Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia

Vol. 6, 2022, Page 81-85

Characteristics of Formal Languages Generated by Bonded Systems

Ahmad Firdaus Yosman[a]*, Wan Heng Fong[a], Hazzirah Izzati Mat Hassim[a]
Department of Mathematical Sciences, Faculty of Science,
Universiti Teknologi Malaysia, 81310 Johor, Malaysia
*Corresponding author: afirdaus49@graduate.utm.my

Abstract
The field of formal language theory involves the construction of words or sentences through a meticulous framework called a system. The set of words derived from the same system is called a language, where it is readily capable of undergoing set operations such as union, concatenation, concatenation closure, λ-free concatenation closure, homomorphism, inverse homomorphism, and intersection. Recently, a system depicting chemical bonding that generates languages in a parallel manner has been introduced, called bonded parallel insertion-deletion systems. The systems utilize the operations of insertion and deletion on letters with integers attached on either side to produce words where the sum of consecutive integers is zero. In this case, the integers are the bonds of each letter. The system is said to be parallel since the application of the insertion and deletion rules occurs simultaneously at each possible position. The aim of this research is to determine the characteristics of the languages generated by these systems under various set operations.
Keywords: insertion-deletion systems; bonded systems; parallel systems; formal language theory; closure properties

Introduction
The advent of parallelism in the modern day has brought upon rapid advancement in technology beyond even the predictions of Moore [1]. The efficiency and work rate of computer processes has seen monumental increment, where even the lowest consumer-oriented processors now completely dwarfing top of the line processors of yesteryears. Parallelism, however, was not a man-made concept. Parallelism occurs naturally in life: in the development of cells, in the growth of plants, and in the lineage succession of humanity.

It is this concept of parallelism in nature that became the impetus for Lindenmayer's work in 1968 that birthed the concept of L-systems [2,3], a mathematical model he called developmental systems that utilizes parallel production rules instead of sequential ones. His work blossomed into a rapidly growing area of formal language theory, with research into the classification of the generative power, closure properties, and complexity of L-systems being the focus of numerous researchers [4-10].

More recently, L-systems were used as a comparison to determine the generative power of parallel variants of bonded systems [11-13], which are systems that execute production rules on letters in a bonding alphabet introduced in [14]. The bonding alphabet consists of letters that are attached to integers on either side to represent its bond number, such that these letters can only be inserted or deleted to form a well-formed word. This concept simulates the chemical bonding between atoms during the formation of molecules or compounds.

In this research, the focus is on determining the characteristics of the languages generated by bonded parallel insertion-deletion systems [13] by proving the closure under the set operations of union, concatenation, concatenation closure, λ-free concatenation closure, homomorphism, inverse homomorphism, and intersection with regular languages.

Materials and methods

We present our findings through direct proofs with reference to previous work in this field. The results on closure properties of L-systems have been studied extensively in [8-10] while the closure under set operations have been clarified in [15]. Based on these, the results in this research are obtained. For further clarification, some fundamental concepts and definitions pertinent to this research are presented. The reader may refer to [14-18] for more in-depth reading.

First off, we recall some basic notations in formal language theory as shown in Table 1.

Table 1: Some basic notations in formal language theory

Notations	**Meaning**
$A \subseteq B$	The *inclusion* of a set A in a set B
$A \subset B$	The *proper inclusion* of a set A in a set B
Σ	An *alphabet*, which is a set of symbols
Σ^*	The set of all *words* over the alphabet Σ
$L \subseteq \Sigma^*$	A *language* over an alphabet Σ
L	A *family* of languages is a set of languages
λ	The *empty* word

Next, we present the definition of bonding alphabet in Definition 1.

Definition 1 [14] Let $\mathbb{Z}$ be the set of integers as well as $\mathbb{Z}_0^- = \{0,-1,-2,\ldots\}$ and $\mathbb{Z}_0^+ = \{0,1,2,\ldots\}$.

Let Σ be an alphabet. Then the set $\mathrm{B}_\Sigma = \mathbb{Z}_0^+ \times \Sigma \times \mathbb{Z}_0^-$ is a *bonding alphabet* over Σ. An element $(i,a,-j)$ of B_Σ is called a *letter* a with left bond i and right bond $-j$. To simplify the presentation, $\left[{}_i a_{-j}\right]$ is written instead of $(i,a,-j)$ for a letter a from B_Σ.

