A Note on Partially Ordered Tree Automata

Pierre-Cyrille Héam

A Note on Partially Ordered Tree Automata

P.-C. Héam

Projet INRIA-CASSIS
Laboratoire d'Informatique de l'Université de Franche-Comté
16 route de Gray, 25030 Besançon Cedex. FRANCE.

Abstract

A recent paper by Bouajjani, Muscholl and Touili shows that the class of languages accepted by partially ordered word automata (or equivalently accepted by Σ_2-formulae) is closed under semi-commutation and it suggested the following open question: can we extend this result to tree languages? This problem can be addressed by proving 1) that the class of tree regular languages accepted by Σ_2 formulae is strictly included in the class of languages accepted by partially ordered automata, and 2) that Bouajjani and the others results can't be extended to tree.

Keywords: Formal languages, formal methods, tree automata.

We assume the reader is familiar with the basic notions on terms and tree automata. For general reference see [6].

1 Introduction

A word automaton is partially ordered if its set of states can be partially ordered by a relation $\leq$ such that if (p, a, q) is a transition, then $p \leq q$. Word languages accepted by partially ordered automata accepts several characterisations.

Proposition 1 *The following propositions are equivalent:*

i) L is accepted by a partially ordered word-automaton.
ii) L is definable by a Σ_2 formula.
iii) The ordered syntactic monoid of L satisfies $x^\omega y x^\omega \leq x^\omega$ for every x and y generated by the same set of letters.

The notion of Σ_2 formula will be presented in the next section. For more information on a logical approach of word regular languages see the survey [12,17]. We will not study point *iii*) of the above proposition which is only given for the reader's information. Indeed, this notion cannot be easily extended to tree languages and is out of the scope of this paper. See [13] for more details and notice that efficient testing algorithms for this class of languages comes from the monoid's characterisation.

A semi-commutation relation R_S on words generated by a subset $S \subseteq \Sigma \times \Sigma$ (where Σ is the alphabet), is the subset of $\Sigma^* \times \Sigma^*$ such that $(u, v) \in R_s$ if and only if there exist two words x and y and two letters a and b such that $u = xaby$, $v = xbay$ and $(a, b) \in S$. The reflexive-transitive closure of R_S is denoted by R_S^*. The following result is proved in [3] (an efficient and extended automata based approach is developed in [5]).

Theorem 2 *For every language L accepted by a partially ordered automaton and every semi-commutation relation R_S, the set $\{v \mid \exists u \in L,\ (u, v) \in R_S^*\}$ is accepted by a computable partially ordered automaton.*

This result is suitable for algorithmic issues related to regular-model-checking or HMSC verification [3]. Partially ordered word automata were also intensively studied in the context of regular languages theory [14,1]. In the conclusion of [3], authors wonder whether Theorem 2 can be extended to tree data structures. But, in this paper, we negatively answer this question.

In this paper $\mathcal{F}$ denotes a finite set of symbols with arity. For the simplicity of notation, if $a_0, \ldots, a_n$ are symbols of arity 1 and if $\#$ is a symbol of arity 0, the term $a_0(\ldots(a_n(\#))\ldots)$ is denoted by $a_0 \ldots a_n(\#)$. For every subset K of terms, K^c denotes the set of terms that are not in K. Throughout this paper, we need the following results that can be easily proved using classical automata constructions.

Lemma 3 *The class of languages accepted by partially ordered automata is closed under union and intersection. The class of languages accepted by deterministic partially ordered automata is closed under union, intersection and complement.*

This paper is organised as follows: in Section 2 is dedicated to prove that the class of tree regular languages accepted by Σ_2 formulae is strictly included in the class of languages accepted by partially ordered automata. In Section 3 it is proved that there exists a tree language accepted by a Σ_2 formula whose closure under a semi-commutation relation is not regular.

2 Partially ordered Tree Automata and Σ_2 formulae

For a general approach on WSkS logic see [6, Section 3.3] or [17, Chapter 7]. This section is dedicated to prove that the class of tree languages accepted by Σ_2 formulae is strictly included in the class of languages accepted by partially ordered tree automata.

