

Weighted Pushdown Systems with Indexed Weight Domains

Yasuhiko Minamide

Faculty of Engineering, Information and Systems
University of Tsukuba

Abstract. The reachability analysis of weighted pushdown systems is a very powerful technique in verification and analysis of recursive programs. Each transition rule of a weighted pushdown system is associated with an element of a bounded semiring representing the weight of the rule. However, we have realized that the restriction of the boundedness is too strict and the formulation of weighted pushdown systems is not general enough for some applications.

To generalize weighted pushdown systems, we first introduce the notion of stack signatures that summarize the effect of a computation of a pushdown system and formulate pushdown systems as automata over the monoid of stack signatures. We then generalize weighted pushdown systems by introducing semirings indexed by the monoid and weaken the boundedness to local boundedness.

1 Introduction

The reachability analysis of weighted pushdown systems is a very powerful technique in verification and analysis of recursive programs [RSJM05]. Each transition rule of a weighted pushdown system is associated with an element of a semiring representing the weight of the rule. To guarantee termination of the analysis, the semiring of the weight must be bounded: there should be no infinite descending sequence of weight. However, recently, we have realized that this restriction of the boundedness is too strict and the formulation of weighted pushdown systems is not general enough for some applications. For the two applications below, the standard algorithm for the reachability analysis of weighted pushdown systems actually works and terminates. However, they require semirings that are not bounded and thus the standard framework of weighted pushdown systems cannot guarantee termination.

The first application is the reachability analysis of conditional pushdown systems. Conditional pushdown systems extend pushdown systems with the ability to check the whole stack content against a regular language [EKS03, LO10]. We proposed an algorithm of their reachability analysis in our previous work on the analysis of HTML 5 parser specification [MM12]. After the development of the algorithm, we realized that the algorithm can be considered as the reachability analysis of weighted pushdown systems. However, it required an unbounded semiring.

The second application is the analysis of recursive programs with local variables. For the efficient analysis of recursive programs, Suwimonteerabuth proposed an encoding of local variables into weight implemented with BDDs [Suw09]. The weight has a structure depending on a configuration of stack and requires a semiring that is not bounded.

To generalize weighted pushdown systems, we first introduce *stack signatures* that summarize the effect of a computation of a pushdown system as a pair of words over stack alphabet. A stack signature w_1/w_2 represents a computation of a pushdown system that pops w_1 and pushes w_2 as its total effect. We show that the set of stack signatures forms an ordered monoid, *i.e.*, a monoid that is equipped with a partial order compatible with the multiplication of the monoid. We then formulate pushdown systems as automata over the monoid of stack signatures.

We extend the structure of weight by introducing semirings indexed by a monoid element. Weighted pushdown systems are generalized to those over a semiring indexed by the monoid of stack signatures. We show that the reachability analysis of weighted pushdown systems can be refined to those over an indexed semiring and the boundedness can be replaced with the *local boundedness*.

Finally, we show two applications of weighted pushdown systems over a semiring indexed by stack signatures. The first one is a simplified version of the structure used by Suwimonteerabuth to encode local variables of a recursive program. The other is an indexed semiring to encode the reachability analysis of conditional pushdown systems into that of weighted pushdown systems. Since both of these indexed semirings are locally bounded, our framework guarantees termination of the two analyses.

2 Semirings and Weighted Automata

We first review the definitions of semirings and weighted automata.

Definition 1. A *semiring* is a structure $\mathcal{S} = \langle D, \oplus, \otimes, 0, 1 \rangle$ where D is a set, 0 and 1 are elements of D , \oplus and \otimes are binary operations on D such that

1. $\langle D, \oplus, 0 \rangle$ is a commutative monoid.
2. $\langle D, \otimes, 1 \rangle$ is a monoid.
3. \otimes distributes over \oplus .

$$(x \oplus y) \otimes z = (x \otimes z) \oplus (y \otimes z) \quad x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$$

4. 0 is an annihilator with respect to \otimes : $0 \otimes x = 0 = x \otimes 0$ for all $x \in D$.

We say that a semiring \mathcal{S} is *idempotent* if its addition \oplus is idempotent (*i.e.*, $a \oplus a = a$). For an idempotent semiring $\langle D, \oplus, \otimes, 0, 1 \rangle$, $\langle D, \oplus \rangle$ can be considered as a join semilattice¹. Then, the partial order \sqsubseteq is defined by $a \sqsubseteq b$ iff $a \oplus b = b$

¹ In [RSJM05], it is considered as a meet semilattice.

for an idempotent semiring. We say that an idempotent semiring is *bounded* if there are no infinite ascending chains with respect to \sqsubseteq .

In this paper, we consider weighted automata without initial and final states.

Definition 2. A *weighted automaton* \mathcal{A} over an idempotent semiring \mathcal{S} and an alphabet Γ is a structure $\langle \Gamma, Q, E \rangle$ where Q is a finite set of states, $E : Q \times \Gamma \times Q \rightarrow \mathcal{S}$ is a set of transition rules each of which associates an element in \mathcal{S} as weight.

For weighted automata over alphabet Γ and semiring $\mathcal{S} = \langle D, \oplus, \otimes, 0, 1 \rangle$, we introduce the transition relation of the form $q \xrightarrow{w|a} q'$ where $w \in \Gamma^*$ and $a \in D$. It is inductively defined as follows.

- $q \xrightarrow{\epsilon|1} q$ for any $q \in Q$.
- $q \xrightarrow{\gamma|a} q'$ if $a = E(\langle q, \gamma, q' \rangle)$.
- $q \xrightarrow{ww'|a \otimes b} q'$ if $q \xrightarrow{w|a} q''$ and $q'' \xrightarrow{w'|b} q'$.

Then, for two states q and q' and a word w , we consider the total weight of the transitions of the form $q \xrightarrow{w|a} q'$ defined as follows².

$$\delta(q, w, q') = \bigoplus \{a \mid q \xrightarrow{w|a} q'\}$$

This is well-defined because there are only finitely many transitions of this form and we assume that the semiring is idempotent. In the general theory of weighted automata, we do not impose that the semiring is idempotent [ÉK09]. However, we impose the condition to adopt the simple and intuitive definition above.

3 Stack Signatures

We introduce stack signatures that summarize the effect of a transition on stack as a pair of words over stack alphabet. It is shown that the set of stack signatures forms a monoid, and then a semiring by introducing a partial order on them. Stack signatures naturally appear in the theory of context-free grammars and pushdown systems [Suw09, MT06, TM07]. We adopt the term ‘stack signature’ introduced by Suwimonterabuth [Suw09].

The proofs of propositions and theorems in this section are not fundamentally difficult, but require detailed case-analysis. Thus, we have formalized and proved them in Isabelle/HOL by extending our previous work on a formalization of decision procedures on context-free grammars [Min07]³.

² This is basically a formal power series, which is used to define the behaviour of weighted automata [ÉK09].

³ The proof script can be found at

<http://www.score.cs.tsukuba.ac.jp/~minamide/stacksig.tar.gz>

The effect of a transition of a pushdown system can be summarized as a pair of sequences of stack symbols written w_1/w_2 where w_1 are the symbols popped by the transition and w_2 are those pushed by the transition. We consider that pushing γ and then popping the same γ cancel the effect, but popping γ and then pushing γ have the effect γ/γ .

Definition 3. We call elements of $\Gamma^* \times \Gamma^*$ stack signatures and write w/w' for a stack signature $\langle w, w' \rangle$.

- We say that w_1/w'_1 and w_2/w'_2 are compatible if either w'_1 is a prefix of w_2 or w_2 is a prefix of w'_1 . Furthermore, they are called strictly compatible if $w'_1 = w_2$.
- For compatible w_1/w'_1 and w_2/w'_2 , we define $w_1/w'_1 \cdot w_2/w'_2$ by

$$w_1/w'_1 \cdot w_2/w'_2 = \begin{cases} w_1/w'_2 w'_1 & \text{if } w'_1 = w_2 w'_1 \\ w_1 w'_2/w'_2 & \text{if } w_2 = w'_1 w'_2 \end{cases}$$

For example, we have $\gamma_1/\gamma_2 \cdot \gamma_2\gamma_3/\gamma_4 = \gamma_1\gamma_3/\gamma_4$. By introducing an element \top and extending the definition \cdot as follows, $\langle (\Gamma^* \times \Gamma^*) \cup \{\top\}, \cdot, \epsilon/\epsilon \rangle$ forms a monoid. We write \mathcal{M}_Γ for this monoid.

$$\begin{aligned} \top \cdot \sigma &= \sigma \cdot \top = \top & \text{for } \sigma \in \mathcal{M}_\Gamma \\ w_1/w'_1 \cdot w_2/w'_2 &= \top & \text{if } w_1/w'_1 \text{ and } w_2/w'_2 \text{ are not compatible} \end{aligned}$$

By relaxing the use of terminology, we call an element of \mathcal{M}_Γ a *stack signature* and an element of the form w/w' a *proper stack signature*.

The following isomorphism is used to relate automata and pushdown systems. It is clear from $w_1/\epsilon \cdot w_2/\epsilon = w_1 w_2/\epsilon$.

Proposition 1. The set $\{w/\epsilon \mid w \in \Gamma^*\}$ is a submonoid of \mathcal{M}_Γ . Furthermore, it is isomorphic to Γ^* by the function projecting w from w/ϵ .

We also introduce a partial order on stack signatures: a transition that pops w_1 and pushes w_2 can be considered as one that pops $w_1 w$ and pushes $w_2 w$ for any $w \in \Gamma^*$.

Definition 4. A partial order \leq on stack signatures is defined by $w_1/w_2 \leq w_1 w/w_2 w$ for $w_1, w_2, w \in \Gamma^*$ and $\sigma \leq \top$ for any stack signature σ .

It is clear that $(\Gamma^* \times \Gamma^*) \cup \{\top\}$ is a join-semilattice. This partial order is compatible with the binary operation \cdot : if $\sigma_1 \leq \sigma'_1$ and $\sigma_2 \leq \sigma'_2$, then $\sigma_1 \cdot \sigma_2 \leq \sigma'_1 \cdot \sigma'_2$. Thus, the monoid of stack signatures is an *ordered monoid*⁴.

Furthermore, we can construct an idempotent semiring by introducing the bottom element \perp and extending \cdot for \perp as follows.

$$\perp \cdot x = x \cdot \perp = \perp \quad \text{for all } x \in (\Gamma^* \times \Gamma^*) \cup \{\top, \perp\}$$

Proposition 2. Let $S = (\Gamma^* \times \Gamma^*) \cup \{\top, \perp\}$. $\langle S, \perp, \cdot, \perp, \epsilon/\epsilon \rangle$ forms an idempotent semiring.

This semiring is not bounded because $\epsilon/\epsilon \leq \gamma/\gamma \leq \gamma\gamma/\gamma\gamma \leq \dots$.

⁴ A monoid is ordered when it is equipped with a compatible partial order.

4 Semirings Indexed by a Monoid

We introduce semirings indexed by a monoid, which is a typed algebraic structure where a type is an element of a monoid. Weighted pushdown systems are generalized by taking this structure as the domain of weight in the next section.

Definition 5. Let $\mathcal{M} = \langle M, \cdot, 1_{\mathcal{M}} \rangle$ be a monoid. An indexed semiring \mathcal{S} over \mathcal{M} is a structure $\langle \{D_m\}, \{\oplus_m\}, \{\otimes_{m_1, m_2}\}, \{0_m\}, 1 \rangle$ such that

- D_m is a set for each $m \in M$.
- $\langle D_m, \oplus_m, 0_m \rangle$ is a commutative monoid for $m \in M$.
- \otimes_{m_1, m_2} is an associative binary operation of type $D_{m_1} \times D_{m_2} \rightarrow D_{m_1 m_2}$ for $m_1, m_2 \in M$.

$$(a \otimes_{m_1, m_2} b) \otimes_{m_1 m_2, m_3} c = a \otimes_{m_1, m_2 m_3} (b \otimes_{m_2, m_3} c)$$

- $1 \in D_{1_{\mathcal{M}}}$ is a neutral element of $\otimes_{m, m'}: a \otimes_{m, 1_{\mathcal{M}}} 1 = 1 \otimes_{1_{\mathcal{M}}, m} a = a$.
- \otimes_{m_1, m_2} distributes over \oplus_m .

$$(a \oplus_{m_1} b) \otimes_{m_1, m_2} c = (a \otimes_{m_1, m_2} c) \oplus_{m_1 m_2} (b \otimes_{m_1, m_2} c)$$

$$a \otimes_{m_1, m_2} (b \oplus_{m_2} c) = (a \otimes_{m_1, m_2} b) \oplus_{m_1 m_2} (a \otimes_{m_1, m_2} c)$$

- 0_m is an annihilator with respect to $\otimes_{m, m'}$.

$$0_{m_1} \otimes_{m_1, m_2} a = 0_{m_1 m_2} = b \otimes_{m_1, m_2} 0_{m_2}$$

We call \mathcal{S} an idempotent indexed semiring if \mathcal{S} is an indexed semiring where \oplus_m is idempotent for all $m \in M$. We introduce partial orders \sqsubseteq_m defined by $a \sqsubseteq_m b$ iff $a \oplus_m b = b$. From distributivity of \otimes , it is clear that \otimes is monotonic with respect to \sqsubseteq_m .

