Reprinted from JOURNAL OF COMPUTER AND SYSTEM SCIENCES
Vol 12, No. 1, February 1976
Printed in Belgium

Interactive Languages

YOSHINORI EZAWA, MASAHARU MIZUMOTO,
JUNICHI TOYODA, AND KOKICHI TANAKA

Department of Computer Science, Osaka University, Toyonaka, Osaka, Japan 560

Received April 2, 1974

The notion of interactive languages generated by interactions between two phrase-structure grammars are proposed and discussed. It is shown that a family of context-free languages does not include a family of interactive languages between two regular grammars, and vice versa. The family of interactive languages, however, is not closed under any of the ordinary operations. The paper also includes discussions about n-cyclic interactive languages among n grammars.

INTRODUCTION

Recently, the problem of characterizing the interactions among various systems, such as "computer network systems" and "time sharing systems," has been proposed and discussed by many researchers. There exists a variety of formal systems that may model some interactions among several systems. For example, Gabrielian's "interacting local automata" [2] are such systems.

The objective of this work is to model such interactions among some formal grammars by applying the well-known concept of control words [3]. In this paper, we propose the notions of *interactive systems* and *interactive languages*. An interactive language is a language defined by an interactive system, a system that consists of two grammars each of which is controlled by the words generated by the other.

Apart from the problem of interactions among several grammars, our study of interactive languages has been stimulated by the recent intense efforts to characterize the derivation of chains of formal languages, among which are the very interesting L-languages [6], SF-languages [8], and associate languages [5]. Interactive languages, too, belong to this category. A particular feature of these languages is that they are generated by interactions between two grammars in a simple way.

This paper is divided into four sections: In Section 1, we define the basic concepts and the classes of interactive languages. Section 2 deals with several characterizations of the classes of interactive languages. In Section 3 the nonclosure properties of the families of interactive languages are given. Section 4 introduces the definition of a cyclic interaction among n grammars ($n \geqslant 3$) and shows some characteristics of the n-cyclic interactive languages.

1. Preliminaries

For notations not explained in the sequel, see [4].

Definition 1. A phrase-structure grammar is a 4-tuple $G = (V_N, V_T, P, S)$, where

(1) V_N is a finite set of nonterminal symbols;

(2) V_T is a finite set of terminal symbols;

(3) P is a finite set of productions $\alpha \to \beta$, with α in $(V_N \cup V_T)^* V_N(V_N \cup V_T)^*$ and β in $(V_N \cup V_T)^*$;

(4) S is a start symbol in V_N.

Let $P_l = \{p_1, ..., p_r\}$ be a set of distinct labels for the productions in P. Let

$$S = Q_0 \overset{p_{j(0)}}{\Rightarrow} Q_1 \overset{p_{j(1)}}{\Rightarrow} \cdots \overset{p_{j(m-1)}}{\Rightarrow} Q_m, \qquad m \geqslant 1,$$

be a left-most derivation according to G, where for each i ($0 \leqslant i \leqslant m-1$) the production labeled by $p_{j(i)}$ is $\alpha \to \beta$ in P and there exist w_1 and w_2 such that $Q_i = w_1\alpha w_2$ and $Q_{i+1} = w_1\beta w_2$, where α occurs only once as a substring of $w_1\alpha$. Then the word $c = p_{j(0)} \cdots p_{j(m-1)}$ over the alphabet P_l is termed a *control word* of the derivation, in symbols $g_G(c) = Q_m$. The mapping g_G from $P_l{}^*$ into $(V_N \cup V_T)^*$ is a function, since the word Q_m is uniquely defined by the control word c.

We now introduce the notion of interaction between two phrasestructure grammars. The interaction is formulated as the process wherein each of the two grammars generates the control words for the other, for the derivation of another grammar.

Definition 2. Let $G_1 = (V_{N_1}, V_{T_1}, P_1, S_1)$ and $G_2 = (V_{N_2}, V_{T_2}, P_2, S_2)$ be two phrase-structure grammars, where $V_{T_1} \subseteq P_{l_2}$, $V_{T_2} \subseteq P_{l_1}$ (P_{l_1} and P_{l_2} are sets of labels of P_1 and P_2, respectively). Let

$$w_1 \underset{G_1G_2}{\Rightarrow} w_2$$

be a relation between w_1 and w_2, where each of the following conditions is satisfied.

(1) $w_1, w_2 \in L(G_1) \subseteq V_{T_1}^* \subseteq P_{l_2}^*$.

(2) $\exists\, y \in L(G_2) \subseteq V_{T_2}^* \subseteq P_{l_1}^*$, $\quad y = g_{G_2}(w_1) \qquad$ and $\qquad g_{G_1}(y) = w_2$.

In this context, the grammar G_2 is named an *associate grammar* and the word y over the alphabet V_{T_2} is called an *associate word* of w_2. The relation $\Rightarrow_{G_1 G_2}$ is termed an *interactive derivation* between G_1 and G_2. A pair of grammars (G_1, G_2) is named an *interactive system*.

Where there is no cause for confusion, this relation will be abbreviated as $\Rightarrow$, and its reflexive and transitive closure will be denoted by $\Rightarrow^*$.

DEFINITION 3. The *interactive language* $L(G_1, G_2; w_0)$ generated by an interactive system (G_1, G_2) consists of those words in $V_{T_1}^*$ that can be interactively derived from the initial word w_0 in $L(G_1)$. Formally,

$$L(G_1, G_2; w_0) = \{w \in L(G_1) | \; w_0 \overset{*}{\underset{G_1 G_2}{\Rightarrow}} w\}.$$

A set of associate words generated by the associate grammar G_2 such that

$$A(G_1, G_2; w_0) = L(G_1, G_2; a_0), \qquad \text{where} \quad a_0 = g_{G_2}(w_0),$$

is called an *associate language* of the interactive language $L(G_1, G_2; w_0)$.

The family of interactive languages is denoted by $\mathscr{I}_{i-j}$ if the grammar G_1 is type i, in Chomsky's sense, and the type of the associate grammar is j ($i, j \in \{0, 1, 2, 3\}$). (A right-linear (or left-linear) grammar is called type 3.)

EXAMPLE 1. The following is an example of an interactive language where $G_1 = (\{A\}, \{a, b\}, P_1, A)$ and $G_2 = (\{B\}, \{1, 2\}, P_2, B)$.

$$P_1 \begin{cases} (1)\; A \to aA \\ (2)\; A \to b, \end{cases} \qquad P_2 \begin{cases} (\text{a})\; B \to 11B \\ (\text{b})\; B \to 2. \end{cases}$$

If ab is an initial word, then there exists an interactive derivation chain as follows:

$$\begin{array}{ccccccccccc} ab & \Rightarrow & aab & \Rightarrow & aaaab & \Rightarrow & \cdots & \Rightarrow & a^{2^n}b & \Rightarrow & \cdots . \\ \searrow & \nearrow & \searrow & \nearrow & \searrow & & & \nearrow & \searrow & & \\ & 112 & & 11112 & & 1^8 2 & & 1^{2^n}2 & & 1^{2^{n+1}}2 & \end{array}$$

The interactive language $L(G_1, G_2; ab)$ is $\{a^{2^n}b \mid n \geqslant 0\}$, and this is a context-sensitive language.

EXAMPLE 2. This example shows that various types of languages can be generated if the appropriate initial words are selected. Let $G_1 = (\{C\}, \{a, b, c\}, P_1, C)$ and $G_2 = (\{D\}, \{1, 2, 3\}, P_2, D)$, where

$$P_1 \begin{cases} (1)\; C \to aC \\ (2)\; C \to bC \\ (3)\; C \to c, \end{cases} \qquad P_2 \begin{cases} (\text{a})\; D \to 12D \\ (\text{b})\; D \to 2D \\ (\text{c})\; D \to 3. \end{cases}$$

Case 1. Let the initial word be bc; then there exists an interactive derivation chain

$$\begin{array}{ccccccccccc} bc & & \Rightarrow & bc & \Rightarrow & \cdots & \Rightarrow & bc & & \Rightarrow & \cdots . \\ & \searrow & \nearrow & & \searrow & & \nearrow & & \searrow & & \\ & & 23 & & 23 & & 23 & & 23 & & \end{array}$$

So the interactive language is $L(G_1, G_2; bc) = \{bc\}$. This is a finite language.

Case 2. Let the initial word be ac; then there exists an interactive derivation chain as follows:

$$\begin{array}{ccccccccccc} ac & & \Rightarrow & abc & \Rightarrow & abbc & \Rightarrow & \cdots & \Rightarrow & ab^nc & \Rightarrow \cdots . \\ & \searrow & \nearrow & \searrow & \nearrow & \searrow & & & \nearrow & \searrow & \\ & 123 & & 1223 & & 12^33 & & & 12^n3 & & 12^{n+1}3 \end{array}$$

So the interactive language is $L(G_1, G_2; ac) = \{ab^nc \mid n \geqslant 0\}$. This is a regular language, but not finite.

Case 3. Let the initial word be aac; then the interactive language $L(G_1, G_2; aac) = \{ab^nab^nc \mid n \geqslant 0\}$ is interactively derived in the same manner as in Case 2. This is a context-free language, but not a regular language.

Case 4. Let the initial word be $aaac$; then the interactive language

$$L(G_1, G_2; aaac) = \{ab^nab^nab^nc \mid n \geqslant 0\}$$

is interactively derived in the same manner as in Case 2. This is a context-sensitive language, but not a context-free language.

In each case, the type of the associate language is shown to be the same as that of the interactive language.

The interactive languages of Examples 1 and 2 are the elements of a family $\mathscr{I}_{3\text{-}3}$, but Example 3 is an element of a family $\mathscr{I}_{2\text{-}3}$.