Let

$$w = \left[{}_{i_0} a_{1i_1}\right]\left[{}_{i_2} a_{2i_3}\right]\left[{}_{i_4} a_{3i_5}\right]\cdots\left[{}_{i_{2n-2}} a_{ni_{2n-1}}\right]$$

be a nonempty sequence of letters from B_Σ. The sequence w is said to be *well-formed* if all bonds fit, i.e. $i_{2j-1} + i_{2j} = 0$, for $1 \le j \le n-1$, where n is an integer. If additionally, $i_0 + i_{2n-1} = 0$ holds, then w is said to be a *balanced word* or for short a *word*. If $i_0 + i_{2n-1} \neq 0$, then the word is said to be *unbalanced*. Moreover, a word is *neutral* if $i_0 = i_{2n-1} = 0$.

For a well-formed word

$$w = \left[{}_{i_0} a_{1-i_1}\right]\left[{}_{i_1} a_{2-i_2}\right]\left[{}_{i_2} a_{3-i_3}\right]\cdots\left[{}_{i_{n-1}} a_{n-i_n}\right],$$

the word w is said to have the left bond i_0 and the right bond $-i_n$ as the *outer bonds* and $i_1,\ldots,i_{n-1}$ as the *inner bonds*. If the inner bonds are not of interest, then the word is written as

$$\left[{}_{i_0} a_1 a_2 a_3 \ldots a_{n-i_n}\right].$$

The set of all well-formed words built from letters of B_Σ including the empty word is referred to as B_Σ^* while the set of all balanced words built from letters of B_Σ including the empty word is referred to as $\mathrm{B}_\Sigma^\#$. By definition, $\mathrm{B}_\Sigma^\# \subset \mathrm{B}_\Sigma^*$.

The *empty word* is the neutral element of both structures B_Σ^* and $\mathrm{B}_\Sigma^\#$, written as $\left[{}_{i_0} \lambda_{-i_0}\right]$ for some number $i_0 \in \mathbb{Z}_0^+$. The empty word is always a balanced word.

The length of a bond word w from B_Σ^* and $\mathrm{B}_\Sigma^\#$ is denoted by $|w|$ and is equal to the number of letters in w. In particular, the empty bond word is of length 0.

Furthermore, the bond erasing homomorphism as presented in Definition 2 removes the bonds on the derived words to enable comparison of bond words with words from common families in formal language theory.

Definition 2 [14] The bond erasing homomorphism is a homomorphism $h_{be} : \mathrm{B}_{\Sigma}^{\#} \to \Sigma^{*}$ defined by $h_{be}\left(\left[{}_{i}a_{-j}\right]\right) = a$ for every $\left[{}_{i}a_{-j}\right] \in \mathrm{B}_{\Sigma}$.

Now, the definition of bonded parallel insertion-deletion systems (bPINSDEL-systems) is presented in Definition 3.

Definition 3 [13] Let Σ be a finite alphabet, $A \subseteq \mathrm{B}_{\Sigma}^{\#}$ be a finite set of *axioms* that contains only neutral words, and $I, D \subseteq \mathrm{B}_{\Sigma}^{*}$ be a finite set of *insertion strings* and *deletion strings*, respectively, such that the insertion strings in I and the deletion strings in D need not be balanced. A *bonded parallel insertion-deletion system* (bPINSDEL-system) is a quadruple $\pi = (\Sigma, A, I, D)$, where the derivation relation $\Rightarrow_{\pi}$ of a bPINSDEL-system $\pi = (\Sigma, A, I, D)$ is defined as follows: for any $\alpha \in A,\ \alpha \Rightarrow_{\pi} \beta,\ \beta \in \mathrm{B}_{\Sigma}^{\#}$ if and only if at least one of the following is true:

- $\alpha = \alpha_1\alpha_2 \ldots \alpha_n,\ n \geq 1$ and $\alpha_i \in \mathrm{B}_{\Sigma}^{*},\ 1 \leq i \leq n$ are nonempty subwords and there are insertion strings $\delta_i \in I,\ 1 \leq i \leq n+1$ such that $\beta = \delta_1\alpha_1\delta_2\alpha_2 \cdots \delta_{n-1}\alpha_{n-1}\delta_n\alpha_n\delta_{n+1}$ and there are no insertion strings that can be inserted into subwords $\alpha_i \in \mathrm{B}_{\Sigma}^{*}$.
- $\alpha = \delta_1'\alpha_1'\delta_2'\alpha_2'\delta_3' \cdots \delta_{n-1}'\alpha_{n-1}'\delta_n'\alpha_n'\delta_{n+1}',\ n \geq 1$ and $\alpha_i' \in \mathrm{B}_{\Sigma}^{*},\ 1 \leq i \leq n$ are nonempty subwords and there are deletion strings $\delta_i' \in D,\ 1 \leq i \leq n+1$ such that $\beta = \alpha_1'\alpha_2' \cdots \alpha_n'$ and there are no removable deletion strings left inside subwords $\alpha_i' \in \mathrm{B}_{\Sigma}^{*},\ 1 \leq i \leq n,$ i.e. $\left|\alpha_i'\right|_{\delta'} = 0,\ \forall \delta' \in D.$ Here, a removable deletion string is a string that once removed results in a well-formed word.

The reflexive and transitive closure of $\Rightarrow_{\pi}$ is denoted by $\Rightarrow_{\pi}^{*}$. If there is no risk of ambiguity, $\Rightarrow$ and $\Rightarrow^{*}$ are written instead of $\Rightarrow_{\pi}$ and $\Rightarrow_{\pi}^{*}$, respectively. The *language generated* by a bPINSDEL-system $\pi = (\Sigma, A, I, D)$ is defined as

$$L(\pi) = \left\{h_{\text{be}}(\beta) \mid \text{there is an axiom } \alpha \in A \text{ such that } \alpha \Rightarrow_{\pi}^{*} \beta\right\},$$

where h_{be} is the bond erasing homomorphism. The family of all languages generated by bPINSDEL-systems is denoted by L(bPINSDEL).

Results and discussion

In this section, the results of this research are presented as new theorems.

Theorem 1

L(bPINSDEL) is closed under union such that for any $L_1, L_2 \in \mathrm{L}$(bPINSDEL), $L_1 \cup L_2 = \left\{a \mid a \in L_1 \text{ or } a \in L_2\right\} \in \mathrm{L}$(bPINSDEL).

Theorem 2

L(bPINSDEL) is closed under concatenation such that for any $L_1, L_2 \in \mathrm{L}$(bPINSDEL), $L_1L_2 = \left\{a_1a_2 \mid a \in L_1 \text{ and } a \in L_2\right\} \in \mathrm{L}$(bPINSDEL).

Theorem 3

L(bPINSDEL) is closed under concatenation closure such that for any $L \in \mathrm{L}$(bPINSDEL), $L^{*} = \bigcup_{i=0}^{\infty} L^{i} \in \mathrm{L}$(bPINSDEL), where $L^{0} = \{\lambda\}$.

Theorem 4
$\mathcal{L}(\text{bPINSDEL})$ is closed under λ-free concatenation closure such that for any $L \in \mathcal{L}(\text{bPINSDEL})$, $L^{+} = \bigcup_{i=1}^{\infty} L^{i} \in \mathcal{L}(\text{bPINSDEL})$.

Theorem 5
$\mathcal{L}(\text{bPINSDEL})$ is closed under homomorphism such that for a homomorphism $h:\Sigma_1 \to \Sigma_2$ and a language $L_1 \subseteq \Sigma_1^{*},\ h(L_1) = \{h(w) \mid w \in L_1\} \in \mathcal{L}(\text{bPINSDEL})$.

Theorem 6
$\mathcal{L}(\text{bPINSDEL})$ is closed under inverse homomorphism such that for a homomorphism $h:\Sigma_1 \to \Sigma_2$ and a language $L_2 \subseteq \Sigma_2^{*},\ h^{-1}(L_2) = \{w \mid h(w) \in L_2\} \in \mathcal{L}(\text{bPINSDEL})$.

Theorem 7
$\mathcal{L}(\text{bPINSDEL})$ is closed under intersection with regular languages such that for any $L \in \mathcal{L}(\text{bPINSDEL})$, and regular language R, $L \cap R = \{a \mid a \in L \text{ and } a \in R\} \in \mathcal{L}(\text{bPINSDEL})$.