Proposition 3. Let $\mathcal{M} = \langle M, \cdot, 1_M \rangle$ be a monoid and \mathcal{S} a semiring indexed by \mathcal{M} . If \mathcal{M}' is a submonoid of \mathcal{M} , then the restriction of \mathcal{S} on \mathcal{M}' is a semiring indexed by \mathcal{M}' .

The notion of weighted automata can be extended for an indexed semiring over the monoid Γ^* in the straightforward manner.

Definition 6. Let \mathcal{S} be an idempotent semiring $\langle \{D_w\}, \{\oplus_w\}, \{\otimes_{w_1, w_2}\}, \{0_w\}, 1 \rangle$ indexed by Γ^* . A weighted automaton \mathcal{A} over \mathcal{S} is a structure $\langle \Gamma, Q, E \rangle$ where Q is a finite set of states, and $E: Q \times \Gamma \times Q \rightarrow \bigcup_{\gamma \in \Gamma} D_{\gamma}$ is a set of transition rules assigning a weight such that $E(\langle q, \gamma, q' \rangle) \in D_{\gamma}$.

The definition of the transition relation is revised as follows. The only revision is that we apply indexed $\otimes_{w, w'}$ to combine two transitions for w and w' .

- $q \xrightarrow{\epsilon | 1} q$ for any $q \in Q$.
- $q \xrightarrow{\gamma | a} q'$ if $a = E(\langle q, \gamma, q' \rangle)$.
- $q \xrightarrow{w w' | a \otimes_{w, w'} b} q'$ if $q \xrightarrow{w | a} q''$ and $q'' \xrightarrow{w' | b} q'$.

5 Weighted Pushdown Systems over an Indexed Semiring and Their Reachability Analysis

We introduce weighted pushdown systems over a semiring indexed by the monoid of stack signatures. The reachability analysis of weighted pushdown systems is refined to those over an indexed semiring and the boundedness is relaxed to the local boundedness. We also show that it is possible to construct an ordinary semiring from an indexed semiring, but the obtained semiring is not bounded.

5.1 Weighted Pushdown Systems over an Indexed Semiring

We basically consider pushdown systems over stack alphabet Γ as automata over the monoid of stack signatures \mathcal{M}_Γ . However, in order to clarify our presentation we introduce the definition of weighted pushdown systems independently.

Definition 7. Let $\mathcal{S} = \langle \{D_\sigma\}, \{\oplus_\sigma\}, \{\otimes_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, 1 \rangle$ be a semiring indexed by \mathcal{M}_Γ . A weighted pushdown system \mathcal{P} over \mathcal{S} is a structure $\langle P, \Gamma, \Delta \rangle$ where P is a finite set of states, Γ is a stack alphabet, and $\Delta \subseteq P \times \Gamma \times P \times \Gamma^* \times \bigcup_{\gamma \in \Gamma, w \in \Gamma^*} D_{\gamma/w}$ is a finite set of transitions such that $a \in D_{\gamma/w}$ for $\langle p, \gamma, p', w, a \rangle \in \Delta$.

A configuration of pushdown system \mathcal{P} is a pair $\langle p, w \rangle$ where $p \in P$ and $w \in \Gamma^*$. We write $\langle p, \gamma \rangle \xrightarrow{a} \langle p', w \rangle$ if $\langle p, \gamma, p', w, a \rangle \in \Delta$.

We consider pushdown systems as automata over stack signatures and define the translation relation as follows:

- $p \xrightarrow{\epsilon/\epsilon|1} p$.
- $p \xrightarrow{\gamma/w|a} p'$ if $\langle p, \gamma \rangle \xrightarrow{a} \langle p', w \rangle$.
- $p \xrightarrow{\sigma_1 \cdot \sigma_2|a} p'$ if $p \xrightarrow{\sigma_1|a_1} p', p' \xrightarrow{\sigma_2|a_2} p', a = a_1 \otimes_{\sigma_1, \sigma_2} a_2$ and $\sigma_1 \cdot \sigma_2 \neq \top$.

where we have $a \in D_\sigma$ if $p \xrightarrow{\sigma|a} p'$.

Traditionally, the transition relation on a pushdown system is defined as a relation between configurations. To introduce such a definition, we need to extend an indexed semiring with an additional operation.

Definition 8. Let \mathcal{M} be an ordered monoid with partial order \leq . By an indexed semiring over \mathcal{M} we shall mean an indexed semiring \mathcal{S} over \mathcal{M} on which there is a family of conversion functions $\uparrow_{m, m'}: D_m \rightarrow D_{m'}$ indexed by pairs of monoid elements $m \leq m'$ such that

- $\uparrow_{m, m} = \text{id}$.
- $\uparrow_{m, m''} = \uparrow_{m', m''} \circ \uparrow_{m, m'}$ for all $m \leq m' \leq m''$.
- $\uparrow_{m, m'}(0_m) = 0_{m'}$ and $\uparrow_{m, m'}(a \oplus_m b) = \uparrow_{m, m'}(a) \oplus_{m'} \uparrow_{m, m'}(b)$.
- $\uparrow_{m_1 m_2, m'_1 m'_2}(a \otimes_{m_1, m_2} b) = \uparrow_{m_1, m'_1}(a) \otimes_{m'_1, m'_2} \uparrow_{m_2, m'_2}(b)$ for all $m_1 \leq m'_1$ and $m_2 \leq m'_2$.