EXAMPLE 3. The following interactive language is a well-known context-sensitive language and the associate language is a regular language. Let $G_1 = (\{S, X, Y, Z\}, \{a, b, c, \#, \$\}, P_1, S)$ and $G_2 = (\{T\}, \{1, 2, 3, 4, 5, 6, 7\}, P_2, T)$, where

$$P_1 \begin{cases} (1)\ S \rightarrow \#XYZ\$ \\ (2)\ X \rightarrow aX \\ (3)\ Y \rightarrow bY \\ (4)\ Z \rightarrow cZ \\ (5)\ X \rightarrow a \\ (6)\ Y \rightarrow b \\ (7)\ Z \rightarrow c, \end{cases} \qquad P_2 \begin{cases} (\#)\ T \rightarrow 1T \\ (a)\ T \rightarrow T \\ (b)\ T \rightarrow T \\ (c)\ T \rightarrow 234T \\ (\$)\ T \rightarrow 567. \end{cases}$$

If the initial word is $\#abc\$$, then there exists an interactive derivation chain

$$\begin{array}{ccccccccc} \#abc\$ & \Rightarrow & \#aabbcc\$ & \Rightarrow & \cdots & \Rightarrow & \#a^nb^nc^n\$ & \Rightarrow & \cdots . \\ & \searrow \; 1234567 \; \nearrow & & \searrow \; 1(234)^2567 & & 1(234)^{n-1}567 \; \nearrow & & & \end{array}$$

Consequently, the interactive language is $L(G_1, G_2; \#abc\$) = \{\#a^nb^nc^n\$ \mid n \geqslant 1\}$ and the associate language is $A(G_1, G_2; \#abc\$) = \{1(234)^n 567 \mid n \geqslant 1\}$.

2. The Families of Interactive Languages and Chomsky's Hierarchy

In this section, the relationship between the classes of interactive languages and Chomsky's hierarchy is discussed.

The type of an interactive language $L(G_1, G_2; w_0)$ depends upon the type of its associate grammar G_2 rather than that of its direct generating grammar G_1. In particular, the interactive language whose associate grammar is a linear grammar has an interesting characteristic.

Lemma 1. *If the language L contains two words w_1aa and w_2bb (where w_1, w_2 in $V_T{}^*$ and a, b in V_T), it is not a member of a family of interactive language whose associate grammars are linear grammars.*

Proof. Consider the case in which the interactive language L contains two words w_1aa and w_2bb, and its associate grammar G_a is a linear grammar. There is no loss of generality in supposing that there exists an interactive derivation chain $w_1aa \Rightarrow^* w_2bb$.

Since the associate grammar G_a is a linear grammar, each of its productions has either form (i) or form (ii):

$$\begin{array}{ll} \text{(i)} & A \rightarrow uBv, \\ & \qquad\qquad A, B \text{ in } V_{N_a} \quad \text{and} \quad u, v \text{ in } V_{T_a}^* . \\ \text{(ii)} & A \rightarrow u, \end{array}$$

If the production whose label is a has the form (i), the derivation in G_a controlled by w_1aa should generate a sentential form that contains a nonterminal symbol. This contradicts the hypothesis $w_1aa \Rightarrow^* w_2bb$.

On the other hand, if the production's form is (ii), there exists no derivation chain whose control word is w_1aa. The reason is that every sentential form derived by linear grammars has at most one nonterminal symbol. This also contradicts the hypothesis.

Q.E.D.

COROLLARY 2. *The interactive language whose associate grammar is a linear grammar never contains two words w_1aw_2a and w_3bw_4b (where w_1, w_2, w_3, w_4 in $V_{T_1}^*$ and a, b in V_{T_1}).*

Proof. It is clear from the proof of Lemma 1. Q.E.D.

THEOREM 3. *The family of interactive languages $\mathscr{I}_{3\text{-}3}$ does not include the usual families $\mathscr{L}_2$, $\mathscr{L}_3$, and $\mathscr{L}_F$, and vice versa ($\mathscr{L}_F$ denotes a family of finite languages).*

Proof. It is easy to prove part (i) by Lemma 1. Part (ii) is proved by Example 1, and it is also known that the interactive families $\mathscr{I}_{3\text{-}3}$ and $\mathscr{L}_F$ are not disjoint. Q.E.D.

LEMMA 4. $\mathscr{I}_{3\text{-}3} \subsetneqq \mathscr{I}_{3\text{-}2}$.

Proof. It is known from Corollary 2 and Example 4. Q.E.D.

EXAMPLE 4. Let $G_1 = (\{A\}, \{a, b, c, d\}, P_1, A)$ and $G_2 = (\{S, T\}, \{0, 1, 2, 3, 4\}, P_2, S)$, where

$$P_1 \begin{cases} (0)\ A \to dA \\ (1)\ A \to aA \\ (2)\ A \to cA \\ (3)\ A \to bA \\ (4)\ A \to b \\ (5)\ A \to c, \end{cases} \qquad P_2 \begin{cases} (a)\ S \to 1ST \\ (b)\ T \to 3 \\ (c)\ S \to 12 \\ (d)\ S \to 0S4. \end{cases}$$

Then the interactive language is $L(G_1, G_2; dc) = \{da^ncb^n \mid n \geqslant 0\}$. This is a well-known context-free language and in a family $\mathscr{I}_{3\text{-}2}$.

LEMMA 5. $\mathscr{I}_{2\text{-}3} \subsetneqq \mathscr{I}_{2\text{-}2}$.

Proof. Obviously it is obtained from Lemma 1 and Example 4. Q.E.D.

THEOREM 6. $\mathscr{I}_{3\text{-}3} \subsetneqq \mathscr{I}_{2\text{-}2}$.

Proof. The relations $\mathscr{I}_{3\text{-}3} \subseteq \mathscr{I}_{2\text{-}3} \subseteq \mathscr{I}_{2\text{-}2}$ are direct consequences of the definitions. Therefore, the proof is clear from Lemma 5. Q.E.D.

In our previous examples, the lengths of the control words increase monotonously. The next proposition shows that this is not true in general.

PROPOSITION 7. *There exists an interactive derivation chain in which the lengths of the control words do not increase monotonously.*

EXAMPLE 5. This example may complete the proof for Proposition 7. Let $G_1 = (\{S\}, \{a_0, a_1, b, \alpha, \beta\}, P_1, S)$ and $G_2 = (\{A\}, \{0, 1, 2, 3, 4\}, P_2, A)$, where

$$P_1 \begin{cases} (0)\ S \to S \\ (1)\ S \to bS \\ (2)\ S \to a_0a_1S \\ (3)\ S \to \alpha \\ (4)\ S \to b\beta, \end{cases} \qquad P_2 \begin{cases} (a_0)\ A \to 0A \\ (a_1)\ A \to 1A \\ (b)\ A \to 2A \\ (\alpha)\ A \to 4 \\ (\beta)\ A \to 3. \end{cases}$$

If the initial word is $b\beta$, there exists an interactive derivation chain as follows.

$$b\beta \Rightarrow a_0a_1\alpha \Rightarrow bb\beta \Rightarrow (a_0a_1)^2\alpha \Rightarrow b^3\beta \Rightarrow \cdots \Rightarrow b^i\beta \Rightarrow (a_0a_1)^i\alpha \Rightarrow b^{i+1}\beta \Rightarrow \cdots.$$

The lengths of the control words do not change monotonously, as shown in Fig. 1.

Let us consider the properties of languages that are not members of an interactive family, to characterize indirectly the family of interactive languages.

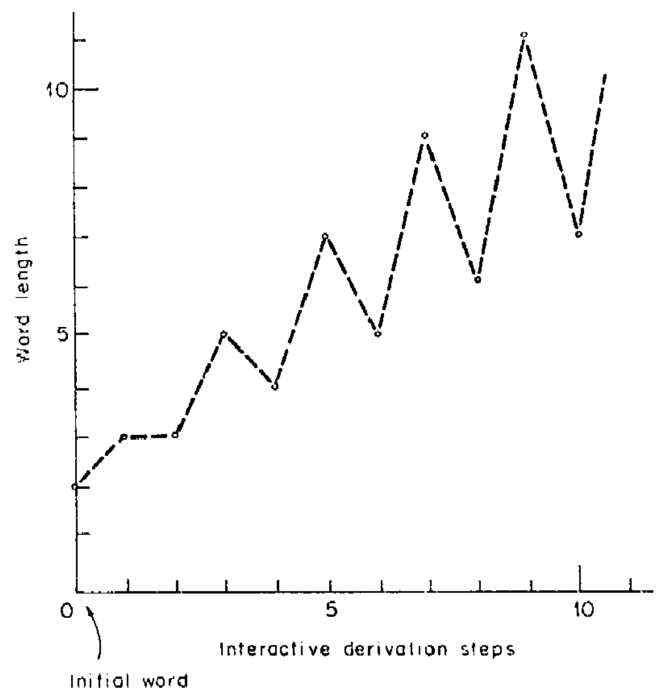

FIG. 1. Relationship between the lengths of words and interactive derivation steps.

LEMMA 8. *The emty set, ϕ, is not an interactive language.*

Proof. Every interactive language contains an initial word. Q.E.D.

LEMMA 9. *The language that contains two words bw_1b and cw_2c (where b, c in V_T and w_1, w_2 in V_T*) is not an interactive language.*

Proof. If the interactive language L contains two words bw_1b and cw_2c, there exists an interactive derivation chain such that $bw_1b \Rightarrow^* cw_2c$. So we can suppose that there exists a word y in $L(G_a)$ such that $y = g_{G_a}(bw_1b)$. That is,

$$S_a \overset{b}{\Rightarrow} z_1 \overset{w_1}{\Rightarrow} z_2 \overset{b}{\Rightarrow} y.$$

The production labeled (b) must have the form

$$\text{(b)} \qquad S_a \to z_1 \in P_a .$$

Since the sentential form z_1 contains some nonterminal symbols, y also contains some nonterminal symbols. This contradicts the hypothesis that y is in $L(G_a)$. Q.E.D.

