

# Regularity for a Large Class of Context-Free Processes is Decidable

D.J.B. Bosscher<sup>1</sup> \*    W.O.D. Griffioen<sup>1,2</sup> \*\*

<sup>1</sup> CWI

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

{doeko,griffioe}@cwi.nl

<sup>2</sup> Computing Science Institute, University of Nijmegen

P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

**Abstract.** Regularity of context-free processes has been proved to be decidable for BPA systems by [MM94] and normed context-free processes by [Kru95]. In this paper the decidable class of regular context-free processes is enlarged to that of context-free processes over so-called NRD specifications (definition in the paper). Furthermore an upper bound is given for the number of states modulo bisimulation.

## 1 Introduction

A classical result from formal language theory is that language equivalence is undecidable for context-free grammars. In [BH64] the correspondence Lemma of Post [Pos46] is used to prove this and the undecidability of regularity of context-free languages. The picture changes if these grammars are studied as process specifications modulo a given equivalence for process graphs.

A well-known equivalence for process theory, bisimulation equivalence is decidable for context-free processes [BBK93, CHS92]. This leads one to believe bisimulation is a sufficiently strong equivalence to allow decidability of regularity of context-free process graphs. In this paper decidability of regularity for a large class of context-free process graphs is proved.

A process graph is regular iff it is (strongly) bisimilar to a process graph with finitely many states. Context-free process graphs are denoted by “guarded recursive specifications” over Basic Process Algebra (BPA) [BW90]. Below three examples of BPA specifications for a process name  $X$  are given.

$$\begin{array}{lll} X = a \cdot Y + c \cdot Z & & X = a \cdot Y \cdot Z \\ (1) Y = d \cdot X + e \cdot Y & (2) X = a \cdot X \cdot Y + c & (3) Y = a \cdot Y \cdot B + c \\ Z = c & Y = b \cdot Y & Z = b \cdot Z \\ & & B = b \end{array}$$

---

\* Research supported by Esprit BRA 7166 CONCUR 2.

\*\* Research supported by the Netherlands Organization for Scientific Research (NWO) under contract SION 612-316-125.

All three specifications define a regular process graph for  $X$ . The process graph for  $X$  is regular with respect to first specification because the specification is linear. It is folklore that the class of regular processes is the same as the class of processes which can be denoted by a linear specification. In fact one of the reasons for using regular processes is that these can be described precisely by such specifications, which allow easy implementation and checking of modal and temporal properties [Hol89].

The second specification defines the process graph for  $X$  also as a regular process, but to see this is already more difficult, since the specification is not linear ( $X$  has summand  $a \cdot X \cdot Y$  in the defining equation). In the paper by [MM94] it is proved that specifications which do not define "normed stackings" define only process names with regular process graphs. The same result is proved for normed process graphs by [Kru95]. Both papers give a method to generate the linear specification for a process graph, provided it is regular.

In this paper we extend the class of specifications further by allowing specifications such as 3. The specification is not linear and allows a normed stacking for  $Y$ . However  $X$ , which depends on  $Y$  defines a regular process graph: the idea is that the "context-free behavior" of process name  $Y$  is somehow neutralized by  $Z$ . Furthermore we give an explicit upper bound for the number of states of the process graph modulo bisimulation and a method to generate the linear specification.

As was pointed out by Didier Caucal our method does not work for every context-free process graph. We must restrict ourselves to specifications which have only so-called weakly deterministic process names in normed repeats (see Section 3 for the definition), so-called process graphs over NRD specifications (process **N**ames occurring in a weakly **N**ormed **R**epeat are weakly **D**eterministic).

The plan of the paper is as follows. In the second section we introduce (context-free) processes and process graphs formally and give some preliminary definitions. In the third section we prove the decidability result. We conclude with a remark on generating a linear specification for a regular process graph.

## 2 Context-free Processes and Regularity

As usual we refer to context-free processes as processes over specifications in Basic Process Algebra. Therefore, in this section we define the basic notions of BPA. We start by giving the syntax and semantics of BPA. Next we define process graphs and bisimulation equivalence as used throughout this paper. For a detailed description of the relation between language and process theory we refer the interested reader to [HM96].

The abstract syntax of BPA is given by

$$p ::= a \mid X \mid p + p \mid p \cdot p$$

where  $a$  ranges over a *finite* set *Act* of atomic actions and  $X$  over *countable*

*infinite* set *Names* of process names. The  $+$  is the usual process algebraic notation for choice and  $\cdot$  for sequential composition. We look at *recursive* processes, processes where the meaning of process names is given by a (finite) set of equations of the form

$$\Delta = \{X_i = p_i \mid 1 \leq i \leq k\}$$

where  $i, k \in \mathbb{N}$ , the  $X_i$  are distinct process names and the process names used in  $p_i$  are those defined in  $\{X_1, \dots, X_k\}$ . In the sequel we will only look at guarded recursive equations, i.e. every summand of the processes  $p_i$  in the equations  $X_i = p_i$  starts with an action.

