)$. We show by induction on the height of the stack that co can be reached from an initial configuration.

If $|\sigma|=0$, by the lemma 5 , we obtain $p=i$ and $r=f$. In particular, this implies $p \in I$ and $r \in F$. Since $(p, q, r) \in \mathrm{WN}$, there is a run $(p, \perp) \rightarrow^{*}(q, \perp) \rightarrow^{*}(r, \perp)$ in $A$, thus by Lemma 7 we have a run $((p, p, r), \perp) \rightarrow^{*}((p, q, r), \perp)$ of $A^{\prime}$. This concludes this case as by the above observation, we have $(p, p, r) \in I \times I \times F=I^{\prime}$.

We now assume for the induction that the property holds when $|\sigma| \leq n$ and we consider a stack $\sigma$ such that $|\sigma|=n+1$. Let denote by ( $\gamma, p^{\prime}, r^{\prime}$ ) the top symbol of $\sigma$, write $\sigma=\sigma^{\prime} .\left(\gamma, p^{\prime}, r^{\prime}\right)$, and consider the first position in the run $\rho$ that pops this symbol from the stack. We denote by $r$ the associated return. More precisely, there
exists a unique decomposition of $\rho$ as follows:

$$
((p, q, r), \sigma) \rightarrow^{*}\left(\left(p^{\prime \prime}, r^{\prime \prime}, r^{\prime \prime}\right), \sigma\right) \xrightarrow{r}\left(\left(p^{\prime}, q^{\prime}, r^{\prime}\right), \sigma^{\prime}\right) \rightarrow^{*}((i, f, f), \perp)
$$

such that the run from $((p, q, r), \sigma)$ to $\left(\left(p^{\prime \prime}, r^{\prime \prime}, r^{\prime \prime}\right), \sigma\right)$ is associated with a well-nested word. By Lemma 5, we obtain $p^{\prime \prime}=p$ and $r^{\prime \prime}=r$. Considering the return transition associated with $r$, and by definition of $\delta_{c}^{\prime}$ and $\delta_{r}^{\prime}$, there exists a call transition $\left(\left(p^{\prime}, q^{\prime \prime}, r^{\prime}\right), c,\left(\gamma, p^{\prime}, r^{\prime}\right),(p, p, r)\right) \in \delta_{c}^{\prime}$ for some letter $c \in \Sigma_{c}$. In addition, as $(p, q, r) \in$ WN , there is a run $(p, \sigma) \rightarrow^{*}(q, \sigma) \rightarrow^{*}(r, \sigma)$ in $A$, and thus by Lemma 7 we have a run $((p, p, r), \sigma) \rightarrow^{*}((p, q, r), \sigma)$. As the top symbol of $\sigma$ is $\left(\gamma, p^{\prime}, r^{\prime}\right)$, the above call transition can be used. As a consequence, we have proven that there exists a run from configuration $\left(\left(p^{\prime}, q^{\prime \prime}, r^{\prime}\right), \sigma^{\prime}\right)$ to configuration co. In particular, configuration $\left(\left(p^{\prime}, q^{\prime \prime}, r^{\prime}\right), \sigma^{\prime}\right)$ is a co-accessible configuration whose height is $n$, and the result follows by induction hypothesis.

To summarize the results presented in this paper, we have the following theorem:
Theorem 9. Let $A$ be a VPA. Then:

- $\operatorname{trim}(A)$ can be built in polynomial time
- $L(A)=L(\operatorname{trim}(A))$
$-\operatorname{trim}(A)$ is trimmed
Note that by definition of trimmed VPA, some useless state may remain in the construction producing $\operatorname{trim}(A)$ : indeed, $\operatorname{trim}(A)$ may contain some state $q$ such that for no stack $\sigma,(q, \sigma)$ is accessible (and thus co-accessible). However, deciding wether for some stack $\sigma,(q, \sigma)$ is accessible can be reduced to the emptiness problem of pushdown automata accepting on final states. This problem is decidable in polynomial time.


## References

1. R. Alur and P. Madhusudan. Visibly pushdown languages. In STOC, pages 202-211, 2004.
2. R. Alur and P. Madhusudan. Adding nesting structure to words. JACM, 56(3):1-43, 2009.
3. D. Girault-Beauquier. Some results about finite and infinite behaviours of a pushdown automaton. In Automata, Languages and Programming, volume 172 of Lecture Notes in Computer Science, pages 187-195. Springer, 1984.