We recall in Section 2.1 some basic logical definitions and notions. In Section 2.2 we point out a language accepted by a partially ordered automaton but not by any Σ_2 formula. Section 2.3 is dedicated to prove that every language accepted by a Σ_2 formula is also accepted by a partially ordered automaton.

2.1 Logic on Terms

The *Atomic formulae* on F are defined as follows:

- $(p_1 = p_2)$, $(p_1 <_i p_2)$ and $R_f(p_1)$, where p_1 and p_2 are elements of $\mathbb{N}^* \cup \mathcal{X}$, $f \in \mathcal{F}$, and $i \in \mathbb{N}$, are atomic formulae.
- If φ_1 and φ_2 are atomic formulae, then $\neg\varphi_1$, $(\varphi_1 \vee \varphi_2)$ are atomic formulae.

In the above definition, if p_1, p_2 and p are required to belong to $\mathcal{X}$, we obtain the subclass of atomic formulae called *Constant-free atomic formulae.* If the set of variables occurring a formula φ is included in $\{x_1, \ldots, x_n\}$, then we also write it $\varphi(x_1, \ldots, x_n)$. For each term t, we inductively define a valuation function ν_t that maps each formula to 0 or 1:

- $\nu_t((p_1 = p_2)) = 1$ iff p_1 and p_2 are positions of t and if $p_1 = p_2$.
- $\nu_t((p_1 <_i p_2)) = 1$ iff $p_1.i$ and p_2 are positions of t and if $p_1.i$ is a prefix of p_2.
- $\nu_t(R_f(p)) = 1$, iff p is a position of t and $t(p) = f$.
- $\nu_t(\neg\varphi) = 1 - \nu_t(\varphi)$.
- $\nu_t((\varphi \vee \varphi_2)) = \max(\nu_t(\varphi_1), \nu_t(\varphi_2))$.

Classically, we say that t satisfies φ, denoted $t \models \varphi$, if $\nu_t(\varphi) = 1$. A formula of the form $\exists x_1 \ldots \exists x_n \forall y_1 \ldots \forall y_k\ \varphi(x_1, \ldots, x_n, y_1, \ldots, y_k)$, where φ is a constant-free atomic formula is called a Σ_2 formula. A Σ_2 formula with no $\exists$ quantifiers is called a Π_1 formula, and a Σ_2 formula with no $\forall$ quantifiers is called a Σ_1 formula. A term t satisfies a Σ_2 formula iff there exist positions $p_1, \ldots, p_n$ of t such that for all positions $p'_1, \ldots, p'_k$ of t, $t \models \varphi(p_1, \ldots, p_n, p'_1, \ldots, p'_k)$. In this case we say that t is accepted by φ. We denote by $L(\varphi)$ the set of terms accepted by φ, which is always regular [15].

2.2 Partially Ordered Automata and Σ_2 formulae do not have the same Expressivity

If we restrict the alphabet to unary or constant symbols, the class of languages accepted by Σ_2 formulae is exactly the class of languages accepted by partially ordered automata [16,1]. Moreover, it is also known that in the word case, the class of languages accepted by Σ_1 formulae, accepts shuffle ideals [16], that is languages which are a finite union of languages of the form $A^*a_1A^* \dots A^*a_nA^*$ where A is the alphabet and the a_i's are letters. It is known [9] that a (finite or infinite) union of shuffle ideal is a shuffle ideal too. Moreover, shuffle ideals are also exactly the class of word languages L satisfying the following property: if u and v are words and if uv is in L, then for every letter a, the word uav is also in L. For a combinatorial approach of shuffle ideals, the reader is referred to [10] and for a regular approach to [8]. Now we can prove the following proposition.

Proposition 4 *There exists a language accepted by a deterministic partially ordered tree automaton but not by any Σ_2 formula.*

Proof. Let $\mathcal{F} = \{f, a, b, \#\}$, f is of arity 2, a and b are both of arity 1 and $\#$ is of arity 0. We consider the partially ordered automaton $\mathcal{A}$ whose states are q_0, q_1 (with $q_0 \leq q_1$), whose transitions are $\# \to q_0$, $a(q_0) \to q_0$, $b(q_0) \to q_0$, $a(q_0) \to q_1$, $f(q_1, q_1) \to q_1$, and whose final state is $\{q_1\}$.