The operational semantics of a BPA expression, given a specification  $\Delta$ , is the smallest transition relation  $\rightarrow_\Delta \subset \text{BPA} \times \text{Act} \times (\text{BPA} \cup \{\epsilon\})$  containing the transitions provable by the following rules:

$$\frac{p \xrightarrow{a} p'}{p + q \xrightarrow{a} p'} \quad \frac{q \xrightarrow{a} q'}{p + q \xrightarrow{a} q'}$$

$$\frac{p \xrightarrow{a} p'}{p \cdot q \xrightarrow{a} p' \cdot q} \quad \frac{}{a \xrightarrow{a} \epsilon} \quad \frac{p \xrightarrow{a} p'}{X \xrightarrow{a} p'}$$

where  $a \in \text{Act}$ ,  $X = p \in \Delta$  and  $\epsilon$  is a special state not in BPA which denotes the empty or *terminated* state. For convenience we identify the states  $\epsilon \cdot p$  and  $p \cdot \epsilon$  with  $p$ .

We will use  $V_\Delta$  for the set of process names defined in  $\Delta$ . We use  $\nu, \rho, \sigma$  to range over  $V_\Delta^*$  and  $|\nu|$  as the length of a sequence  $\nu$  of process names. We use  $\nu^k$  as an abbreviation for a sequence  $\nu \cdot \dots \cdot \nu$  of length  $k$ . Notice that we thus identify string concatenation with sequential composition and use that concatenation is associative. We speak of the head and the tail of a sequence in the obvious way. As a convention we use that  $\nu^0$  is equal to  $\epsilon$ .

Most of the time we omit actions in transitions. We use the Greek letter  $\phi$  to range over transition sequences starting with a process name in  $V_\Delta$ .  $\phi(0)$  denotes the first state in transition sequence,  $\phi(i-1)$  the  $i$ -th state in  $\phi$ . The Greek letters  $\chi$  and  $\psi$  are used to range over transition *subsequences*, i.e. choppings off of transition sequences. We mean by  $|\phi|$  the length of a transition sequence given by the number of transitions. We say that a process-expression  $p$  is weakly normed, also denoted by  $p \downarrow$ , if there is a transition sequence  $p \rightarrow_\Delta \dots \rightarrow_\Delta \epsilon$ <sup>3</sup>. A process-expression  $p$  is perpetual, also denoted by  $p \uparrow$  if it is not weakly normed. By  $|V_\Delta|$  we mean the size of the set  $V_\Delta$ . If no confusion can arise we use  $|\nu|, |\phi|, |V_\Delta|$  without stating which kind of length or size is used.

<sup>3</sup> Notice that a weakly normed process can arrive after one or more transitions in a state which is *not* weakly normed, i.e. it is not *normed* in the usual meaning.

In rest of this paper we mostly look at process graphs of process-expressions defined in  $\Delta$  in the context of bisimulation equivalence. As usual we define process graphs as rooted, labeled transition systems.

**Definition 2.1** Let  $Act$  be a given set of actions.

1. A *labeled transition system* over  $Act$  is a pair  $(S, \rightarrow)$  where  $S$  is a set of states and  $\rightarrow \subseteq S \times Act \times S$  is the transition relation,
2. A *process graph* over  $Act$  is a triple  $g = (r, S, \rightarrow)$  with  $(S, \rightarrow)$  a labeled transition system over  $Act$  and  $r \in S$  is the root state.

In this paper the equivalence between process graphs is strong bisimulation equivalence.

**Definition 2.2** Let  $g = (r, S, \rightarrow)$  and  $g' = (r', S', \rightarrow)$  be process graphs.

1.  $R \subseteq S \times S'$  is a *bisimulation* iff for all  $(p, q) \in R$  it holds that
  - (a) If  $p \xrightarrow{a} p'$ , then there is a  $q'$  such that  $q \xrightarrow{a} q'$  and  $(p', q') \in R$ ,
  - (b) If  $q \xrightarrow{a} q'$ , then there is a  $p'$  such that  $p \xrightarrow{a} p'$  and  $(p', q') \in R$ .
2.  $g$  and  $g'$  are *bisimilar* iff there is a bisimulation relating the roots of  $g$  and  $g'$ .

Now we are ready to define what we mean by regularity.

**Definition 2.3** A process  $p$  is *regular* iff the process graph  $(p, \text{BPA} \cup \{\epsilon\}, \rightarrow_\Delta)$  is bisimilar to a process graph with a finite number of states.

### 3 Decidability of Regularity

In this section we prove that regularity for a large class of context-free processes is decidable.

For the proof we can restrict ourselves to BPA systems in restricted Greibach Normal Form (rGNF), i.e. systems of the form

$\Delta = \{X_i = a_{i1} \cdot \sigma_{i1} + \dots + a_{in_i} \cdot \sigma_{in_i} \mid 1 \leq i \leq m\}$ , where  $\sigma_{ij}$  is a process name sequence containing *at most two* process names. It is folklore that any guarded BPA specification can effectively be represented in restricted Greibach Normal Form (rGNF)<sup>4</sup> maintaining bisimulation equivalence of the process graphs of the defined processes. We have the following useful properties.

