

A Review of Aryabhata's Symbolic System and its Analogy with AI Encoding

Tabsum B

Assistant Professor, Department of Mathematics, ¹Government college of Arts, Science and Commerce Quepem Goa, India

Abstract: This Research note examines Aryabhata's Varga–Avarga system as an early form of symbolic encoding, revealing structural parallels with modern principles of artificial intelligence, particularly in representation and reasoning. It highlights how Aryabhata's integration of sound, number, and rule reflects a cognitive framework that anticipates core aspects of linguistic computation and knowledge modeling in AI. By analyzing the manual methods employed by Aryabhata, the study demonstrates that his mathematical and linguistic codifications functioned similarly to algorithmic coding in the modern digital era. The findings suggest that Aryabhata's work embodies a foundational logic of computation long before the advent of computers, offering valuable insights into the continuity between ancient mathematical thought and contemporary AI systems.

Keywords: Varga, Avarga, syllable, encoding, tokens, virama, vowels

1. Introduction

Aryabhata began his work to compress large astronomical numbers (158223500) into pronounceable syllables (beat of sound). These syllables preserve both meaning (numerical value) and structure of the numbers. One of the most fascinating concepts in Aryabhata's work is how he connected language and number. To begin with my study in this topic I mainly focused on [1] and [2]. The concept the Varga and Avarga is recorded in Chapter II Sections 1, 2 and 3 in 59 places in [1]. That shows the importance of the work done by him. The detailed work done by him motivated me to learn his working and some numbers by following his method. The work was tedious but was worth learning. While working on it, I realized it was nothing but coding (in modern world). Today coding is programmed but, in those days, there were no computers. Aryabhata manually could do this work to successfully represent a large number into a pronounceable small word. This paper shows the analogy between his work and coding done in AI.

1) Varga (Grouped)

ka (क), kha (ख), ga (ग), gha (घ), ṅa (ङ) (1, 2, 3, 4, 5 respectively)
ca (च), cha (छ), ja (ज), jha (झ), ña (ञ) (6, 7, 8, 9, 10 respectively)
ṭa (ट), ṭha (ठ), ḍa (ड), ḍha (ढ), ṇa (ण) (11, 12, 13, 14, 15 respectively)
ta (त), tha (थ), da (द), dha (ध), na (न) (16, 17, 18, 19, 20 respectively)
pa (प), pha (फ), ba (ब), bha (भ), ma (म) (21 to 25 respectively)
so 1 – 25 (five per group, from 1 = ka to 25 = ma respectively)

2) Avarga (Ungrouped)

ya (य), ra (र), la (ल), va (व), śa (श), ṣa (ष), sa (स), ha (ह) (30, 40, 50, 60, 70, 80, 90, 100 respectively)

From Aryabhata's literature, it is clear that the letters *ka* to *ma* have the values of 1 – 25. The letters *ya* to *ha* would have the values of 3 – 10, but since a short *a* is regarded as inherent in a consonant when no other vowel sign is attached and when the virama (pause) is not used, and since short *a* refers the Avarga letters to the place of tens, the signs *ya, ra, la, va* etc. really have the values of 30 – 100. The vowels themselves have no numerical values. They merely serve to refer the consonants (which do have numerical

2. Content of the Research Note

What is meant by “Varga” and “Avarga”?

In Sanskrit, Varga (वर्ग) literally means grouped, classified, or square and Avarga (अवर्ग) means unclassified or non-grouped. Aryabhata used these terms to classify consonants in Sanskrit phonetics and assign them numerical values. This forms the basis of his alphabet–numeral system, a compact code used to write large numbers inside verses of poetry (called Giti stanzas).

The Sanskrit Alphabet as Numbers

Varga consonants (*ka* → *ma*) are grouped into fives — hence “classified.”

Avarga consonants (*ya* → *ha*) are “unclassified” because they fall outside the main phonetic groups and denote higher multiples (tens, hundreds).

Let's elaborate

Aryabhata divided the 25 consonants from *ka* (क) to *ma* (म) into two groups:

values) to certain places. So Avarga letters represent multiples of ten (30 – 100). Aryabhata's system encoded numbers using Sanskrit syllables- consonants (Varga-Avarga) combined with vowels to represent place values.

How did Aryabhata number the vowels?