For an indexed semiring over the ordered monoid \mathcal{M}_Γ , we write \uparrow_w for $\uparrow_{w_1/w_2, w_1 w/w_2 w}$ if w_1 and w_2 are clear from the context. Then, the standard definition of the transition relation of a weighted pushdown system is given as follows.

- $\langle p, w \rangle \xRightarrow{\uparrow_w(1)} \langle p, w \rangle$.
- $\langle p, \gamma w' \rangle \xRightarrow{\uparrow_{w'}(a)} \langle p', w w' \rangle$ if $\langle p, \gamma \rangle \xrightarrow{a} \langle p', w \rangle$.
- $\langle p, w \rangle \xRightarrow{a} \langle p', w' \rangle$ if $\langle p, w \rangle \xRightarrow{a_1} \langle p'', w'' \rangle$, $\langle p'', w'' \rangle \xRightarrow{a_2} \langle p', w' \rangle$, and $a = a_1 \otimes_{w/w'', w''/w'} a_2$.

Then, these two definitions of transition relations are equivalent in the following sense.

Proposition 4. *If $\langle p, w \rangle \xRightarrow{a} \langle p', w' \rangle$, then there exist σ and a' such that $\sigma \leq w/w'$, $p \xRightarrow{\sigma|a'} p'$, and $a = \uparrow_{\sigma, w/w'}(a')$. Conversely, if $p \xRightarrow{\sigma|a'} p'$, then $\langle p, w \rangle \xRightarrow{\uparrow_{\sigma, w/w'}(a')} \langle p', w' \rangle$ for all $\sigma \leq w/w'$.*

As a special case of this proposition, we have $\langle p, w \rangle \xRightarrow{a} \langle p', \epsilon \rangle$ iff $p \xRightarrow{w/\epsilon|a} p'$.

5.2 Reachability Analysis

We show that the reachability analysis of weighted pushdown systems can be generalized for those over an indexed semiring, where we adopt a localized version of the boundedness of a semiring. We say an indexed idempotent semiring over \mathcal{M}_Γ is *locally bounded* if $D_{\gamma/\epsilon}$ is bounded for all $\gamma \in \Gamma$.

First, we focus on the (generalized) backward reachability to a configuration with the empty stack and consider the problem that computes the following function:

$$\delta(p, w, p') = \bigoplus \{a \mid p \xRightarrow{w/\epsilon|a} p'\}$$

where the addition above is the extension of $\oplus_{w/\epsilon}$ for a set. This function is well-defined if the indexed semiring is locally bounded. It is clear from the following equation:

$$\delta(p, \gamma w', p') = \bigoplus_{p'' \in P} (\delta(p, \gamma, p'') \otimes_{\gamma/\epsilon, w'/\epsilon} \delta(p'', w', p'))$$

where we have $\delta(p, \gamma, p'') \in D_{\gamma/\epsilon}$ for all $p'' \in P$. Although there are infinitely many transitions of the form $p \xRightarrow{\gamma/\epsilon|a} p''$, $\delta(p, \gamma, p'')$ is well-defined because $D_{\gamma/\epsilon}$ is bounded.

We generalize the reachability analysis of weighted pushdown automata for those over an indexed semiring. The algorithm is a generalization of the saturation procedure on \mathcal{P} -automata [BEM97, FWW97].

Let us consider a weighted pushdown system $\mathcal{P} = \langle P, \Gamma, \Delta \rangle$ over a semiring \mathcal{S} indexed by \mathcal{M}_Γ . We apply the procedure to a weighted automaton over the restriction of \mathcal{S} on $\{w/\epsilon \mid w \in \Gamma^*\}$ and start from $\mathcal{A}_0 = \langle P, \Gamma, E_0 \rangle$, which has no transition, i.e., $E_0(\langle p, \gamma, p' \rangle) = 0_{\gamma/\epsilon}$ for $p, p' \in P$ and $\gamma \in \Gamma$. Then, the weighted automaton $\mathcal{A}_{\text{pre}^*}$ representing $\delta_{\mathcal{P}}(p, \gamma, p')$ can be obtained by applying the *standard rule* for weighted pushdown systems to \mathcal{A}_0 until saturation. The following is the saturation rule of Reps *et al.* for the backward reachability analysis adapted to our framework [RSJM05].

- If $\langle p, \gamma \rangle \xrightarrow{a_1} \langle p', w \rangle$ and $p' \xrightarrow{w|a_2} p''$ in the current automaton, add a transition rule $p \xrightarrow{\gamma|a} p''$ where $a = a_1 \otimes_{\gamma/w, w/\epsilon} a_2$.

When we add $p \xrightarrow{\gamma|a} p''$, if there already exists transition $p \xrightarrow{\gamma|a'} p''$, then we replace it with $p \xrightarrow{\gamma|a \oplus_{\gamma/\epsilon} a'} p''$.

Since there is only a finite number of (one-step) transitions in $\mathcal{A}_{\text{pre}^*}$, it is clear that the application of the rule terminates if the indexed semiring is locally bounded.

Theorem 1. *Let \mathcal{P} be a weighted pushdown system over a locally bounded idempotent semiring indexed by \mathcal{M}_Γ and $\mathcal{A}_{\text{pre}^*}$ be a weighted automaton obtained by the saturation procedure. Then, we have $p \xrightarrow[\mathcal{A}_{\text{pre}^*}]{\gamma|a} p'$ for $a = \delta_{\mathcal{P}}(p, \gamma, p')$.*

As a corollary, we have $p \xrightarrow[\mathcal{A}_{\text{pre}^*}]{w|a} p'$ for $a = \delta_{\mathcal{P}}(p, w, p')$. The theorem is proved from the following two lemmas.

Lemma 1. *If $p \xrightarrow[\mathcal{P}]{w/\epsilon|a} p'$, then $p \xrightarrow[\mathcal{A}_{\text{pre}^*}]{w|a'} p'$ and $a \sqsubseteq_{w/\epsilon} a'$ for some a' .*

Let \mathcal{A}_{i+1} be a weighted automaton obtained by applying the saturation rule once to \mathcal{A}_i .

Lemma 2. *If $p \xrightarrow[\mathcal{A}_i]{\gamma|a} p'$, then $a \sqsubseteq_{\gamma/\epsilon} \delta_{\mathcal{P}}(p, \gamma, p')$.*

5.3 Reachability to a Regular Set of Configurations

In previous works of the reachability analysis of pushdown systems, it is common to consider the reachability problem to a regular set of configurations. For a weighted pushdown automaton over an indexed semiring, this problem must be generalized for a regular set with weight represented by a weighted automaton.