**Proposition 3.1** Let  $\Delta$  be a BPA system in rGNF and  $X \in V_\Delta$ .

1. Let  $\phi \equiv X \rightarrow_\Delta \dots$  be a transition sequence. For every  $i \in \mathbb{N}$ ,  $\phi(i)$  is a sequence of process names in  $V_\Delta$ ,

<sup>4</sup> For context-free processes see e.g. [BBK87, BP95]. Similar result for context-free languages and language equality [HU79].

2. If  $\sigma \rightarrow_{\Delta} \rho$ , then  $|\rho| \leq |\sigma| + 1$ .

**Remark 3.2** Notice that Proposition 3.1 gives that the states reachable from a process name are *sequences* of process names.

**Definition 3.3** Let  $X = p \in \Delta$ . An action is *unique in X*, iff  $p$  has only one summand starting with  $a$ .

**Definition 3.4** A transition (sub)sequence  $\chi$  is *weakly deterministic* iff for every transition  $\chi(i) \xrightarrow{a} \chi(i+1)$ ,  $a$  is unique in the head of  $\chi(i)$ .

**Definition 3.5** Let  $W_0, W_1, \dots \subseteq V_{\Delta}$  be sets inductively defined as follows

1.  $W_0 = \emptyset$ ,
2.  $X \in W_{i+1}$  iff either  $X \in W_i$  or  $X = p \in \Delta$  and there is an action  $a$  such that  $a$  is unique in  $p$ , and the summand starting with  $a$  has all process names in  $W_i$ .

Let  $W_i = W_{i+1}$  for some  $i \in \mathbb{N}$ . Then  $W_i$  is the smallest set of *weakly deterministic* process names for  $\Delta$ , denoted as  $W_{\Delta}$ .

In Lemma 3.17 we prove that elements of  $W_{\Delta}$  all start a weakly deterministic and terminating transition (sub)sequence.

**Remark 3.6** Notice that  $W_{\Delta}$  can be effectively computed (as  $W_{|V_{\Delta}|}$ ).

We define *weakly normed repeat invariance* as the criterion for regularity. Informally it states that looping through a specification while stacking does not result in infinitely many non bisimilar states.

**Definition 3.7** Let  $\Delta$  be a specification in rGNF. Let  $\phi$  be a transition sequence  $\phi(0) \xrightarrow{\alpha_0}_{\Delta} \phi(1) \xrightarrow{\alpha_1}_{\Delta} \dots$ .

1. The transition sequence  $\phi(0) \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \phi(i)$  is a *repeat* iff  $\phi(i) \equiv \phi(0) \cdot \sigma$ .
  - (a) which is *cyclic* iff  $\phi(i) \equiv \phi(0)$ ,
  - (b) which is *perpetual* iff  $\phi(i) \uparrow$  and  $\phi(i) \not\equiv \phi(0)$ ,
  - (c) which is *weakly normed* iff  $\phi(i) \downarrow$  and  $\phi(i) \not\equiv \phi(0)$ ,
2. The transition subsequence  $\chi(0) \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \chi(i)$  *corresponds* to the repeat  $\phi(0) \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \phi(i)$  iff  $\chi(0) \equiv \phi(0) \cdot \sigma, \dots, \chi(i) \equiv \phi(i) \cdot \sigma$ ,
3. A transition sequence *uses* a repeat when it has a transition subsequence that corresponds to a repeat,
4. A repeat  $\phi(0) \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \phi(i)$  is *lonely* iff  $\phi(0) \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \phi(i)$  uses no other repeats,
5. A transition sequence  $\phi$  is a possible entry to a weakly normed repeat, *entry* for short, iff
  - (a)  $\phi$  starts in the root  $r$ ,
  - (b) uses no weakly normed repeats and,
  - (c) passes only through pair-wise non bisimilar states,

- i. An entry  $\phi$  is *weakly normed repeat invariant* for a (subsequence corresponding to a) lonely weakly normed repeat  $\chi$  iff
  - (a) the last state of  $\phi$  is equal to the first state of  $\chi$ ,
  - (b) the first and last state of  $\chi$  are bisimilar.
- 7.  $r$  is *weakly normed repeat invariant* iff every entry to a (subsequence corresponding to a) lonely weakly normed repeat is weakly normed repeat invariant.

**Definition 3.8** Let  $\Delta$  be a specification in rGNF. A process name  $Y$  occurs in *weakly normed repeat*  $\phi(0) \rightarrow \dots \rightarrow \phi(i)$ , iff  $\phi(i) \equiv \phi(0) \cdot \sigma$  and  $Y$  occurs in  $\sigma$ .

In the sequel we assume a fixed but arbitrary system of equations  $\Delta$  in restricted Greibach Normal Form with the root process name  $r$  in  $V_\Delta$  in a process graph  $(r, \text{BPA} \cup \{\epsilon\}, \rightarrow_\Delta)$ , all as defined in the previous section.

Finally we are able to define the class of process for which the decidability proof holds.