Vowels in Sanskrit:

a (अ), i (इ), u (उ), ṛ (ऋ), ̣ (ॠ), e (ए), ai (ऐ), o (ओ), au (औ) (recorded in Aryabhata's work in page number 4) are

Volume 14 Issue 10, October 2025

Fully Refereed | Open Access | Double Blind Peer Reviewed Journal

www.ijssr.net

not given numerical values themselves, but they modify the consonants (Vargas and Avargas) and act as place-value markers (like zeros or positional multipliers). Aryabhata used them to show tens, hundreds, thousands, etc.

For example:

$a \rightarrow 1 (10^0)$ Unit place

$\bar{a} \rightarrow 10 (10^1)$ Tens place

$i \rightarrow 100 (10^2)$ Hundreds place

$\bar{i} \rightarrow 1,000 (10^3)$ Thousands place

$u \rightarrow 10,000 (10^4)$ Ten – thousands place

$\bar{u} \rightarrow 100,000 (10^5)$ Lakhs place

$\bar{r} \rightarrow 1,000,000 (10^6)$ Millions place

$\bar{\bar{r}} \rightarrow 10,000,000 (10^7)$ Ten – millions place

$\bar{l} \rightarrow 100,000,000 (10^8)$ Hundred – millions place

$e \rightarrow 1,000,000,000 (10^9)$ Billions place

$ai \rightarrow 10,000,000,000 (10^{10})$ Ten – billions place

$o \rightarrow 100,000,000,000 (10^{11})$ Hundred – billions place

$au \rightarrow 1,000,000,000,000 (10^{12})$ Trillions place

So, a single consonant + vowel can represent very large numbers.

Example

consider gha (घ),

gha (घ), is the 4th varga letter = 4.

With vowel “a” (अ), it stays as 4 (units place).

With vowel “i” (इ), it becomes $4 \times 100 = 400$. (ghi = 400)

With vowel “u” (उ), it becomes $4 \times 10,000 = 40,000$. (ghu = 40000)

So by changing the vowel, Aryabhata could express numbers in compact syllables (beat of sound).

It's a phonetic number system where sound = symbol = number. This is what makes it so relevant to symbolic computation and AI — it encodes meaning by systematic rules.

Application in Aryabhata's

Aryabhata used this system to write huge astronomical numbers poetically.

Instead of writing “1,582,237,500”, he could encode it as a short, metrical Sanskrit verse using the varga–avarga letter rules.

Let's see the rules

While following what Aryabhata did in his literature I found something very fascinating.

He only considered the following vowels

a(अ), i(इ), u(उ), r(ऋ), l(ॠ), e(ए), ai(ऐ), o(ओ), au(औ) that means

$1, 10^2, 10^4, 10^6, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}$ respectively

in that case $200000 = 2 \times 10^5 = 20 \times 10^4 = na u = nu$

but not $200000 = 2 \times 10^5 = kha \bar{u} = kh\bar{u}$ (example)

Examples

$$\begin{aligned} 15,822,375 &= 1 \times 10^7 + 5 \times 10^6 + 8 \times 10^5 + \\ &2 \times 10^4 + 2 \times 10^3 + 3 \times 10^2 + 7 \times 10^1 + 5 \\ &= 10 \times 10^6 + 5 \times 10^6 + 80 \times 10^4 + 2 \times 10^4 \\ &\quad + 20 \times 10^2 + 3 \times 10^2 + 70 + 5 \\ &= \bar{n}a \bar{r} + \bar{n}a \bar{r} + \bar{s}a u + kha u + na u + ga i + \bar{s}a \\ &\quad + \bar{n}a \end{aligned}$$

$$= \bar{n}ra + \bar{n}ra + \bar{s}u + khu + nu + gi + \bar{s}a + \bar{n}a$$

Reading it from right to left we get

$$\bar{n}a - \bar{s}a - gi - nu - khu - \bar{s}u - \bar{n}ra - \bar{n}ra$$

$$\begin{aligned} \bullet \quad 158,223,750 &= 1 \times 10^8 + 5 \times 10^7 + 8 \times 10^6 + \\ &2 \times 10^5 + 2 \times 10^4 + 3 \times 10^3 \\ &\quad + 7 \times 10^2 + 5 \times 10^1 \\ &= 1 \times 10^8 + 50 \times 10^6 + 8 \times 10^6 + 20 \times 10^4 \\ &\quad + 2 \times 10^4 + 30 \times \\ &\quad 10^2 + 7 \times 10^2 + 50 \\ &= ka \bar{l} + la \bar{r} + ja \bar{r} + na u + kha u + ya i + cha i \\ &\quad + la \end{aligned}$$

$$= k\bar{l}a + l\bar{r}a + j\bar{r}a + nu + khu + yi + chi + la$$

Reading it from right to left we get

$$la - chi - yi - khu - nu - j\bar{r}a - l\bar{r}a - k\bar{l}a$$

$$\begin{aligned} \bullet \quad 1,582,237,500 &= 1 \times 10^9 + 5 \times 10^8 + 8 \times 10^7 + \\ &2 \times 10^6 + 2 \times 10^5 + 3 \times 10^4 \\ &\quad + 7 \times 10^3 + 5 \times 10^2 \\ &= 1 \times 10^9 + 5 \times 10^8 + 80 \times 10^6 + 2 \times 10^6 \\ &\quad + 20 \times 10^4 + 3 \times 10^4 + 70 \times 10^2 \\ &\quad + 5 \times 10^2 \\ &= ka e + \bar{n}a \bar{l} + \bar{s}a \bar{r} + kha \bar{r} + na u + ga u + \bar{s}a i \\ &\quad + \bar{n}a i \end{aligned}$$

$$= ke + \bar{n}la + \bar{s}ra + kh\bar{r}a + nu + gu + \bar{s}i + \bar{n}i$$

Reading it from right to left we get

$$\bar{n}i - \bar{s}i - gu - nu - kh\bar{r}a - \bar{s}ra - \bar{n}la - ke$$

Modern Relevance with AI

In modern AI terms, Consonants = symbolic tokens (like feature codes). Vowels = positional encoders (like context markers). Rules of combination = algorithmic syntax. Thus, the Varga–Avarga system is an early symbolic data model, where language directly encodes structured quantitative information — a 5th-century precursor to computational linguistics and symbolic AI. Thus, there is a deep connection between Aryabhata's Varga–Avarga sound–number system and certain principles in Artificial Intelligence, especially in symbolic representation, encoding, and language processing. Aryabhata created a symbolic information system — where sound represents numbers according to logical rules. That is essentially what AI systems do today: They represent information (meaning) through symbols (codes). They manipulate those symbols using rules (algorithms). Aryabhata's system aimed at achieving compact representation of large numbers into few syllables. This is exactly the aim of AI, to reduce dimensions while preserving meaning.

Symbolic Encoding and analogy with AI

Aryabhata's Method	Modern AI
- Letters (Varga–Avarga) encode numerical values systematically. - Rules define how vowels modify consonants ($\times 100^n$). - Each syllable = compact representation of large data.	- Tokens or symbols encode data in rule-based AI (e.g., symbolic reasoning, expert systems). - Rules define how symbols combine in logic or programming (e.g., inference rules, syntax). - Each token = compact representation of information (in NLP models, embeddings).

3. Conclusion

Aryabhata's system is a symbolic encoding model — a manually constructed language–number mapping, similar to what AI uses in knowledge representation. He used Sanskrit phonology (structured sound system) as a computational medium. Modern AI does the same: it treats language as data, processed logically or statistically. Aryabhata's encoding is purely rule-based, not empirical which parallels classical (symbolic) AI, the pre-neural stage of AI (before deep learning). His Varga–Avarga scheme anticipates the symbolic encoding logic used in early AI systems — rule-based reasoning, symbol manipulation, and hierarchical data representation. This gives it direct conceptual relevance to symbolic AI (e.g., expert systems, logic programming). In Indian philosophy: Shabda = Sound or linguistic form, Samkhya = number or measurable quantity, Buddhi = cognition, reasoning power. Aryabhata's unifies all three: Sound (Shabda) encodes Number (Samkhya) through Rule (Buddhi). This rule of Three is identical to AI cognition: Input (language) → Representation (symbol) → Processing (reasoning). Aryabhata's system symbolically integrates the three cognitive elements of artificial intelligence — language, logic, and computation — through the Sanskritic synthesis of Shabda, Samkhya, and buddhi. Aryabhata's linguistic numerical coding anticipates key principles of artificial intelligence such as symbolic representation, efficient data encoding and algorithmic reasoning. His system represents an early attempt at cognitive computation where sound, number and rule interact much like modern AI models integrate language and logic.

References

- [1] Aryabhata. The Aryabhatiya of Aryabhata: An Ancient Indian Work on Mathematics and Astronomy. Translated with notes by Walter Eugene Clark. University of Chicago Press, 1930.
- [2] Aryabhata; K. V. Sarma & K. S. Shukla (eds.). Aryabhatiya of Aryabhata. Indian National Science Academy, New Delhi, 1976 (3 volumes).
- [3] Online Sanskrit text: Aryabhatiya of Aryabhata. Wiki source/online edition