Now we will prove that there is no Σ_2 formula accepting $L(\mathcal{A})$. Indeed assume that there exists a Σ_2 formula $\phi = \exists x_1 \dots \exists x_n \forall y_1 \dots \forall y_k\, \varphi(x_1, \dots, x_n, y_1, \dots, y_k)$ accepting $L(\mathcal{A})$.

(1) Firstly, let $u \in a\{a, b\}^*$. We first define a sequence of terms $(t_i^u)_{i\geq 1}$ by $t_1^u = f(u(\#), u(\#))$ and $t_{i+1}^u = f(t_i^u, t_i^u)$. Each t_i^u is in $L(\mathcal{A})$. Let ℓ be an integer such that $\ell > \log_2 n$. Since $t_\ell \models \phi$, there exist positions $p_1, \dots, p_n$ of t_ℓ such that $t_\ell^u \models \forall y_1 \dots \forall y_k \varphi'(y_1, \dots, y_k)$ where $\varphi'(y_1, \dots, y_k)$ denotes $\varphi(p_1, \dots, p_n, y_1, \dots, y_k)$. Now, since t_ℓ^u admits $2^\ell > n$ positions of length ℓ, there exists a position p of t_ℓ^u of length ℓ which is not a prefix of any p_i's. Now let

$$K_u = \{v \in \{a, b\}^* \mid t_\ell^u[v(\#)]_p \models \forall y_1 \dots \forall y_k \varphi'(y_1, \dots, y_k)\},$$

where $t_\ell^u[v(\#)]_p$ denotes the term obtained by replacing in t_ℓ the subterm at position p by $v(\#)$. We claim that if $v_1v_2 \notin K_u$, then $v_1av_2 \notin K$ and $v_1bv_2 \notin K_u$. Assuming that $t_\ell^u[v_1v_2(\#)]_p \models \neg\varphi'(p'_1, \dots, p'_k)$, let $p''_i = p'_i$ if $p.1^r$ (where r is the length of v_1) is not a prefix of p'_i and let $p''_i = p'_i.1$ otherwise. By a direct induction, one has

$$\nu_{t_\ell^u[v_1v_2(\#)]_p}(\varphi'(p''_1, \dots, p''_k)) = \nu_{t_\ell^u[v_1xv_2(\#)]_p}(\varphi'(p'_1, \dots, p'_k)),$$

with $x = a$ or $x = b$, proving the claim. It follows that $\{a,b\}^* \setminus K_u$ is a shuffle ideal.

(2) Secondly, let $H = \{a,b\}^* \setminus \bigcup_{u \in a\{a,b\}^*} K_u$. The set H is a shuffle ideal. Moreover, by construction of K_u, for every u in $a\{a,b\}^*$, $u \in K_u$. Furthermore, since ϕ accepts $L(\mathcal{A})$, for every u, $K_u \subseteq a\{a,b\}^*$. It follows that $H = \{a,b\}^* \setminus a\{a,b\}^* = b\{a,b\}^* \cup \{\varepsilon\}$, which is not a shuffle ideal. This is a contradiction which very much completes the proof. □

2.3 Σ_2 Formulae are less Expressive than Partially Ordered Automata

In this section we prove that the class of languages accepted by Σ_2 formulae is included in the class of languages accepted by partially ordered automata. First of all, by using a construction translating a formula into a tree regular language. This construction (and its correctness) is, for instance, presented in [17, Chapter 7]. Next by giving some technical lemmas in order to prove the motioned result.

For every positive integer ℓ, the set of terms t over $\mathcal{F} \times \{0,1\}^\ell$ (arities are the same as in $\mathcal{F}$) such that for every $i \in \{1, \ldots, \ell\}$ there exists a unique position p, denoted $\overline{t}^{\ell,i}$, or $\overline{t}^{i}$ if there is no ambiguity on ℓ, such that $t(p) \in \mathcal{F} \times \{0,1\}^{i-1}1\{0,1\}^{n-i}$, is denoted by $\mathcal{K}_\ell$. Informally it means that if the i-th components of symbols at position p_1 and p_2 are equal to 1, then $p_1 = p_2$. Let π_i $(1 \leq i \leq \ell)$ be the function from $\mathcal{F} \times \{0,1\}^\ell$ into $\mathcal{F} \times \{0,1\}^{\ell-1}$ such that $\pi_i((f,u)) = (f,v)$ where v is deduced from u by deleting the i-th component. The projection π_i is naturally extended to terms t in the following way: $\pi_i(t)$ and t have the same set of positions and for every position p, $\pi_i(t)(p) = \pi_i(t(p))$.