**Definition 3.9** We say that the specification  $\Delta$  is NRD (weakly Normed Repeat variables are weakly Deterministic) iff the set of process names of  $V_\Delta$  occurring in a weakly normed repeat is a subset of  $W_\Delta$ .

**Remark 3.10** Note that it is decidable if a specification is a NRD specification.

*Example 1.* Let  $\Delta$  be defined as

$$\Delta = \left\{ \begin{array}{l} X = a \cdot Y, \\ Y = a \cdot Y \cdot B + d, \\ B = b \end{array} \right\}$$

where  $X$  is the root process name. The transition sequence  $\chi \equiv Y \xrightarrow{a} Y \cdot B$  is a weakly normed repeat, because  $Y \cdot B \downarrow$  and  $Y \not\equiv Y \cdot B$ . Furthermore  $\chi$  is a lonely repeat. Contrast this with  $\chi' \equiv Y \xrightarrow{a} Y \cdot B \xrightarrow{a} Y \cdot B \cdot B$  which is a weakly normed but not a lonely repeat, because it uses  $\chi$ .

The transition sequence  $\phi \equiv X \xrightarrow{a} Y$  is an entry, because it uses no weakly normed repeat and  $X \not\equiv Y$ .  $\phi$  is not weakly normed repeat invariant for  $\chi$  because the last state  $Y$  of  $\phi$ , is not bisimilar to the last state  $Y \cdot B$  of  $\chi$ .

The root process name  $X$  is not weakly normed repeat invariant, because the entry  $\phi$  is not weakly normed repeat invariant for the lonely, weakly normed repeat  $\chi$ . The reader can verify easily that the specification is NRD. It allows only process names from  $W_\Delta$  in a weakly normed repeat: the process name  $B$  occurring in the weakly normed repeat  $Y \xrightarrow{a} Y \cdot B$ , is in  $W_1$ .

The first part of the proof establishes that a weakly normed repeat invariant root cannot start an *infinite* transition sequence of pair-wise non bisimilar states. This implies regularity with the following well-known fact.

**Proposition 3.11** If  $r$  is not regular, then there exists an *infinite* transition sequence  $r \rightarrow_\Delta \dots$  where all states are pair-wise non bisimilar.

**Lemma 3.12** The length of a transition subsequence  $\chi \equiv \sigma \rightarrow_{\Delta} \dots$  which uses  $no$  repeats is maximally  $|\sigma| \cdot (2^{|\mathcal{V}_{\Delta}|} - 1)$ .

**Proof.** By a well-founded simultaneous induction on the number of different process names  $h$  that occur left-most in states of  $\chi$  and the length  $l$  of  $|\sigma|$ . Now assume the Induction Hypothesis holds for all tuples which are lexicographically smaller than tuples  $(h, l)$ . Distinguish the following cases in the Induction Step.

1. If  $l = 1$ , then  $\chi \equiv X \rightarrow_{\Delta} \rho \rightarrow_{\Delta} \dots$  for some  $X \in \mathcal{V}_{\Delta}$ . We only prove it for the case that  $\rho \not\equiv \epsilon$ , for  $\rho \equiv \epsilon$  the result is immediate. The Induction Hypothesis holds for  $\psi \equiv \rho \rightarrow_{\Delta} \dots$  because  $X$  cannot occur left-most in states of  $\psi$  any more, otherwise  $\psi$  uses a repeat starting with  $X$ . By Proposition 3.1  $|\rho| \leq 2$  and so  $|\chi| = 1 + |\psi| \leq 1 + 2 \cdot (2^{|\mathcal{V}_{\Delta} - \{X\}|} - 1) = |\sigma| \cdot (2^{|\mathcal{V}_{\Delta}|} - 1)$ ,
2. If  $l > 1$ , then  $\chi \equiv X \cdot \rho \rightarrow_{\Delta} \dots$ , for some  $X \in \mathcal{V}_{\Delta}, \rho \in \mathcal{V}_{\Delta}^+$ . By the Induction Hypothesis any transition sequence without using repeats  $X \rightarrow_{\Delta} \dots \epsilon$  has a length not exceeding  $2^{|\mathcal{V}_{\Delta}|} - 1$  and so does the associated transition (sub)sequence  $X \cdot \rho \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \rho$ , which does not use the presence of the tail  $\rho$ . Also by the Induction Hypothesis the transition (sub)sequence  $\rho \rightarrow_{\Delta} \dots$  has a maximal length of  $|\rho| \cdot (2^{|\mathcal{V}_{\Delta}|} - 1)$ , which gives a total maximal length of  $|\sigma| \cdot (2^{|\mathcal{V}_{\Delta}|} - 1)$ .

We use that a perpetual repeat passes through a state with a perpetual process name at the first or second position.

**Lemma 3.13** If  $\phi$  is a perpetual repeat, then there is a state  $\rho$  in  $\phi$  so that  $\rho \equiv P \cdot \sigma$  or  $\rho \equiv Y \cdot P \cdot \sigma$ , where  $Y, P \in \mathcal{V}_{\Delta}$ ,  $Y \downarrow$  and  $P \uparrow$ .