In the many characterizations of the interactive languages, the property mentioned in the next lemma is very interesting and inherent.

LEMMA 10. *If the interactive language L contains two words w and z, where* $w \Rightarrow z$, *then* $|z| \leqslant K \cdot |w|$. *Moreover, the length of the derivation chain of z is less than* $M \cdot |w|$ *(K and M are constant numbers).*

Proof. Suppose that the interactive language L is generated by an interactive system (G_1, G_2). Let K_1 and K_2 be the maximum lengths of the strings on the right-hand sides of the production sets P_1 and P_2, respectively. From the assumption, there exists a relation $w \Rightarrow_{G_1G_2} z$. Therefore,

$$|g_{G_2}(w)| \leqslant K_2 \cdot |w|,$$

and

$$|z| = |g_{G_1}(g_{G_2}(w))| \leqslant K_1 \cdot |g_{G_2}(w)| \leqslant K_1 \cdot K_2 \cdot |w|.$$

As the word $g_{G_2}(w)$ is a control word for z, if we set $K = K_1 \cdot K_2$ and $M = K_2$, then the lemma is clear. Q.E.D.

COROLLARY 11. *The infinite language in which each word's length is represented by* $n!$ $(n \geqslant 0)$ *is not an interactive language.*

Proof. Obvious from Lemma 10. Q.E.D.

Considering Lemma 10, it is known that the following well-known context-sensitive grammar G_s cannot generate the language $L(G_s)$ in the interactive way by any type 0 associate grammar.

EXAMPLE 6. Let $G_s = (\{X_0, X_1, X_2\}, \{a, b, c\}, P_s, X_0)$, where

$$P_s \begin{cases} (1)\ X_0 \to aX_0X_1X_2 \\ (2)\ X_0 \to abX_2 \\ (3)\ X_2X_1 \to X_1X_2 \\ (4)\ bX_1 \to bb \\ (5)\ bX_2 \to bc \\ (6)\ cX_2 \to cc. \end{cases}$$

The language generated by this grammar is $L(G_s) = \{a^n b^n c^n \mid n \geqslant 1\}$ [7]. In general, a word $a^k b^k c^k$ has the derivation chain

$$X_0 \overset{*}{\Rightarrow} a^{k-1}X_0(X_1X_2)^{k-1} \Rightarrow a^k b(X_2X_1)^{k-1} X_2 \overset{*}{\Rightarrow} a^k b X_1^{k-1} X_2^{k} \overset{*}{\Rightarrow} a^k b^k X_2^{k} \overset{*}{\Rightarrow} a^k b^k c^k.$$

Thus, it requires a control word whose length is greater than $(k^2 + 5k - 2)/2$ to generate interactively the word $a^k b^k c^k$ by this grammar and a certain associate grammar. Accordingly, it is impossible to generate interactively the words longer than $K \cdot |w_0|$, where K is a constant number and $|w_0|$ is the length of the initial word.

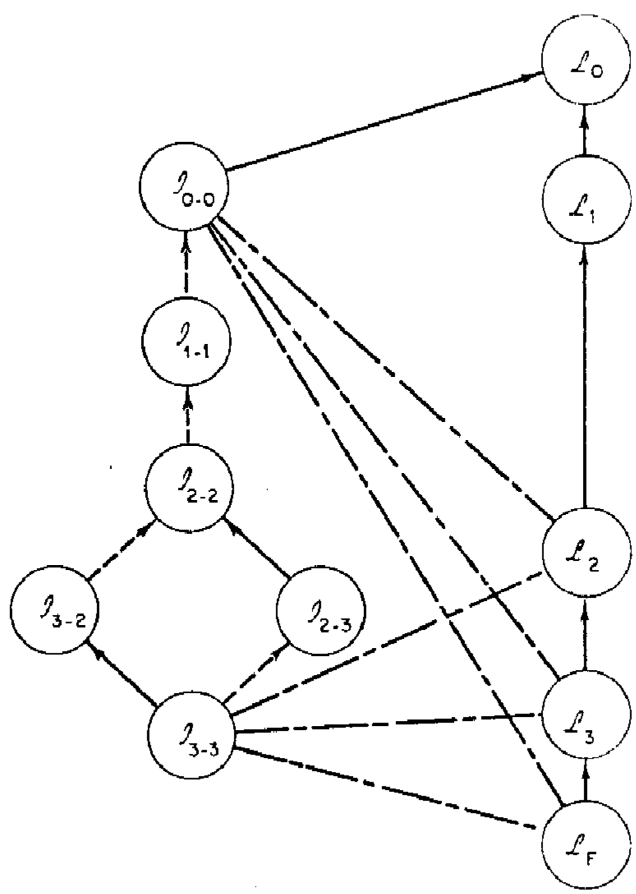

FIG. 2. The hierarchy of interactive languages. ——→, $\not\subseteq$; - - -→, $\subsetneq$; — — —, incomparable.

LEMMA 12. *An infinite language* $\{ww \mid w \text{ in } V_T^*\}$ *is not an interactive language.*

Proof. Suppose that an infinite language $\{ww \mid w \text{ in } V_T^*\}$ is an interactive language. There must exist a derivation whose control word is $ww = a_1 \cdots a_n a_1 \cdots a_n$ (a_i in V_T, $1 \leqslant i \leqslant n$). The sentential form derived from a start symbol S with a control word w must contain at least one nonterminal symbol S so as to apply a production labeled a_1. In that case, the sentential form derived from S with the control word ww still contains at least one nonterminal symbol S. This is a contradiction. Q.E.D.

COROLLARY 13. *The language L on a one-letter alphabet* $\{a\}$ *is an interactive language if and only if* $|L| = 1$ *or* $|L| = 2$; *in the latter case the initial word is a.*

THEOREM 14. $\mathscr{I}_{0\text{-}0} \subsetneqq \mathscr{L}_0$.

Proof. The proper part of the relation is obvious from Lemma 10. Q.E.D.

Figure 2 summarizes the results described in this section.

3. CLOSURE PROPERTIES

Interactive languages are remarkable for their nearly complete lack of closure properties under the operations usually considered. Similar to the classes of L-languages [6] and SF-languages [8], it seems to be due to the fact that every string appearing in the interactive derivation processes is also an element of the language.

THEOREM 15. *The family of interactive languages* $\mathscr{I}_{0\text{-}0}$ *is not closed with respect to*

(i) *union,*
(ii) *complement,*
(iii) *intersection,*
(iv) *the star operator* $(*)$,
(v) *concatenation,*
(vi) *homomorphisms,*
(vii) *inverse homomorphisms,*
(viii) *intersection with regular sets,*
(ix) *mirror image (reverse).*

Proof. We make use of the following interactive languages to provide a counterexample for each operation.

[I]

$$G_1 = (\{A\}, \{a\}, P_1, A), \qquad G_2 = (\{B\}, \{0, 1\}, P_2, B),$$

$$P_1 \begin{cases} (0)\ A \to aA \\ (1)\ A \to a, \end{cases} \qquad P_2 = \{(a)\ B \to 12\}.$$

Thus, $H_1 = L(G_1, G_2; a) = \{a, aa\}$ and $H_2 = L(G_1, G_2; aaa) = \{aaa\}$.

$$G_3 = (\{C\}, \{a, b, \alpha\}, P_3, C), \qquad G_4 = (\{D\}, \{0, 1, 2\}, P_4, D)$$

[II]
$$P_3 \begin{cases} (0)\ C \to aC \\ (1)\ C \to \alpha C \\ (2)\ C \to \alpha b, \end{cases} \qquad P_4 \begin{cases} (\text{a})\ D \to 0D \\ (\text{b})\ D \to 2 \\ (\alpha)\ D \to 1D. \end{cases}$$

Thus, $H_3 = L(G_3, G_4\ a\alpha b) = \{a\alpha^n b \mid n \geqslant 1\}$.

(i) A trivial counterexample is $H_1 \cup H_2$; the component sets are in $\mathscr{I}_{0\text{-}0}$, but their union is not, for see Corollary 13.

(ii) Immediate from Corollary 13, the complement of H_1, $V_{T_1}^* - H_1$, is not an interactive language.

(iii) The intersection of H_1 and H_2 is equal to $\varnothing$; it follows directly from Lemma 8.

(iv) Again considering Corollary 13, it is obvious that $H_1{}^*$ is not in $\mathscr{I}_{0\text{-}0}$.

(v) Obviously, $H_1 \cdot H_2 = \{a^4, a^5\}$, but this is not in $\mathscr{I}_{0\text{-}0}$.

(vi) Consider a homomorphism defined by $h(\text{a}) = c$, $h(\alpha) = \alpha$, and $h(\text{b}) = c$; $h(H_3) = \{c\alpha^n c \mid n \geqslant 1\}$. This is not an interactive language.

(vii) Again $\varnothing$ can be used as a proof: If $h_2(\text{b}) = aa$, then $h_2^{-1}(H_2) = \varnothing$.

(viii) $\varnothing$ is a regular set: so intersection with it provides a counterexample.