We inductively defined the languages for any formula whose set of variables is $\{x_1, \ldots, x_\ell\}$, by

- $\mathcal{K}_\ell(R_f(x_i)) = \{t \in \mathcal{K}_\ell \mid \exists u \in \{0,1\}^\ell t(\overline{t}^{\ell,i}) = (f,u)\}$, for every $f \in \mathcal{F}$,
- $\mathcal{K}_\ell(x_i = x_j) = \{t \in \mathcal{K}_\ell \mid \overline{t}^{\ell,i} = \overline{t}^{\ell,j}\}$,
- $\mathcal{K}_\ell(x_i <_r x_j) = \{t \in \mathcal{K}_\ell \mid \overline{t}^{\ell,i} <_r \overline{t}^{\ell,j}\}$,
- $\mathcal{K}_\ell(\neg\varphi) = \mathcal{K}_\ell \setminus \mathcal{K}_\ell(\varphi)$, $\mathcal{K}_\ell(\varphi_1 \vee \varphi_2) = \mathcal{K}_\ell(\varphi_1) \cup \mathcal{K}_\ell(\varphi_2)$,
- $\mathcal{K}_\ell(\exists x_i \varphi) = \pi_i(\mathcal{K}_\ell(\varphi))$,
- $\mathcal{K}_\ell(\forall x_i \varphi) = \mathcal{K}_\ell(\neg\exists x_i \neg\varphi)$,

It is known [15] that if ϕ has no free variable, then $\mathcal{K}_\ell(\phi)$ is exactly the language accepted by ϕ.

For technical issues, the languages $H_\ell((f,u))$, with $(f,u) \in \mathcal{F} \times \{0,1\}^\ell$, of terms t over $\mathcal{F} \times \{0,1\}^\ell$ such that there exists at least a position p of t such that $t(p) = (f,u)$ are introduced.

Lemma 5 *The language K_ℓ and the languages $H_\ell((f,u))$ are accepted by deterministic partially ordered automata.*

PROOF. If u and v are two elements of $\{0,1\}^\ell$, we denote by $u \sqcap v$ (resp. $u \sqcup v$) the element w of $\{0,1\}^\ell$ defined by $w(i) = 1$ iff $u(i) = 1$ and $v(i) = 1$ (resp. either $u(i) = 1$ or $v(i) = 1$). We consider the automaton $\mathcal{A}_\ell$ whose set of states is $\{q_u \mid u \in \{0,1\}^\ell\}$, whose set of final states is reduced to $\{q_{1^\ell}\}$ and whose set of transitions is defined by the following:

- For every $A \in \mathcal{F}$ of arity 0 and for every $u \in \{0,1\}^\ell$, $(A,u) \to q_u$ is a transition of $\mathcal{A}_\ell$,
- For every $f \in \mathcal{F}$ of arity $r \geq 1$, and for every $u_0, u_1, \ldots, u_r \in \{0,1\}^\ell$, if $u_i \sqcap u_j = 0^\ell$ for all $i \neq j$, then $(f,u_0)(q_{u_1}, \ldots, q_{u_r}) \to q_{u_0 \sqcup u_1 \sqcup \ldots \sqcup u_r}$ is a transition of $\mathcal{A}_\ell$,
- there is no other transition.

One can easily check that for every term t accepted by $\mathcal{A}_\ell$ and each $1 \leq i \leq \ell$, there exists a unique position p, such that $t(p) \in \mathcal{F} \times \{0,1\}^{i-1}1\{0,1\}^{n-i}$. It follows that $\mathcal{A}_\ell$ accepts $\mathcal{K}_\ell$. Moreover, $\mathcal{A}_\ell$ is partially ordered by the relation $q_u \leq q_v$ iff $v(i) = 1 \Rightarrow u(i) = 1$.