**Proof.** Let  $N_{\Delta}$  and  $P_{\Delta}$  be the subsets of weakly normed and perpetual process names of  $\mathcal{V}_{\Delta}$  respectively. Suppose  $\phi(i)$  is the last state of the perpetual repeat  $\phi$ . By definition  $\phi(i)$  is of the form  $X \cdot \rho$  with  $X \cdot \rho$  perpetual. It is easily verified that either  $X$  is perpetual, in which case we are finished, or there is a  $P \in P_{\Delta}$  so that  $\rho \equiv \nu \cdot P \cdot \sigma$  with  $\nu \in N_{\Delta}^+$ . Let  $\phi(j)$  be the first state so that  $P$  is present in the process name sequence. Suppose the preceding state  $\phi(j-1)$  has the process name  $Y \in N_{\Delta}$  at the first position in the process name sequence. Using the definition of  $Y$  in  $\Delta$ ,  $P$  is introduced. Because  $\Delta$  is in rGNF, this implies that the perpetual process name  $P$  is introduced at the first or second position.

**Lemma 3.14** The length of a transition sequence, which uses no weakly normed repeat and passes only through pair-wise non bisimilar states is maximally  $|\mathcal{V}_{\Delta}|^2 \cdot 2^{|\mathcal{V}_{\Delta}|}$ .

**Proof.** Let  $N_{\Delta}$  and  $P_{\Delta}$  be as in the previous proof. Suppose  $\phi \equiv X \rightarrow_{\Delta} \dots$  is a transition sequence which uses no weakly normed repeat. With Lemma 3.12 we know that after  $2^{|\mathcal{V}_{\Delta}|}$  transitions  $\phi$  has used at least one cyclic or perpetual repeat. This repeat cannot be cyclic, because  $\phi$  has then at least two bisimilar states. So  $\phi$  has used a perpetual repeat in the first  $2^{|\mathcal{V}_{\Delta}|}$  transitions. By Lemma

3.13 it has either passed through a state  $\rho$  of the form (1)  $\rho \equiv P \cdot \sigma$  or (2)  $\rho \equiv X \cdot P \cdot \sigma$ , where  $X \in N_\Delta$  and  $P \in P_\Delta$ .

In case (1) in the state  $\rho'$  following  $\rho$  in  $\phi$ , there is a perpetual process name at the first or second position. Because  $P$  is perpetual the tail  $\sigma$  cannot shift left-most in the transition from  $\rho$  to  $\rho'$ .  $\Delta$  is in rGNF and therefore  $\rho'$  is of the form  $\rho' \equiv \nu \cdot \sigma$  and  $1 \leq |\nu| \leq 2$ . Every transition from  $P$  is to another perpetual state, so  $\nu$  has one perpetual process name. But then  $\rho'$  has a perpetual process name at the first or second position.

In case (2)  $\phi$  passes through a state with a perpetual process name at the first or second position in  $2^{|V_\Delta|}$  transitions. If  $\phi$  does not use the presence of the tail of  $\rho$ , then this implies with the previous argument that in  $2^{|V_\Delta|}$  steps from  $\rho$ ,  $\phi$  passes through a state with a perpetual process name at the first or second position. If  $\phi$  does use the presence of the tail of  $\rho$ , then by Lemma 3.12 in  $2^{|V_\Delta|}$  transitions  $X$  and its reducts have disappeared and is  $P$  the left-most process name. So in the gaps between perpetual states with a perpetual process name at the first or second position in  $\phi$  there are at most  $2^{|V_\Delta|}$  connecting transitions.

Notice that there are maximally  $|P_\Delta| + |N_\Delta| \cdot |P_\Delta|$  of such non bisimilar perpetual states in  $\phi$ , because  $P \cdot \nu \not\equiv P \cdot \sigma$  and  $X \cdot P \cdot \nu \not\equiv X \cdot P \cdot \sigma$  for every  $P \in P_\Delta$ ,  $\nu$  and  $\sigma$ . If  $|N_\Delta| = 0$  we have a maximal length of  $|V_\Delta| - 1$ , see Remark below. If  $|N_\Delta| > 0$  simple arithmetic gives us that  $|P_\Delta| + |N_\Delta| \cdot |P_\Delta| \leq |V_\Delta|^2 - 1$ . Hence the maximal length is the number of different perpetual states with a perpetual process name at the first or second position plus one times the maximal length of a gap, i.e.  $|V_\Delta|^2 \cdot 2^{|V_\Delta|}$ .

**Remark 3.15** Actually we conjecture the maximal length to have an upper bound of  $2^{|V_\Delta|} - 1$ . The reader can verify this in the two simple cases that either *no* process names are perpetual (Lemma 3.12) or all process names in  $V_\Delta$  are perpetual. In the last case the maximum number of states modulo bisimulation is even  $|V_\Delta|$  and hence the maximal length  $|V_\Delta| - 1$ .

We show the soundness of the criterion by proving that for *all* context-free processes, not only over NRD specifications, it disallows infinite transition sequences passing through pair-wise non bisimilar states and hence implies regularity.

**Theorem 3.16** If  $\tau$  is weakly normed repeat invariant, then  $\tau$  is regular.

**Proof.** Assume that  $\tau$  is weakly normed repeat invariant, and suppose  $\tau$  is not regular. The non regularity implies with Proposition 3.11 that there is a transition sequence  $\phi$  starting in  $\tau$  passing through infinitely many pair-wise non bisimilar states.

First suppose  $\phi$  uses no weakly normed repeat. By Lemma 3.14 we conclude that  $\phi$  has a finite length. Contradiction. Therefore  $\phi$  uses a weakly normed repeat.

Let  $\chi$  be the first (subsequence corresponding to a) weakly normed repeat used in  $\phi$ .  $\chi$  cannot use a cyclic repeat, because then it would have two bisimilar



states. Also it cannot use a perpetual repeat, because then the last state of  $\chi$  would be perpetual. Therefore  $\chi$  is a (subsequence corresponding to a) *lonely*, weakly normed repeat.

Let  $\phi'$  be the part of  $\phi$  until the first state of the first lonely weakly normed repeat  $\chi$ .  $\phi$  (and therefore  $\phi'$ ) starts in the root and passes only through pairwise non bisimilar states. So  $\phi'$  is a entry for  $\chi$  and by assumption  $\phi'$  is weakly normed repeat invariant for  $\chi$  as well. But then  $\chi$  has two bisimilar states, i.e. its first and last state and hence  $\phi$  too. Contradiction.

Unlike Theorem 3.16 we do not prove for every process that is regular, that it is weakly normed repeat invariant. Here we do need that the specifications allow only process names in  $W_\Delta$  occurring in weakly normed repeats. Without this restriction Lemma 3.18 does not hold. In [BCS95] an ingenious counter example to that extent is given. Based on it, Bernhard Steffen supplied us a specification of a process name which is not weakly normed repeat invariant, i.e. specification  $\{S = a \cdot X \cdot Z, X = b + c \cdot X \cdot Y, Y = d + d \cdot Y + d \cdot Z, Z = e \cdot Z\}$ . As the reader can verify  $S$  has a regular process graph, but  $X \cdot Y \cdot Z \not\equiv X \cdot Z$ , i.e.  $S$  is not weakly normed repeat invariant and the cancellation property does not hold:  $Y \cdot Z \equiv Y \cdot Y \cdot Z \not\equiv Y \cdot Z \equiv Z$ .

Now we begin with a proof of the completeness of the criterion.

**Lemma 3.17** If  $\nu \in W_\Delta^+$ , then there is a weakly deterministic transition subsequence  $\nu \xrightarrow{\alpha_0}_\Delta \dots \xrightarrow{\alpha_n}_\Delta \epsilon$ .

**Proof.** Assume without loss of generality that  $\nu \equiv X \cdot \rho$ . First we prove that a weakly deterministic subsequence from  $\nu$  to  $\rho$  exists. Then by induction on the number of names in  $\nu$  we are done.

With each variable we associate the smallest  $i$  such that it is a member of the set  $W_i$  in Definition 3.5.

Now the sequence is constructed as follows: For each transition we pick the a summand  $a \cdot X \cdot Y$  ( $a$  and  $a \cdot X$  likewise) such that the initial action is unique and the process names  $Y, Z$  are in "lower"  $W_i$ .

**Lemma 3.18** Let  $\nu, \rho, \sigma \in V_\Delta^*$  and  $\nu$  weakly deterministic. If  $\nu \cdot \rho \equiv \nu \cdot \sigma$ , then  $\rho \equiv \sigma$ .

**Proof.** Because  $\nu$  is weakly deterministic, from Lemma 3.17 follows that a weakly deterministic transition subsequence from  $\nu \cdot \rho$  to  $\rho$  exists. A corresponding transition subsequence starts in  $\nu \cdot \sigma$  and ends in  $\sigma$ , because each transition is labeled with an unique action (unique with respect to the head process name in the state). By Definition of bisimulation we conclude that  $\rho \equiv \sigma$ .

In the proof we need the Approximation Induction Principle (AIP), which states that if the unfolded graphs of processes are bisimilar down to an arbitrary depth, then the processes are bisimilar. The principle and its proof are described in [BBK87].

**Proposition 3.19** Let  $\pi_i : \text{BPA} \rightarrow \text{BPA}$ ,  $i \in \mathbb{N}^*$  be the projection operators defined as follows. Let  $a \in \text{Act}$ ,  $p, p_1, p_2 \in \text{BPA}$ ,

$$\begin{aligned}\pi_1(a \cdot p) &= a \\ \pi_i(a) &= a \\ \pi_{i+1}(a \cdot p) &= \pi_i(p) \\ \pi_i(p_1 + p_2) &= \pi_i(p_1) + \pi_i(p_2).\end{aligned}$$

If  $\pi_m(p_1) \Leftrightarrow \pi_m(p_2)$  for all  $m \in \mathbb{N}^*$ , then  $p_1 \Leftrightarrow p_2$ .

The completeness of our criterion is proved by showing for a regular process that the first and last state of a (subsequence corresponding to) a lonely, weakly normed repeat have to be bisimilar.

**Theorem 3.20** If  $r$  is regular, then  $r$  is weakly normed repeat invariant.

**Proof.** Suppose  $\phi$  is an arbitrary entry and  $\chi$  a connecting transition subsequence corresponding to a lonely, weakly normed repeat. Suppose that the last state  $X \cdot \rho$  of  $\phi$  is equal to the first state of  $\chi$  and the last state of  $\chi$  is  $X \cdot \nu \cdot \rho$ . It is easily verified that  $\chi$  can be extended with a transition subsequence  $X \cdot \nu \cdot \rho \rightarrow_{\Delta} \dots \rightarrow_{\Delta} X \cdot \nu^2 \cdot \rho \dots$  etc.. Because  $X$  is weakly normed,  $r$  can reach states  $\nu^n \cdot \rho$  for every  $n \in \mathbb{N}$ . By assumption  $\nu \in W_{\Delta}^+$  and hence  $\nu^m \in W_{\Delta}^+$  for  $m > 0$ . Because  $r$  is regular, the pigeon hole principle gives that there are smallest  $k, l \in \mathbb{N}$ ,  $k > l$  so that  $\nu^k \cdot \rho \Leftrightarrow \nu^l \cdot \rho$ . A (repeated) application of Lemma 3.18 gives that  $\nu^{k-l} \cdot \rho \Leftrightarrow \rho$ . This implies that for every  $m \in \mathbb{N}$ , (1)  $X \cdot \nu \cdot \nu^{(k-l) \cdot m} \cdot \rho \Leftrightarrow X \cdot \nu \cdot \rho$  and (2)  $X \cdot \nu^{(k-l) \cdot (m+1)} \cdot \rho \Leftrightarrow X \cdot \rho$ . Because  $\nu$  is weakly normed, by Lemma 3.17 there exists a terminating transition (sub)sequence  $\chi \equiv \nu^m \rightarrow_{\Delta} \dots \rightarrow_{\Delta} \epsilon$ . Because  $\nu \not\equiv \epsilon$ ,  $\chi$  has *at least* length  $m$  and so (3) for all  $m, k \in \mathbb{N}^*$ ,  $\pi_m(\nu^{m+1} \cdot \rho) \Leftrightarrow \pi_m(\nu^{m+1+k} \cdot \rho)$ . Therefore for all  $m \in \mathbb{N}^*$ ,  

$$\pi_m(X \cdot \nu \cdot \rho) \stackrel{(1)}{\Leftrightarrow} \pi_m(X \cdot \nu \cdot \nu^{(k-l) \cdot m} \cdot \rho) \stackrel{(3)}{\Leftrightarrow} \pi_m(X \cdot \nu^{(k-l) \cdot (m+1)} \cdot \rho) \stackrel{(2)}{\Leftrightarrow} \pi_m(X \cdot \rho).$$
An application of AIP (Proposition 3.19) now gives that the last states of  $\phi$  and  $\chi$  are bisimilar. So  $\phi$  is a weakly normed repeat invariant entry for  $\chi$ .

It remains to be proved that the criterion is decidable.

**Theorem 3.21** It is decidable if  $r$  is weakly normed repeat invariant.

**Proof.** BPA is finitely branching and the maximal length of an entry is bounded (Lemma 3.14). It is easily verified that the length of a lonely, weakly normed repeat is bounded too. Therefore there are only finitely many possible connecting combinations and they are of bounded length too. Verifying that such a transition sequence is weakly normed repeat invariant for a (transition subsequence corresponding to a) lonely weakly normed repeat can be done effectively, using that bisimulation equivalence is decidable for all context-free processes [CHS92].

Now we can finish with the decidability result.

**Corollary 3.22** The regularity of context-free processes over NRD specifications is decidable.

**Proof.** Immediate from Theorems 3.16, 3.20 and 3.21.

**Lemma 3.23** The number of non bisimilar states of a regular process graph given by a process name over a NRD specification is bounded by  $\sum_{i=0}^{|V_{\Delta}|^2 \cdot 2^{|V_{\Delta}|}} bf^i$ , where  $bf$  is the branching factor <sup>5</sup>.

**Proof.** Let  $R$  be such a process name. Each state can be reached from  $R$  passing only through pair-wise non bisimilar states with some transition sequence  $\phi$ . Lemma 3.14 gives that the maximal length of  $\phi$ , if it uses no weakly normed repeats is  $|V_{\Delta}|^2 \cdot 2^{|V_{\Delta}|}$ . If  $\phi$  is one longer,  $\phi$  has to use a weakly normed repeat. By Theorem 3.20 and the regularity of the process graph,  $R$  is weakly normed repeat invariant. Therefore there are *at least* two bisimilar states in  $\phi$ . This gives rise to a tree with branching factor  $bf$  and a maximal depth bounded by  $|V_{\Delta}|^2 \cdot 2^{|V_{\Delta}|}$ .

## 4 Conclusions

Difference with the previously mentioned papers [MM94, Kru95] <sup>6</sup> is that we distinguish a root process name and allow the process to have perpetual "sub processes". Whereas the criterion for regularity for the classes in [MM94, Kru95] is relatively cheap to check, our criterion has a substantial computational complexity. If we want to use our criterion of weakly normed repeat invariance, this involves checking bisimulation equivalence between *perpetual* context-free processes <sup>7</sup>. Checking bisimulation equivalence of *normed* processes is known to have a polynomial time complexity [HM95], whereas the complexity of checking bisimulation of arbitrary context-free processes is (doubly) exponential [BCS95].

In principle our proof also gives a method to generate the equivalent linear specification for a given specification satisfying our constraints. The naïve way is simply to "list" all non bisimilar states. Start with the root process name and generate new non bisimilar states using the specification and the algorithm in [BCS95]. Lemma 3.23 tells us that this could be very expensive.

**Acknowledgements.** We thank Jan Friso Groote, Alban Ponse and Frits Vaandrager for proof-reading. Olaf Burkhart, Didier Caujal, Bernhard Steffen, Faron Moller, Sjouke Mauw and Colin Stirling are thanked for advice.

<sup>5</sup> i.e. the maximum number of summands in a definition for a process name in the specification.

<sup>6</sup> We refer to the results in BPA, in the last paper is also dealt with parallelism, but this is not treated in this paper.

<sup>7</sup> Note that in our criterion we have to check if entries are repeat normed invariant for lonely normed repeats. This comes down to checking if  $X \cdot \rho$  the last state of the entry and  $X \cdot \sigma \cdot \rho$ , the last state of the lonely normed repeat are bisimilar. This implies that  $X \cdot \rho \uparrow$  and  $X \cdot \sigma \cdot \rho \uparrow$ .

## References

- [BBK87] J.C.M. Baeten, J.A. Bergstra, and J.W. Klop. On the consistency of Koomen's fair abstraction rule. *Theoretical Computer Science*, 51(1/2):129–176, 1987.
- [BBK93] J.C.M. Baeten, J.A. Bergstra, and J.W. Klop. Decidability of Bisimulation Equivalence for Processes generating Context-free Languages. *Journal of the ACM*, 40(3):653–682, 1993.
- [BCS95] O. Burkart, D. Caucal, and B. Steffen. An Elementary Bisimulation Decision Procedure for Arbitrary Context-Free Processes. In *MFCS '95*, volume 969 of *Lecture Notes in Computer Science*, pages 423–433. Springer-Verlag, 1995.
- [BH64] Y. Bar-Hillel. *Language and Information*. Series in Logic. Addison-Wesley, 1964.
- [BP95] D.J.B. Bosscher and A. Ponse. Translating a Process Algebra with Symbolic Data Values to Linear Format. In Uffe H. Engberg, Kim G. Larsen, and Arne Skou, editors, *Proceedings of the Workshop on Tools and Algorithms for the Construction and the Analysis of Systems*, volume NS-95-2 of *BRICS Notes Series*, pages 119–130, 1995.
- [BW90] J.C.M. Baeten and W.P. Weijland. *Process Algebra*. Cambridge Tracts in Theoretical Computer Science 18. Cambridge University Press, 1990.
- [CHS92] S. Christensen, H. Hüttel, and C. Stirling. Bisimulation is Decidable for all Context-free Processes. In W.R. Cleaveland, editor, *Proceedings of CONCUR 92*, volume 630 of *Lecture Notes in Computer Science*, pages 138–147. Springer-Verlag, 1992.
- [HM95] Y. Hirshfeld and F. Moller. Deciding Equivalences in Simple Process Algebras. In A. Ponse, M. de Rijke, and Y. Venema, editors, *Modal Logic and Process Algebra*, volume 53 of *CSLI Lecture Notes*, pages 151–169. CSLI Publications, Stanford, 1995.
- [HM96] Y. Hirshfeld and F. Moller. Decidability Results in Automata and Process Theory. In *Logics for Concurrency: Automata vs Structure*, Springer Lecture Notes in Computer Science, 1996. To appear. Previously presented as lecture notes at the VIII-th Banff Higher Order Workshop "Theories of Concurrency: Structure vs Automata" in 1994.
- [Hol89] Uno Holmer. Translating Static CCS Agents into Regular Form. PMG report 51, Department of Computer Science, Chalmers University of Technology and the University of Göteborg, 1989.
- [HU79] J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley, 1979.
- [Kru95] A. Kručera. Deciding Regularity in Process Algebras. Technical Report RS-95-52, BRICS (Basic Research in Computer Science, Centre of the Danish National Research Foundation), 1995.
- [MM94] S. Mauw and H. Mulder. Regularity of BPA-Systems is Decidable. In Bengt Jonsson and Joachim Parrow, editors, *CONCUR'94: Concurrency Theory*, volume 836 of *Lecture Notes in Computer Science*, pages 34–47. Springer-Verlag, 1994.
- [Pos46] E.L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52:264–268, 1946.