Let us consider an indexed semiring \mathcal{S} over \mathcal{M}_Γ and a weighted pushdown system \mathcal{P} over \mathcal{S} . We also consider a weighted automaton \mathcal{A} over the restriction of \mathcal{S} on $\{w/\epsilon \mid w \in \Gamma^*\}$ with the initial states q_0 and the set of final states F . Then, the generalized reachability problem to a regular set of configuration $\{\langle p', w' \rangle \mid w' \text{ is accepted by } \mathcal{A}\}$ is to compute the following function⁵.

⁵ For simplicity, we consider the set of configurations whose state is p' . It is easy to extend the discussion for the general case.

$$\delta_{\mathcal{P}, \mathcal{A}}(p, w, p') = \bigoplus_{q \in F} \{a \otimes_{\sigma, w'/\epsilon} a' \mid p \xrightarrow[\mathcal{P}]{\sigma \mid a} p', q_0 \xrightarrow[\mathcal{A}]{w' \mid a'} q, \text{ and } \sigma \cdot w'/\epsilon = w/\epsilon\}$$

This function can be computed by applying the saturation procedure to the pushdown system \mathcal{P}' obtained by combining \mathcal{P} and \mathcal{A} with the identification of p' and q_0 . This corresponds to the saturation procedure using \mathcal{P} -automata.

The condition $\sigma \cdot w'/\epsilon = w/\epsilon$ above is equivalent to $\sigma \leq w/w'$. Furthermore, if the indexed semiring is equipped with the conversion functions $\uparrow_{\sigma_1, \sigma_2}$, we have the following.

$$\begin{aligned} &= \bigoplus_{q \in F} \{\uparrow_{\sigma, w/w'}(a) \otimes_{w/w', w'/\epsilon} a' \mid p \xrightarrow[\mathcal{P}]{\sigma \mid a} p', q_0 \xrightarrow[\mathcal{A}]{w' \mid a'} q, \text{ and } \sigma \leq w/w'\} \\ &= \bigoplus_{q \in F} \{a \otimes_{w/w', w'/\epsilon} a' \mid \langle p, w \rangle \xrightarrow[\mathcal{P}]{a} \langle p', w' \rangle \text{ and } q_0 \xrightarrow[\mathcal{A}]{w' \mid a'} q\} \end{aligned}$$

5.4 Constructing a Semiring from an Indexed Semiring over Stack Signatures

We show that an ordinary semiring can be constructed from a semiring indexed by the ordered monoid of stack signatures. However, the semiring obtained by the construction is not bounded in general even for a locally bounded indexed semiring. Thus, the standard framework of the reachability analysis of weighted pushdown systems cannot guarantee termination of the saturation procedure. Although a similar construction appears in [Suw09], the definition of \oplus differs from ours and it fails to satisfy the distributivity of \otimes over \oplus .

In this section, we assume that D_\top is a singleton set and $D_\top = \{\bullet\}$.

Theorem 2. *Let $\mathcal{S} = \langle \{D_\sigma\}, \{\oplus_\sigma\}, \{\otimes_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, 1_\mathcal{S}, \uparrow_{\sigma, \sigma'} \rangle$ be a semiring indexed by the ordered monoid \mathcal{M}_Γ and $D = \bigcup_{\sigma \in \mathcal{M}_\Gamma} \{(\sigma, a) \mid a \in D_\sigma\} \cup \{\perp\}$. Then, $\langle D, \oplus, \otimes, \perp, 1 \rangle$ defined as follows forms a semiring.*

- 1 is $(\epsilon/\epsilon, 1_\mathcal{S})$.
- \oplus is defined by $\perp \oplus x = x = x \oplus \perp$ for all $x \in D$ and

$$(\sigma_1, a) \oplus (\sigma_2, b) = \begin{cases} (\sigma_1, a \oplus_{\sigma_1} \uparrow_{\sigma_2, \sigma_1}(b)) & \text{if } \sigma_2 \leq \sigma_1 \\ (\sigma_2, \uparrow_{\sigma_1, \sigma_2}(a) \oplus_{\sigma_2} b) & \text{if } \sigma_1 \leq \sigma_2 \\ (\top, \bullet) & \text{otherwise} \end{cases}$$

- \otimes is defined by $(\sigma_1, a) \otimes (\sigma_2, b) = (\sigma_1 \sigma_2, a \otimes_{\sigma_1, \sigma_2} b)$ and $x \otimes \perp = \perp = \perp \otimes x$ for all $x \in D$.

Suwimonteerabuth did not consider the partial order on stack signatures and defined the addition of the semiring \oplus' in the following manner [Suw09].

$$(\sigma_1, a) \oplus' (\sigma_2, b) = \begin{cases} (\sigma_1, a \oplus_{\sigma_1} b) & \text{if } \sigma_1 = \sigma_2 \\ (\top, \bullet) & \text{otherwise} \end{cases}$$

However, \otimes does not distribute over \oplus' , and thus fails to form a semiring.

$$((\epsilon/\epsilon, a) \oplus' (\gamma/\gamma, b)) \otimes (\gamma/\gamma, c) = (\top, \bullet) \otimes (\gamma/\gamma, c) = (\top, \bullet)$$

$$\begin{aligned} ((\epsilon/\epsilon, a) \otimes (\gamma/\gamma, c)) \oplus' ((\gamma/\gamma, b) \otimes (\gamma/\gamma, c)) \\ = (\gamma/\gamma, a \otimes_{\epsilon/\epsilon, \gamma/\gamma} c) \oplus' (\gamma/\gamma, b \otimes_{\gamma/\gamma, \gamma/\gamma} c) \\ = (\gamma/\gamma, a \otimes_{\epsilon/\epsilon, \gamma/\gamma} c \oplus_{\gamma/\gamma} b \otimes_{\gamma/\gamma, \gamma/\gamma} c) \end{aligned}$$

It should be noted that the semiring constructed in Theorem 2 is not bounded as the following sequence shows.

$$(\epsilon/\epsilon, a) \sqsubset (\gamma/\gamma, \uparrow_\gamma(a)) \sqsubset (\gamma\gamma/\gamma\gamma, \uparrow_{\gamma\gamma}(a)) \sqsubset \dots$$

This is one of the reasons why we refine the formulation of the reachability analysis of weighted pushdown systems in this paper.