(ix) As shown in Example 1, the language $\{a^{2^n}b \mid n \geqslant 0\}$ is in $\mathscr{I}_{0\text{-}0}$. Consider its mirror image, $\{ba^{2^n} \mid n \geqslant 0\}$. Is this an interactive language? If that is the case, there exists next, an interactive derivation chain for some positive integers, k, m, r.

$$(**) \qquad ba^k \Rightarrow ba^m \Rightarrow ba^r \qquad (k < m < r).$$

It means that there are two derivation chains according to the associate grammar.

$$S_2 \overset{b}{\Rightarrow} z \overset{a^k}{\Rightarrow} y_k, \qquad S_2 \overset{b}{\Rightarrow} z \overset{a^k}{\Rightarrow} y_k \overset{a^{m-k}}{\Rightarrow} y_m.$$

Depending upon the former relation in (**), the sentential form y_k does not contain any nonterminal symbols. According to the latter chain in (**), however, y_k must contain at least one nonterminal symbol. Consequently, the desired relationship (**) is not realized, and the family of interactive languages is not closed under mirror image operation. Q.E.D.

4. Cyclic Interactions among n Grammars

In the preceding sections, we discussed the interaction between two grammars. But in the real systems, many cases contain the problem of interactions among more than two systems. In this section, cyclic interactions among n grammars are defined and some characteristics of them are also discussed.

DEFINITION 4. Let $G_1, \ldots, G_n$ ($n \geqslant 3$) be phrase-structure grammars, and extend the relation $\Rightarrow_{G_1 G_2}$ in Definition 2 to

$$w_1 \underset{G_1 \cdots G_n}{\mapsto} w_2 ,$$

where the following conditions are satisfied.

$$\begin{aligned} w_1 &= x_1 \text{ in } L(G_1), \\ g_{G_2}(x_1) &= x_2 \text{ in } L(G_2), \\ &\vdots \\ g_{G}(x_{n-1}) &= x_n \text{ in } L(G_n), \\ g_{G_1}(x_n) &= w_2 \text{ in } L(G_1). \end{aligned}$$

Note that $V_{T_i} \subseteq P_{l_{i+1}}$ ($i < n$) and $V_{T_n} \subseteq P_{l_1}$.

The n-cyclic interactive language generated by an ordered n-tuple of grammars $(G_1, \ldots, G_n)$ is defined in the same manner as in Definition 3.

$$CL^n(G_1, \ldots, G_n ; w_0) = \{w \in L(G_1) \mid w_0 \overset{*}{\underset{G_1 \cdots G_n}{\Rightarrow}} w\}.$$

Then an n-cyclic associate language is also defined, as

$$\begin{aligned} CA^n(G_1, \ldots, G_n ; w_0) = \{x_{2_i}, \ldots, x_{n_i} \mid w_0 \overset{*}{\underset{G_1 \cdots G_n}{\Rightarrow}} w_i \underset{G_1 \cdots G_n}{\mapsto} w_{i+1} , g_{G_2}(w_i) \\ = x_{2_i}, \ldots, g_{G_n}(x_{n_i - 1}) = x_{n_i} , g_{G_1}(x_{n_i}) = w_{i+1}\}. \end{aligned}$$

In this context, the n-tuple of grammars is called an n-cyclic interactive system.

DEFINITION 5. An n-cyclic interactive system $(G_1, \ldots, G_n)$ is called nonblocked if, whenever there exists an interactive derivation chain from an initial word w_0 to a word w in $L(G_1)$, there exists a direct interactive derivation from the word w to some word w' in $L(G_1)$.

THEOREM 16. *Let* $(G_1, \ldots, G_n)$ *be a nonblocked n-cyclic interactive system, where* $n > 2$ *and one grammar in* $\{G_i \mid 1 \leqslant i \leqslant n, i \neq 2\}$ *is type* 3. *Then there exists a nonblocked n-cyclic interactive system* $(G_1', \ldots, G_{n-1}')$ *such that* $C\ L^n(G_1, \ldots, G_n ; w_0) = CL^{n-1}(G_1', \ldots, G_{n-1}' ; w_0')$.

Proof. Let the cyclic interaction among $G_1, \ldots, G_n$ be shown as in Fig. 3a and let G_i ($i \neq 2$) be type 3. It is sufficient to construct a new grammar G_{new} that simulates

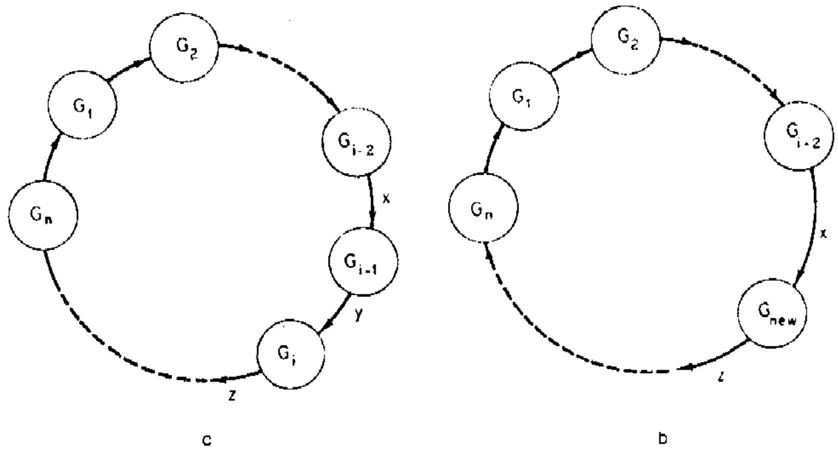

FIG. 3. An n-cyclic interaction among n grammars.

the interaction from G_{i-1} to G_i (see Fig. 3b). That is, if $\forall x$ in $L(G_{i-2})$, $\exists! y$ in $L(G_{i-1})$, $\exists! z$ in $L(G_i)$; $g_{G_{i-1}}(x) = y$ and $g_{G_i}(y) = z$, then $g_{G_{\text{new}}}(x) = z$ (where $y = y_1 \cdots y_m$, y_j in P_{l_i} for $1 \leqslant j \leqslant m$).

⟨i⟩ If G_i is a right-linear grammar, a production labeled y_j in P_i has the form

$$(y_j) \qquad A \rightarrow \alpha_j B$$

or

$$(y_j) \qquad A \rightarrow \alpha_j ,$$

where α_j in $V_{T_{i-2}}^* \subseteq P_{l_{i-1}}^*$ and A, B in $V_{N_{i-2}}^*$. Therefore,

$$g_{G_i}(y) = g_{G_i}(y_1 \cdots y_m) = \alpha_1 \cdots \alpha_m = z.$$

In general, the productions in P_{i-1} are supposed to be as

$$(x_k) \qquad \beta \rightarrow \delta,$$

where β and δ are in $(V_{T_{i-1}} \cup V_{N_{i-1}})^*$. The substituted grammar G_{new} has the productions

$$(x_k) \qquad \tau(\beta) \rightarrow \tau(\delta),$$

where $\tau(\beta)$ is a string in which each terminal symbol y_j in β is replaced by α_j. In the case where $i = 1$, let $i - 1 = n$ and $i - 2 = n - 1$. Furthermore, if $i = 1$, then the following production should be contained in P_{new} since the initial word w_0 must be in $L(G_{\text{new}})$.

$$(\$) \qquad S_{i-1} \rightarrow w_0 , \qquad \text{where} \quad \{\$\} \cap P_{l_{i-2}} = \varnothing .$$

Finally, we have a new grammar $G_{\text{new}} = (V_{N_{i-1}}, V_{T_i}, P_{\text{new}}, S_{i-1})$.

⟨ii⟩ If the grammar G_i is a left-linear grammar, the above grammar G_{new} should have the productions

$$(x_k) \qquad \tau(\beta^R) \to \tau(\delta^R),$$

where β^R is a mirror image of the string β. Q.E.D.

THEOREM 17. *For any nonblocked n-cyclic interactive system* $(G_1, \ldots, G_n)$, *where* $n > 2$ *and each of the grammars is type* 3, *there exists a nonblocked* $(n-1)$*-cyclic interactive system* $(G_1', \ldots, G_{n-1}')$ *with* $CL^n(G_1, \ldots, G_n; w_0) = CL^{n-1}(G_1', \ldots, G_{n-1}'; w_0')$, *where* $G_1', \ldots, G_{n-1}'$ *are also type* 3.

Proof. This theorem is really a corollary to Theorem 16. Since the substitution used in the former proof is a homomorphism, the type of the new grammar G_{new} is just the same as that of the grammar G_{i-1}. Q.E.D.

It is an open problem whether Theorems 16 and 17 hold or not in the case where the nonblocked condition is not satisfied.

CONCLUSION

An interaction between various systems is formalized by the use of formal language techniques. The classes of interactive languages are discussed and compared with Chomsky's hierarchy. It is shown that the family of interactive languages is not closed under usual operations. However, it is an open problem whether the interactive languages between context-sensitive grammars are recursive. Furthermore, the cyclic interactions among n grammars are also discussed.

For a further study, the characterization of the cyclic interactions among n grammars of any types and the reasonable different definitions of the interactions among many grammars will be interesting topics.

ACKNOWLEDGMENT

The authors gratefully acknowledge the referee's detailed comments and suggestions, which led to considerable improvement.

REFERENCES

1. S. ABRAHAM, Compound and serial grammars, *Inform. Contr.* **20** (1972), 432–438.
2. A. GABRIELIAN, The theory of interacting local automata, *Inform. Contr.* **16** (1970), 360–377.

3. S. Ginsburg and E. H. Spanier, Control sets on grammars, *Math. Systems Theory* 2 (1968), 159–177.
4. J. E. Hopcroft and J. D. Ullman, "Formal Languages and Their Relation to Automata," Addison-Wesley, Reading, Mass., 1969.
5. E. Moriya, Associate languages and derivational complexity of formal grammars and languages, *Inform. Contr.* 22 (1973), 139–162.
6. A. Lindenmayer, Mathematical models for cellular interaction in development, I, II, *J. Theoret. Biol.* 18 (1968), 280–315.
7. A. Salomaa, "Formal Languages," Academic Press, New York, 1973.
8. A. Salomaa, On sentential forms of context-free grammars, *Acta Inform.* 2 (1973), 40–49.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.