The construction of $H_\ell((f,u))$ is easy and left to the reader. □

Lemma 6 *If $\varphi(x_1, \ldots, x_\ell)$ is a constant free atomic formula and $0 \leq k \leq \ell$ then the languages $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \varphi)$ and $\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k \varphi)$ are accepted by a deterministic partially ordered automata.*

PROOF. The lemma is proved by an induction on φ. Assuming firstly that $\varphi(x_1, \ldots, x_\ell)$ is a constant free atomic formula such that $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \varphi)$ and $\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k \varphi)$ are respectively accepted by the deterministic partially ordered automata $\mathcal{A}_\exists$ and $\mathcal{A}_\forall$, one has

$$\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k \neg\varphi) = \mathcal{K}_\ell(\neg\exists x_1 \ldots \exists x_k \neg\neg\varphi) = \mathcal{K}_\ell(\neg\exists x_1 \ldots \exists x_k \varphi) = L(\mathcal{A}_\exists)^c \cap \mathcal{K}_{\ell-k}.$$

Since the class of languages accepted by deterministic partially ordered automata is closed under intersection and complement, $\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k \neg\varphi)$ is accepted by a deterministic partially ordered automaton. The same argument holds for $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \neg\varphi)$.

Secondly, if $\varphi_1(x_1, \ldots, x_\ell)$ and $\varphi_2(x_1, \ldots, x_\ell)$ are constant free atomic formulae, then $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k (\varphi_1 \vee \varphi_2)) = \mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \varphi_1) \cup \mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \varphi_2)$. Since the class of languages accepted by deterministic partially ordered automata is closed under union, if $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k(\varphi_1)$ and $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k(\varphi_2)$ are accepted by deterministic partially ordered automata, then $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k(\varphi_1 \vee \varphi_2))$ is accepted by a deterministic partially ordered automaton too. Since the class of languages accepted by deterministic partially ordered automata is closed under intersection, a similar construction may be done in order to prove that $\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k(\varphi_1 \vee \varphi_2))$ is accepted by a deterministic partially ordered automaton.

It remains to prove that $\mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \varphi)$ and $\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k \varphi)$ are accepted by deterministic partially ordered automata for $\varphi \in \{R_f(x_i), x_i = x_j, x_i <_r x_j \mid 1 \leq i \leq \ell, 1 \leq j \leq \ell\}$. We only give the proof for $R_f(x_i)$.

1) Let $H = \mathcal{K}_\ell(\exists x_1 \ldots \exists x_k R_f(x_i))$. If $i \leq k$, then

$$H = \{t \in K_{\ell-k} \mid \exists p,\ t(p) = (f,u)\} = \mathcal{K}_{\ell-k} \cap \bigcup_{u \in \{0,1\}^{\ell-k}} H_{\ell-k}((f,u)).$$

Then, using Lemma 3 and 5, H is accepted by a deterministic partially ordered automaton. Now if $i > k$, then

$$\begin{aligned} H &= \{t \in K_{\ell-k} \mid \exists p,\ t(p) = (f,u) \text{ and } u(i-k) = 1\} \\ &= \mathcal{K}_{\ell-k} \cap \bigcup_{u \in \{0,1\}^{\ell-k},\ u(i-k)=1} H_{\ell-k}((f,u)). \end{aligned}$$

Similarly, H is accepted by a deterministic partially ordered automaton.

2) Let $S = \mathcal{K}_\ell(\exists x_1 \ldots \exists x_k \neg R_f(x_i))$. If $i \leq k$, then

$$S = \{t \in K_{\ell-k} \mid \exists p,\ t(p) \neq (f,u)\} = \mathcal{K}_{\ell-k} \cap \bigcup_{u \in \{0,1\}^{\ell-k},\ g \neq f} H_{\ell-k}((g,u)).$$

Moreover, if $i > k$, then

$$\begin{aligned} H &= \{t \in K_{\ell-k} \mid \exists p,\ t(p) \neq (f,u) \text{ and } u(i-k) = 1\} \\ &= \mathcal{K}_{\ell-k} \cap \bigcup_{u \in \{0,1\}^{\ell-k},\ u(i-k)=1, g \neq f} H_{\ell-k}((g,u)). \end{aligned}$$

As for 1), It follows that S is accepted by a deterministic partially ordered automaton. Consequently $\mathcal{K}_\ell(\forall x_1 \ldots \forall x_k R_f(x_i))$ also is accepted by a deterministic partially ordered automaton.

□

From Lemma 6 one can deduced the main result of this section.

Proposition 7 *If a language is accepted by a Σ_2 formula, then it is accepted by a partially ordered automaton.*

PROOF. First of all, notice that if $\mathcal{A} = (\mathcal{F} \times \{0,1\}^j, Q, \Delta, Q_f)$ is partially ordered automaton accepting a language of $\mathcal{K}_j$, then for all $1 \leq i \leq j$, $\pi_i(L(\mathcal{A}))$ is accepted by the partially ordered automaton $(\mathcal{F} \times \{0,1\}^{j-1}, Q, \Delta', Q_f)$ where

$$\Delta' = \{\pi_i((f,u))(q_1,\ldots,q_k) \to q \mid (f,u)(q_1,\ldots,q_k) \to q \in \Delta\}$$

Let $\exists x_1 \ldots \exists x_n \forall y_1 \ldots \forall y_k \ \varphi(x_1,\ldots,x_n,y_1,\ldots,y_k)$ accepting a language L. By Lemma 6 one knows that $\mathcal{K}_\ell(\forall y_1 \ldots \forall y_k \ \varphi(x_1,\ldots,x_n,y_1,\ldots,y_k))$ is accepted by a deterministic partially ordered automaton. Using the above remark, it follows that $L = \pi_1 \ldots \pi_n(\mathcal{K}_\ell(\forall y_1 \ldots \forall y_k \ \varphi(x_1,\ldots,x_n,y_1,\ldots,y_k)))$ is accepted by a partially ordered automaton. □

3 Counterexample

The set of functional symbols of arity k. is denoted by $\mathcal{F}_k$. The semi-commutation relation R_S generated by a finite subset S of $\cup_{k\geq 1}\mathcal{F}_k \times \mathcal{F}_k$ is defined by: $(s,t) \in R_s$ if and only if s and t have the same set of positions and if there exist positions p and $p.i$ of s, such that $(s(p), s(p.i)) \in S$, $s(p) = t(p.i)$, $s(p.i) = t(p)$ and if for every other position p' of s, $s(p') = t(p')$.

Let $\mathcal{A}$ be the bottom-up tree automaton whose set of states is $\{q_A, q_B, q_F\}$, whose set of transitions is $\{A \to q_A, B \to q_B, f(q_A,q_A) \to q_A, g(q_B,q_B) \to q_B, f(q_A,q_B) \to q_f\}$ and whose unique final state is q_f. Let $S = \{(f,g),(g,f)\}$. Finally let K be the set of terms

$$K = \{t \mid \exists s \text{ accepted by } \mathcal{A} \text{ s.t. } (s,t) \in R_S^*\}.$$

Proposition 8 *The automaton $\mathcal{A}$ is partially ordered and $L(\mathcal{A})$ is accepted by a Σ_2 formula. Moreover the tree language K is not regular.*

PROOF. (sketch) One can easily check that $\mathcal{A}$ is partially ordered by the relation $\leq$ defined by: $q_A \leq q_A$, $q_B \leq q_B$, $q_A \leq q_f$ and $q_B \leq q_f$.

One can also easily check that $L(\mathcal{A})$ is accepted by the Σ_2 formula:

$$\exists x \forall y \ [R_f(x) \wedge ((x <_1 y) \vee (x <_2 y)) \wedge ((x <_1 y \Rightarrow (R_f(y) \vee R_A(y)))) \wedge ((x <_2 y \Rightarrow (R_g(y) \vee R_B(y))))],$$

where the symbols $\wedge$ and $\Rightarrow$ are classically defined by $(\varphi_1 \wedge \varphi_2) \stackrel{\text{def}}{=} \neg(\neg\varphi_1 \vee \varphi_2))$ and $(\varphi_1 \Rightarrow \varphi_2) \stackrel{\text{def}}{=} (\neg\varphi_1 \vee \varphi_2)$.

We claim that K is the set of terms t such that $t(\varepsilon) \in \{f, g\}$, all leaves of $t_{|1}$ are labelled by A, all leaves of $t_{|2}$ are labelled by B and

- if $t(\varepsilon) = f$, then the number of g occurring in $t_{|1}$ is equal to the number of f occurring in $t_{|2}$ and
- if $t(\varepsilon) = g$, then the number of g occurring in $t_{|1}$ is equal to the number of f occurring in $t_{|2}$ minus 1.

The proof of this claim is straightforward by an induction on the number of g occurring in $t_{|1}$.

One can prove that K is not regular using the classical and scholarly *pumping* techniques. Intuitively, we are not able to count in the context of regular languages. The details are left to the reader. □

4 Conclusion

The result exposed in this paper is not surprising: several works [4,7,2] show that tree transformations are difficult to analyse. Notice that the language K defined above is neither regular nor accepted by standard extended classes of tree automata, like automata with constraints between brothers. One can now use the approach developed in [11] and wonder whether there exists a partially ordered tree automaton accepting a tree language whose image by the reflexive-transitive closure of a semi-commutation relation is regular but not accepted by a partially ordered automaton. A question remains unanswered: does a fine logical characterisation of tree languages accepted by partially ordered automata exists? Moreover, as far as we know, the question on whether a regular tree language given by its minimal automaton can be accepted by a partially ordered tree automata is still open and, in that event, how to compute such an automaton. Notice that this question requires strong arguments to be solved in the word automaton cases [1,13]

The author wishes to express his thanks to the anonymous referee for drawing author attention to logical aspects that greatly improve the paper. The author would like to acknowledge Eugène Pamba Capo-Chichi for valuable comments.

References

[1] M. Arfi. Polynomial operations on rational languages. In *STACS: Annual Symposium on Theoretical Aspects of Computer Science*, 1987.

[2] Bouajjani and Touili. On computing reachability sets of process rewrite systems. In *RTA: Rewriting Techniques and Applications, 1987, LNCS 256*, 2005.

[3] A. Bouajjani, A. Muscholl, and T. Touili. Permutation rewriting and algorithmic verification. *Information and Computation*, 205(2):199–224, 2007.

[4] A. Bouajjani and T. Touili. Extrapolating tree transformations. *Lecture Notes in Computer Science*, 2404:539–??, 2002.

[5] G. Cécé, P.-C. Héam, and Y. Mainier. Efficiency of automata in semi-commutation verification techniques, 2003. To appear in ITA-RAIRO.

[6] H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. *Tree Automata Techniques and Applications.* 1999.

[7] G. Feuillade, T. Genet, and V. V. T. Tong. Reachability analysis over term rewriting systems. *J. Autom. Reasoning*, 33(3-4):341–383, 2004.

[8] P.-C. Héam. On shuffle ideals. *ITA*, 36(4):359–384, 2002.

[9] G. Higman. Ordering by divisibility in abstract algebras. *Proc. of the London Mathematical Society*, 2:326–336, 1952.

[10] M. Lothaire. *Combinatorics on words.* Cambridge Mathematical Library, 1983.

[11] Muscholl and Petersen. A note on the commutative closure of star-free languages. *IPL: Information Processing Letters*, 57, 1996.

[12] J.-E. Pin. Logic, semigroups and automata on words. *Ann. Math. Artif. Intell*, 16:343–384, 1996.

[13] J.-É. Pin and P. Weil. Polynominal closure and unambiguous product. *Theory Comput. Syst*, 30(4):383–422, 1997.

[14] H. Straubing and D. Thérien. Partially ordered finite monoids and a theorem of I. Simon. *Journal of Algebra*, 119:161–183, 1985.

[15] J. Tatcher and J. Wright. Generalized finite automata with an application to a decision problem of second oreder logic. *Mathematical System Theory*, 2:57–88, 1968.

[16] W. Thomas. Classifying regular events in symbolic logic. *Journal of Computer and System Sciences*, 25(3):360–376, Dec. 1982.

[17] W. Thomas. Languages, automata and logic. In A. Salomaa and G. Rozenberg, editors, *Handbook of Formal Languages*, volume 3, Beyond Words. Springer, Berlin, 1997.