6 Simplified Structure: Multiplication on Strictly Compatible Signatures

An indexed semiring has a multiplication indexed by two stack signatures. However, it is often simpler to consider and implement a restricted multiplication defined only for strictly compatible signatures. We show that an indexed semiring over the ordered monoid of stack signatures can be constructed from such a structure.

We introduce *weight structures* that have a restricted multiplication $\odot_{\sigma_1, \sigma_2}$ for strictly compatible σ_1 and σ_2 .

Definition 9 (Weight Structure). *A weight structure \mathcal{W} over stack alphabet Γ is $\langle \{D_\sigma\}, \{\oplus_\sigma\}, \{\odot_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, \{1_w\}, \{\uparrow_{\sigma, \sigma'}\} \rangle$ such that*

- D_σ is a set for each proper stack signature σ .
- $\langle D_\sigma, \oplus_\sigma, 0_\sigma \rangle$ is a commutative monoid for proper stack signature σ .
- $\odot_{\sigma_1, \sigma_2}$ is an associative binary operation of $D_{\sigma_1} \times D_{\sigma_2} \rightarrow D_{\sigma_1 \sigma_2}$ for strictly compatible signatures σ_1 and σ_2 .
- $1_w \in D_{w/w}$ is an indexed unit of $\odot_{\sigma_1, \sigma_2}$: $a \odot_{w'/w, w/w} 1_w = a$ and $1_w \odot_{w/w, w/w'} b = b$.
- 0_σ is an annihilator with respect to $\odot_{\sigma, \sigma'}$: $0_{\sigma_1} \odot_{\sigma_1, \sigma_2} a = 0_{\sigma_1 \sigma_2} = b \odot_{\sigma_1, \sigma_2} 0_{\sigma_2}$.
- \odot distributes over \oplus .

$$\begin{aligned} (a \oplus_{\sigma_1} b) \odot_{\sigma_1, \sigma_2} c &= (a \odot_{\sigma_1, \sigma_2} c) \oplus_{\sigma_1 \sigma_2} (b \odot_{\sigma_1, \sigma_2} c) \\ a \odot_{\sigma_1, \sigma_2} (b \oplus_{\sigma_2} c) &= (a \odot_{\sigma_1, \sigma_2} b) \oplus_{\sigma_1 \sigma_2} (a \odot_{\sigma_1, \sigma_2} c) \end{aligned}$$

- $\uparrow_{\sigma, \sigma'}$ is a conversion function of $D_\sigma \rightarrow D_{\sigma'}$ for $\sigma \leq \sigma'$ such that
 - $\uparrow_{\sigma, \sigma} = \text{id}$ and $\uparrow_{\sigma, \sigma''} = \uparrow_{\sigma', \sigma''} \circ \uparrow_{\sigma, \sigma'}$ for all $\sigma \leq \sigma' \leq \sigma''$.
 - $\uparrow_{\sigma, \sigma'}(0_\sigma) = 0_{\sigma'}$ and $\uparrow_{\sigma, \sigma'}(a \oplus b) = \uparrow_{\sigma, \sigma'}(a) \oplus \uparrow_{\sigma, \sigma'}(b)$
 - $\uparrow_{w_1/w_2, w_1 w'/w_2 w'}(a \odot b) = \uparrow_{w_1/w, w_1 w'/w w'}(a) \odot \uparrow_{w/w_2, w w'/w_2 w'}(b)$
 - $\uparrow_{w/w, w w'/w w'}(1_w) = 1_{w w'}$

We show that the multiplication of an indexed semiring over \mathcal{M}_Γ can be obtained from that of a weight structure. Let $\{D'_\sigma\}$ be a family of $\{D_\sigma\} \cup \{D_\top\}$ where $D_\top = \{\bullet\}$. Then, the multiplication on D'_σ is defined as follows.

$$x \otimes_{\sigma_1, \sigma_2} y = \begin{cases} \uparrow_{\sigma_1, \sigma'_1}(x) \odot_{\sigma'_1, \sigma_2} y & \text{if } \sigma_1 \leq \sigma'_1 \text{ and } \sigma'_1 \text{ is strictly compatible with } \sigma_2 \\ x \odot_{\sigma_1, \sigma'_2} \uparrow_{\sigma_2, \sigma'_2}(y) & \text{if } \sigma_2 \leq \sigma'_2 \text{ and } \sigma_1 \text{ is strictly compatible with } \sigma'_2 \\ \bullet & \text{otherwise} \end{cases}$$

The other operations are extended for D_\top in a straightforward manner. Then, we obtain a semiring indexed by the ordered monoid \mathcal{M}_Γ .

Theorem 3. *Let $\langle \{D_\sigma\}, \{\oplus_\sigma\}, \{\odot_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, \{1_w\}, \{\uparrow_{\sigma, \sigma'}\} \rangle$ be a weight structure. Then, $\langle \{D'_\sigma\}, \{\oplus_\sigma\}, \{\otimes_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, 1_\epsilon, \{\uparrow_{\sigma, \sigma'}\} \rangle$ is an indexed semiring over an ordered monoid \mathcal{M}_Γ .*

7 Applications

7.1 Encoding of Local Variables into Weight

Suwimonteerabuth applied a semiring similar to one constructed from an indexed semiring to encode local variables of a recursive program into weight [Suw09]. Although his implementation worked without any problem, it is actually not in the standard framework of weighted pushdown systems because the semiring is not bounded.

We show that his encoding can be formulated more naturally with an indexed semiring. In order to simplify our presentation, we give an encoding of a pushdown system into a weighted pushdown system with a singleton stack alphabet. Since local variables can be encoded into stack alphabet, the same approach can be applied for the encoding of local variables.

Let us consider a singleton stack alphabet $\Gamma' = \{\#\}$. We write m/n for a stack signature $\#^m/\#^n$. We will construct a weight structure to translate pushdown systems over stack alphabet Γ . We define weight structure $\mathcal{W}_\Gamma = \langle \{D_\sigma\}, \{\oplus_\sigma\}, \{\odot_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, \{1_w\}, \{\uparrow_{\sigma_1, \sigma_2}\} \rangle$ as follows.

- $D_{m/n}$ is the set of relations over Γ^m and Γ^n : $D_{m/n} = 2^{\Gamma^m \times \Gamma^n}$.
- $0_{m/n} = \emptyset$ and $1_m = \{(x, x) \mid x \in \Gamma^m\}$.
- $R_1 \odot_{l/m, m/n} R_2$ is a composition of relations: $R_1 \circ R_2$ where $R_1 \subseteq \Gamma^l \times \Gamma^m$ and $R_2 \subseteq \Gamma^m \times \Gamma^n$.
- $R_1 \oplus_{m/n} R_2$ is the union of two relations R_1 and R_2 : $R_1 \cup R_2$ where $R_1, R_2 \subseteq \Gamma^m \times \Gamma^n$.
- $\uparrow_{l/m, l+1/m+1}$ extends the domain of a relation and is defined by

$$\uparrow_{l+1/m+1}(R) = \{((x, z), (y, z)) \mid (x, y) \in R \wedge z \in \Gamma\}$$

where we consider $\Gamma^{k+1} = \Gamma^k \times \Gamma$.

It is straightforward to show this structure forms a weight structure. Furthermore, it induces a locally bounded indexed semiring because $D_{m/n}$ is the power set of a finite set and ordered by the set inclusion.

We show how to simulate a pushdown system $\mathcal{P} = \langle P, \Gamma, \Delta \rangle$ by a weighted pushdown system \mathcal{P}' over the weight structure \mathcal{W}_Γ . Let \mathcal{P}' be $\langle P, \Gamma', \Delta' \rangle$ such that

$$(q, \#, q', \#^m, a) \in \Delta' \quad \text{iff} \quad (q, \gamma, q', w) \in \Delta$$

where $|w| = m$ and $a = \{(\gamma, w)\}$. Then, \mathcal{P} and \mathcal{P}' are equivalent in the following sense:

$$p \xrightarrow[\mathcal{P}]{w/w'} p' \quad \Longleftrightarrow \quad p \xrightarrow[\mathcal{P}']{m/m'} \bigwedge^a p' \wedge (w, w') \in a$$

where $m = |w|$ and $m' = |w'|$. Then, we can check the reachability in \mathcal{P} by checking that in \mathcal{P}' .

7.2 Reachability Analysis of Conditional Pushdown Systems

Esparza *et al.* introduced pushdown systems with checkpoints that have the ability to inspect the whole stack contents against a regular language [EKS03]. Li and Ogawa reformulated their definition and called them conditional pushdown systems [LO10]. We review conditional pushdown systems and then formulate the reachability analysis in our previous work [MM12] as that of weighted pushdown systems.

Definition 10. A conditional pushdown system \mathcal{P} is a structure $\langle P, \Gamma, \Delta \rangle$ where P is a finite set of states, Γ is a stack alphabet, and $\Delta \subseteq P \times \Gamma \times P \times \Gamma^* \times \text{Reg}(\Gamma)$ is a set of transitions where $\text{Reg}(\Gamma)$ is the set of regular languages over Γ .

We write $\langle p, \gamma \rangle \xrightarrow{R} \langle p', w \rangle$ if $\langle p, \gamma, p', w, R \rangle \in \Delta$ as weighted pushdown systems. The transition relation of a conditional pushdown system is defined as follows.

- $\langle p, w \rangle \Longrightarrow \langle p, w \rangle$.
- $\langle p, \gamma w' \rangle \Longrightarrow \langle p', ww' \rangle$ if $\langle p, \gamma \rangle \xrightarrow{R} \langle p', w \rangle$ and $w' \in R$.
- $\langle p, w \rangle \Longrightarrow \langle p', w' \rangle$ if $\langle p, w \rangle \Longrightarrow \langle p'', w'' \rangle$ and $\langle p'', w'' \rangle \Longrightarrow \langle p', w' \rangle$.

In the second case above, the transition can be taken only when the current stack contents excluding its top is included in the regular language R given as the condition of the rule.

We show that the transition of a conditional pushdown system can be simulated by that of a weighted pushdown system without conditional rules. Let us design a weight structure for this simulation: we use the same domain for all proper stack signatures σ : $D_\sigma = 2^{\Gamma^*}$. Then, the weight structure $(\{D_\sigma\}, \{\oplus_\sigma\}, \{\odot_{\sigma_1, \sigma_2}\}, \{0_\sigma\}, \{1_w\}, \{\uparrow_{\sigma, \sigma'}\})$ is given as follows.

- $0_\sigma = \emptyset$ and $1_w = \Gamma^*$.
- $a \oplus_\sigma b = a \cup b$.

- $a \odot_{\sigma_1, \sigma_2} b = a \cap b$ for strictly compatible signatures σ_1 and σ_2 .
- $\uparrow_{w_1/w_2, w_1 w/w_2 w}(a) = w^{-1}a$ where $w^{-1}a$ is left quotient defined by $w^{-1}a = \{w' \mid ww' \in a\}$.

It is clear that this structure is a weight structure from the basic properties of left quotient and set operations. Then, for a conditional pushdown system \mathcal{P} we obtain a weighted pushdown system \mathcal{P}' over the indexed semiring above by considering a conditional transition rule $\langle p, \gamma \rangle \xrightarrow{R} \langle p', w \rangle$ as a weighted one.

A conditional pushdown system \mathcal{P} is simulated by a weighted pushdown system \mathcal{P}' in the following sense.

- If $\langle p_1, w_1 \rangle \xRightarrow{\mathcal{P}} \langle p_2, w_2 \rangle$, then there exist w and σ such that $p_1 \xRightarrow[\mathcal{P}']{\sigma|a} p_2$, $w \in a$, and $\uparrow_w(\sigma) = w_1/w_2$.
- If $p_1 \xRightarrow[\mathcal{P}']{w_1/w_2|a} p_2$ and $w \in a$, $\langle p_1, w_1 w \rangle \xRightarrow{\mathcal{P}} \langle p_2, w_2 w \rangle$.

Please note that this weight structure is not locally bounded because 2^{Γ^*} is not bounded with respect to the set inclusion. However, D_σ can be restricted to the set $D \subseteq 2^{\Gamma^*}$ inductively defined as follows.

