Inheriting the glory of PāṆinian grammar into computer science

Ashwini S N
Research Scholar
School of Computer Science and Information Technology
JAIN(Deemed to be University)
Bangalore, India
ashwinisn.rao@gmail.com

Dr M N Nachappa
Professor,
School of Computer Science and Information Technology
JAIN(Deemed to be University),
Bangalore, India
mn.nachappa@jainuniversity.ac.in

Dr Geetha Madhusudan
Professor,
School of Computer Science and Information Technology,
JAIN(Deemed to be University),
Bangalore, India
geetha.madhusudan@jainuniversity.ac.in

Abstract— the integration of Pāṇinian grammar principles into computer science represents a significant advancement in computational linguistics and natural language processing. Pāṇini, an ancient Indian grammarian, formulated an elaborate system of grammar known as *Aṣṭādhyāyī* around the 4th century BCE, which has been revered for its systematic and concise rules governing Sanskrit language structure. In recent years, there has been a growing interest in adapting Pāṇinian principles to enhance computational frameworks for understanding human languages. This paper explores the theoretical foundations of Pāṇinian grammar and its applicability in the context of modern computational linguistics. The integration of Paninian grammar into computer science enhances the accuracy and efficiency of natural language processing system and contributes to preservation and dissemination of ancient linguistic knowledge. Through this the researchers aim to bridge gap between traditional linguistic scholarship and modern computational methodologies. By embracing the systematic and rigorous framework laid out by Pāṇini millennia ago, researchers are poised to achieve new breakthroughs in understanding and processing human language through innovative computational techniques. The journey from ancient Sanskrit grammar to innovative computer science exemplifies a harmonious blend of tradition and innovation in the quest for linguistic knowledge and technological advancement.

Keywords - *Pāṇini, Aṣṭādhyāyī, Computational linguistics, Machine translation, semantic analysis, syntactic rules.*

I.   INTRODUCTION

In Sanskrit the term Vyākaraṇa is defined as *vyakriyate anena iti vyakarana*: Grammar is that which enables us to form and examine words and sentences. Vyākaraṇa, from earliest times was considered to help safeguard correct transmission of the scriptural knowledge. Once we understand the foundational rules of Sanskrit grammar, we can effectively generate wide range of expressions. Hence Sanskrit Language has a greater stock of words than any other language ever known to the philological world[1].

The origin of Sanskrit Grammar cannot be pinpointed. Though Maharishi Pāṇini is recognized as the earliest known grammarian, it is evident that he was preceded by a long line of distinguished grammarians. And Pāṇini refers to several grammarians previous to his time. But very little is known about those ancient masters. But regardless of the origin, "*Aṣṭādhyāyī*" is the most significant, most authoritative and the only available source of Sanskrit grammar for methodical text analysis of Sanskrit vocabulary today.

G G Joseph, specialist in the history of mathematics, in his book "The crest of the Peacock" mentions that Sanskrit's potential for scientific use was greatly enhanced as a result of the thorough systemization of its grammar by Pāṇini. Brevity of form and comprehensiveness of content are the two unparalleled features of Pāṇini's work.

When we refer to Theory of formal language, for computer scientists, the "Formal" refers to the fact that the rules of the language are stated with absolute clarity without any ambiguity. This fact was recognized too early by Sanskrit Grammarians and thus has contributed for the mystique of the language [2] . Pāṇini should be thought of as the forerunner of the modern formal language theory used to specify computer languages. In a Letter to the Editor of CACM, American computer scientist and mathematician Donald Ervin Knuth highlights an observation he made. Pāṇini's most significant work was the compilation of the grammar of Sanskrit. In order to describe the rules of grammar, he invented a notation which is equivalent in its power to that of Backus.

Hence the objective of this paper is to provide brief fundamentals into *Aṣṭādhyāyī*, its comprehensive and concise nature, and to draw parallels between modern computer applications and *Aṣṭādhyāyī* thereby highlighting Pāṇini's influence in the realm of computing even after around 2000 years of its composition

II.   PĀṆINI GRAMMAR

A.   Overview of Pāṇini 's work

The author of the magnificent text "*Aṣṭādhyāyī*" is Maharshi Pāṇini. Maharshi Pāṇini's timeline is assumed to be between 4th and 5th century BCE. Very little is known about his life beyond this or the inadequate details of Pāṇini's life are completely overshadowed by the importance of his work "*Aṣṭādhyāyī*". It is a system of Rules (Sutrani) which regulates the right form of Sanskrit. Hence, Patanjali calls it Siṣṭa-jñānārthāAṣṭādhyāyī. (M. Bh. VI. 3.109) [3].

---

[1] Dr. Arindam Mondal- "The development of Sanskrit grammar on pre-pāṇinian stage" ISSN: 2394-7519 IJSR 2020; 6(6): 154-158

[2] T.R.N. Rao - The Pāṇini-Backus Form in Syntax of Formal Languages

[3] Patanjali Maharshi has written commentary on Pāṇini's *Aṣṭādhyāyī* called Mahabhashya (M. Bh) - Sreenivasarao, "Who was Dhirgatamas?" [Online] Available: Who was Dhirgatamas …? – Part One | sreenivasarao's blogs (sreenivasaraos.com)

B. Structure of Pāṇini Grammar

As the name implies, Aṣṭādhyāyī comprises of Ashta (Eight) Adhyaya (Chapters). Each chapter is divided into four Paadaa (quarters) summing to 32 paadaas. Each paada consists of grammar rules in the *Sutra* format. Sutras are brief and compact statements, often consisting of just a few words or syllables but capture essential truths or principles in a concise and memorizable form. The *Aṣṭādhyāyī is* organized into a series of around 4000 such sutras formulated in a highly condensed and systematic manner often using Meta linguistic symbols and abbreviations. Economy (Lāghava) of expression is Pāṇini's primary concern and he achieves it using technical terms like Pratyāhāras in place of lengthy expressions that saves words and even letters in a rule. Logical organization, generality and other aspects of explanation largely follow from brevity and that brevity can be regarded as the substantial aspect of the explanatory power of the theory behind the grammar[4]

Pāṇini revolutionized the arrangements of Sanskrit alphabets (Varṇa) differently from the traditional arrangements. He *changed the pre existing order of arrangement of Sanskrit alphabets to his own set so that these Pratyāhāras can be constructed seamlessly.* He arranged the letters as 14 sets called the Māheshwara Sutra. These 14 sets form the basis for Pāṇinian Sutras.

TABLE 1. Māheshwara Sutra

1.	अ इ उ **ण्**	8.	झ भ **ञ्**
2.	ऋ ऌ **क्**	9.	घ ढ ध **ष्**
3.	ए ओ **ङ्**	10.	ज ब ग ड द **श्**
4.	ऐ औ **च्**	11.	ख फ छ ठ थ च ट त **व्**
5.	ह य व र **ट्**	12.	क प **य्**
6.	ल **ण्**	13.	श ष स **र्**
7.	ञ म ङ ण न **म्**	14.	ह **ल्**

TABLE 2. Maheshwara sutra representation

1. a i u ṇ		8. jha bha **ñ**	
2. ri li **k**		9. gha ḍha dha **ṣ**	
3.	e o **ng**	10. Ja ba ga ḍa da ś	
4. ai au **c**	11. Kha pha cha ṭha tha ca ṭa ta **v**		
5. ha ya va ra **T**		12. ka pa **y**	
6. la **Ṇ**		13. Śa ