

# DIRECT PROOFS OF THE UNDECIDABILITY OF THE EQUIVALENCE PROBLEM FOR SENTENTIAL FORMS OF LINEAR CONTEXT-FREE GRAMMARS AND THE EQUIVALENCE PROBLEM FOR OL SYSTEMS

G. ROZENBERG

*Department of Computer Science, State University of New York at Buffalo, 4226 Ridge Lea Road,  
Amherst, New York 14226, USA*

*and  
Mathematical Institute, Utrecht University, Utrecht-Uithof, Netherlands*

Received 7 August 1972

formal languages      linear context-free grammars      OL systems      equivalence problems

## 1. Introduction

L-languages were introduced by Lindenmayer [3] for the description of the development of filamentous organisms. Originally they were described in terms of linear arrays of automata, but in later works the formalism was changed to the more linguistic notion of an L-system. This gave rise to various families of developmental languages with an already fairly developed literature (see, e.g., ref. [4] and its references).

One type of L-systems, the so called OL systems (see, e.g., ref. [6]), differs from context-free grammars in the following:

- (i) There is no terminal alphabet; every string derived in the system is an element of its language.
- (ii) The axiom of a OL system is a word of length one or more.
- (iii) Productions are always applied simultaneously; in other words if a word derives another word, productions are applied to all the letters in it.

The above differences, in particular the lack of a terminal alphabet, imply in fact that standard techniques of formal language theory do not apply to OL systems even when standard questions concerning them are asked. A particular example of such a question is the decidability of the *equivalence problem* for OL systems: Is it decidable whether two arbitrary OL systems generate the same language?

In fact, even within the theory of context-free languages, if one abandons the terminal alphabet a number of standard questions become more difficult to answer. A particular example of such a question is the decidability of the *equivalence problem* for sentential forms of context-free grammars:

Is it decidable whether two arbitrary context-free grammars generate the same set of sentential forms?

Both of the above mentioned problems remained open for quite a long time. Recently, Blattner [1] proved that both of the above problems are undecidable. Her elegant solution to these problems is however quite involved (it uses the undecidability of the equivalence problem for a-transducers).

In this note we show very simple direct proofs of the undecidability of both of the above problems. The technique we use is similar to that used in ref. [5].

We assume that the reader is familiar with the notions of context-free and linear context-free grammars (e.g. in the scope of Ginsburg [2]), OL systems [6] and the Post Correspondence Problem (see, e.g., ref. [2]).

We use standard formal language notations (as in ref. [2]).

1. Results

**Theorem 1.** The equivalence problem for sentential forms of linear context-free grammars is undecidable.

*Proof.* Let  $\Sigma$  be an alphabet containing at least two letters. Let  $K = (\alpha_1, \dots, \alpha_n), L = (\beta_1, \dots, \beta_n)$  be an instance of the Post Correspondence Problem over the alphabet  $\Sigma$ . Let  $V = \Sigma \cup \{S, A, B, C, D, E, F\}$ , where  $\Sigma \cap \{S, A, B, C, D, E, F\} = \emptyset$ . Let  $G_{K,L} = (V, \Sigma, P, S)$  be a linear context-free grammar where  $P$  consists of the following productions (in what follows  $\bar{x}$  denotes the mirror image of the word  $x$ ):

- $S \rightarrow a_i E \bar{a}_i$  for every  $i \in \{1, \dots, n\}$ ,
- $S \rightarrow x A x$  for every  $x \in \Sigma$ ,
- $S \rightarrow x B y$  for every  $x, y \in \Sigma$  such that  $x \neq y$ ,
- $A \rightarrow x A x$  for every  $x \in \Sigma$ ,
- $A \rightarrow x b y$  for every  $x, y \in \Sigma$  such that  $x \neq y$ ,
- $A \rightarrow x C$  for every  $x \in \Sigma$ ,
- $A \rightarrow F x$  for every  $x \in \Sigma$ ,
- $B \rightarrow x B$  for every  $x \in \Sigma$ ,
- $B \rightarrow B x$  for every  $x \in \Sigma$ ,
- $B \rightarrow L$ ,
- $C \rightarrow x C$  for every  $x \in \Sigma$ ,
- $C \rightarrow D$ ,

- $E \rightarrow a_i E \bar{a}_i$  for every  $i \in \{1, \dots, n\}$ ,
- $F \rightarrow F x$  for every  $x \in \Sigma$ ,
- $F \rightarrow D$

Let  $H_{K,L} = (V, \Sigma, P_1, S)$  be a linear context-free grammar, where  $P_1 = P \cup \{E \rightarrow D\}$ . Let  $\mathcal{S}(G)$  denote the set of sentential forms of the context-free grammar  $G$ .

It is straightforward to show that

$$\begin{aligned} \mathcal{S}(G_{K,L}) = & \{S\} \cup \{w A \bar{w} : w \in \Sigma^+\} \cup \{w D \bar{u} : w, u \in \Sigma^+, w \neq u\} \cup \{w x z B u \bar{b} \bar{w} : w \in \Sigma^+, u, z \in \Sigma^+, a, b \in \Sigma, a \neq b\} \cup \\ & \cup \{w x C \bar{w} : w, z \in \Sigma^+\} \cup \{w F z \bar{w} : w, z \in \Sigma^+\} \cup \{w E \bar{u} : w = \alpha_{i_1} \dots \alpha_{i_r}, u = \beta_{i_1} \dots \beta_{i_r} \text{ for some } r \geq 1, \\ & i_1, \dots, i_r \in \{1, \dots, n\}\}, \end{aligned}$$

and

$$\mathcal{S}(H_{K,L}) = \mathcal{S}(G_{K,L}) \cup \{w D \bar{u} : w = \alpha_{i_1} \dots \alpha_{i_r}, u = \beta_{i_1} \dots \beta_{i_r} \text{ for some } r \geq 1, i_1, \dots, i_r \in \{1, \dots, n\}\}.$$

Thus  $\mathcal{S}(G_{K,L}) \neq \mathcal{S}(H_{K,L})$  if, and only if, the given instance  $K, L$  of the Post Correspondence Problem has a solution. Hence if the equivalence problem for sentential forms of linear context-free grammars is decidable then so is the Post Correspondence Problem. However the Post Correspondence Problem is not decidable and so theorem 1 holds.

**Corollary 1.** The equivalence problem for sentential forms of context-free grammars is undecidable.

**Theorem 2.** The equivalence problem for OL systems is undecidable.

**Proof:** Let  $\Sigma$  be an alphabet containing at least two letters. Let  $K = \langle \alpha_1, \dots, \alpha_n \rangle$ ,  $L = \langle \beta_1, \dots, \beta_n \rangle$  be an instance of the Post Correspondence Problem over the alphabet  $\Sigma$ . Let  $I_{K,L} = \langle V, P \cup R, S \rangle$ ,  $J_{K,L} = \langle V, P_1 \cup R, S \rangle$  be OL systems  $\dagger$ , where  $V, P, P_1, S$  are defined as in the proof of theorem 1 and  $R = \{a \rightarrow x: a \in \Sigma\} \cup \{D \rightarrow D\}$ . Let  $\mathcal{L}(G)$  denote the language of the OL system  $G$ . Again,  $\mathcal{L}(I_{K,L}) \neq \mathcal{L}(J_{K,L})$  if, and only if, the given instance  $K, L$  of the Post Correspondence Problem has a solution. Hence if the equivalence problem for OL systems is decidable then so is the Post Correspondence Problem. But the Post Correspondence Problem is not decidable and so theorem 2 holds.

#### Note added in proof

The reader may be interested to know that another simple solution to the problem discussed in this paper was (independently) obtained by A. Salomaa (see "On sentential forms of context-free grammars" – manuscript to be published).

#### Acknowledgement

The research for this paper has been supported by NSF Grant GT-998.

The author is indebted to G.T. Herman and K.P. Lee for useful discussions on the subject, and to the referee for a number of useful remarks concerning the presentation of this paper.

#### References

- [1] M. Blattner, The unsolvability of the equality problem for sentential forms of context-free languages, *J. Comp. & Syst. Sciences*, (1972) to appear.
- [2] S. Ginsburg, *The mathematical theory of context-free languages*, McGraw-Hill, New York (1966).
- [3] A. Lindenmayer, Mathematical models for cellular interactions in development, Parts I and II, *J. Theoret. Biol.*, 18 (1968) 280–315.
- [4] A. Lindenmayer, G. Rozenberg, Developmental systems and languages, *Proc. 4 ACM Symp. Theory Comp.*, (1972) 214–221.
- [5] G. Rozenberg, The equivalence problem for deterministic TOL-systems is undecidable, *Information Processing Letters*, 1 (1972) 201 (see also Errata in this issue).
- [6] G. Rozenberg and P. Doucet, On OL languages, *Information & Control*, 19 (1971) 302–318.

$\dagger$  The rule  $D \rightarrow D$  is included in  $R$  because in the definition of a OL system (see ref. [6]) it is required that a rule is provided for every letter in the alphabet of the system considered.