- $\emptyset \in D$ and $\Gamma^* \in D$.
- $R \in D$ if $\langle p, \gamma \rangle \xrightarrow{R} \langle p', w \rangle$ for some p, γ, p', w .
- $R_1 \cap R_2 \in D$ and $R_1 \cup R_2 \in D$ if $R_1 \in D$ and $R_2 \in D$.
- $w^{-1}R \in D$ if $R \in D$ and $w \in \Gamma^*$.

This set D is finite because the set of transitions is finite, there are finitely many languages obtained from each regular language with left quotient, and left quotient distributes over union and intersection. Thus, we obtain a locally bounded indexed semiring by using D . This gives the algorithm of the backward reachability analysis for conditional pushdown systems that we used to analyse the HTML5 parser specification [MM12].

8 Related Work

An automaton over a monoid M is called a generalized M -automaton by Eilenberg [Eil74]. The textbook of Sakarovitch discusses automata over several classes of monoids including free groups and commutative monoids [Sak09]. As far as we know, this paper is the first work that discusses the reachability analysis of pushdown systems by considering them as automata over the monoid of stack signatures.

Let us consider a paired alphabet $\tilde{\Gamma} = \Gamma \cup \overline{\Gamma}$ where $\overline{\Gamma} = \{\bar{a} \mid a \in \Gamma\}$. Letters γ and $\bar{\gamma}$ correspond to a push and a pop of γ , respectively. Then, the monoid \mathcal{M}_Γ is closely related to the monoid over $\tilde{\Gamma}^*$ obtained by Shamir congruence [Sha67], which is generated by $\gamma\bar{\gamma} = \epsilon$. If we add the relation $\gamma\bar{\gamma}' = \top$ for $\gamma \neq \gamma'$, then the reduced form of a word over $\tilde{\Gamma}$ has the following form: $\overline{w_1}w_2$ or \top . If we write w_1/w_2^R for $\overline{w_1}w_2$, we obtain a stack signature⁶.

⁶ w_2^R is the reverse of w_2 .

Esparza *et al.* showed that conditional pushdown systems can be translated to ordinary pushdown systems [EKS03]. Hence, the reachability can be decided via the translation. However, it is not practical to apply the translation because of exponential blowup of the size of pushdown systems. The algorithm formulated in Section 7.2 as the reachability analysis of weighted pushdown systems has also an exponential complexity. However, it avoids the exponential blowup by the translation before applying the reachability analysis and worked well for the analysis of the HTML5 parser specification.

9 Conclusions

We have introduced the monoid of stack signatures to treat pushdown systems as automata over the monoid. Then, weighted pushdown systems are generalized by adopting a semiring indexed by stack signatures as weight. This generalization makes it possible to relax the restriction of boundedness and extend the applications of the reachability analysis of weighted pushdown systems.

The indexed semirings for the two applications in this paper are given through weight structures. We consider that it is simpler to construct and implement indexed semirings through weight structures than to directly construct them. However, we are not completely satisfied with the formulation of weight structures because their definition looks rather ad hoc mathematically. We would like to investigate more abstract notion corresponding to weight structures.

Acknowledgments. I would like to thank Stefan Schwoon for the inspiring discussions and suggestions. Schwoon informed me of the work of Suwimon-teerabuth on the encoding of local variables into weight. This work has been partially supported by JSPS Grant-in-Aid for Science Research (C) 24500028 and the Kayamori Foundation of Informational Science Advancement.

References

- [BEM97] Bouajjani, A., Esparza, J., Maler, O.: Reachability Analysis of Pushdown Automata: Application to Model-Checking. In: Mazurkiewicz, A., Winkowski, J. (eds.) CONCUR 1997. LNCS, vol. 1243, pp. 135–150. Springer, Heidelberg (1997)
- [Eil74] Eilenberg, S.: Automata, Languages, and Machines, vol. A. Academic Press (1974)
- [ÉK09] Ésik, Z., Kuich, W.: Finite automata. In: Droste, M., Kuich, W., Vogler, H. (eds.) Handbook of Weighted Automata, ch. 3, pp. 69–104. Springer (2009)
- [EKS03] Esparza, J., Kucera, A., Schwoon, S.: Model checking LTL with regular valuations for pushdown systems. Information and Computation 186(2), 355–376 (2003)
- [FWW97] Finkel, A., Willems, B., Wolper, P.: A direct symbolic approach to model checking pushdown systems. In: INFINITY 1997. ENTCS, vol. 9, pp. 27–39 (1997)

- [LO10] Li, X., Ogawa, M.: Conditional weighted pushdown systems and applications. In: Proceedings of the 2010 ACM SIGPLAN Workshop on Partial Evaluation and Program Manipulation, pp. 141–150 (2010)
- [Min07] Minamide, Y.: Verified Decision Procedures on Context-Free Grammars. In: Schneider, K., Brandt, J. (eds.) TPHOLs 2007. LNCS, vol. 4732, pp. 173–188. Springer, Heidelberg (2007)
- [MM12] Minamide, Y., Mori, S.: Reachability Analysis of the HTML5 Parser Specification and Its Application to Compatibility Testing. In: Giannakopoulou, D., Méry, D. (eds.) FM 2012. LNCS, vol. 7436, pp. 293–307. Springer, Heidelberg (2012)
- [MT06] Minamide, Y., Tozawa, A.: XML Validation for Context-Free Grammars. In: Kobayashi, N. (ed.) APLAS 2006. LNCS, vol. 4279, pp. 357–373. Springer, Heidelberg (2006)
- [RSJM05] Reps, T., Schwoon, S., Jha, S., Melski, D.: Weighted pushdown systems and their application to interprocedural dataflow analysis. *Science of Computer Programming* 58, 206–263 (2005)
- [Sak09] Sakarovitch, J.: *Elements of Automata Theory*. Cambridge University Press (2009)
- [Sha67] Shamir, E.: A representation theorem for algebraic and context-free power series in non commuting variables. *Information and Control* 11(1/2), 239–254 (1967)
- [Suw09] Suwimonteerabuth, D.: *Reachability in Pushdown Systems: Algorithms and Applications*. PhD thesis, Technischen Universität München (2009)
- [TM07] Tozawa, A., Minamide, Y.: Complexity Results on Balanced Context-Free Languages. In: Seidl, H. (ed.) FOSSACS 2007. LNCS, vol. 4423, pp. 346–360. Springer, Heidelberg (2007)