A corollary of these results is presented as follows.

Corollary 1
$\mathcal{L}(\text{bPINSDEL})$ is a full abstract family of languages (AFL).

Conclusion
The characteristics of the family of languages generated by bPINSDEL-systems $(\mathcal{L}(\text{bPINSDEL}))$, specifically their closure properties have been determined, where it is closed under the operations of union, concatenation, concatenation closure, λ-free concatenation closure, homomorphism, inverse homomorphism, and intersection with regular languages. Thus, $\mathcal{L}(\text{bPINSDEL})$ is a full AFL. These results are essential in the advancement of theoretical knowledge in the field of formal language theory as well as contributing to progress in DNA computing through mathematical models of atomic behavior. In the future, variants of bPINSDEL-systems can be introduced by restricting the rules of insertion or deletion. From there, the generative power and characteristics of the new systems can be determined. Other than that, the computational complexity of bPINSDEL-systems is yet to be studied.

Acknowledgement
The first author would like to acknowledge The Ministry of Higher Education (MOHE) for his MyBrainSc scholarship. The authors would also like to MOHE and Research Management Centre (RMC) of Universiti Teknologi Malaysia (UTM) for the financial funding through UTMShine 09G89 and Matching Grant 4B711.

References

[1] Moore G. E. 2006. Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. IEEE Solid-State Circuits Society Newsletter. 11(3): 33-35.

[2] Lindenmayer A. 1968. Mathematical models for cellular interactions in development I. Filaments with one-side inputs. Journal of Theoretical Biology. 18: 280-299.

[3] Lindenmayer A. 1968. Mathematical models for cellular interactions in development II. Simple branching filaments with two-sided inputs. Journal of Theoretical Biology. 18: 300-315.

[4] Herman G. T. 1969. Computing ability of a developmental model for filamentous organisms. Journal of Theoretical Biology. 25: 421-435.

[5] Van Dalen D. 1971. A note on some systems of Lindenmayer. Mathematical Systems Theory. 5: 128-140.

[6] Vitanyi P. M. 1973. Structure of growth in Lindenmayer systems. Indagationes Mathematicae. 76: 247-253.

[7] Pollul W. and Schutt D. 1975. Characterization of growth in deterministic Lindenmayer systems. Indagationes Mathematicae. 78: 183-188.

[8] Vitanyi P. M. 1976. Deterministic Lindenmayer languages, nonterminals and homomorphisms. Theoretical Computer Science. 2: 49-71.

[9] Vitanyi P. M. 1977. Context sensitive table Lindenmayer languages and a relation to the LBA problem. Information and Control. 33: 217-226.

[10] Rozenberg G. and Salomaa A. 1980. Mathematical Theory of L Systems. New York, NY: Academic Press Inc.

[11] Holzer M., Truthe B. and Yosman A. F. 2018. On bonded sequential and parallel insertion systems. RAIRO - Theoretical Informatics and Applications. 52: 127-151.

[12] Yosman A. F., Holzer M., Truthe B., Fong W. H. and Turaev S. 2017. On bonded Indian and uniformly parallel insertion systems and their generative power. Malaysian Journal of Fundamental and Applied Sciences. 13: 769-773.

[13] Yosman A. F., Fong W. H. and Mat Hassim H. I. 2022. Bonded parallel insertion-deletion systems and their generative power. AIP Conference Proceedings. In press.

[14] Fong W. H., Holzer M., Truthe B., Turaev S. and Yosman A. F. 2016. On bonded sequential and parallel insertion systems. 8th Workshop on Non-Classical Models of Automata and Applications (NCMA); 2016; Debrecen, Hungary. Austria: OCG.

[15] Salomaa A. 1973. Formal Languages. New York, NY: Academic Press Inc.

[16] Rozenberg G. and Salomaa A. 1997. Handbook of Formal Languages Vol 1. Berlin: Springer-Verlag.

[17] Paun G., Rozenberg G. and Salomaa A. 1998. DNA Computing: New Computing Paradigms. Berlin: Springer-Verlag.

[18] Hopcroft J. E., Motwani R. and Ullman J. D. 2007. Introduction to Automata Theory, Languages, and Computation. 3rd Edition. Boston, MA: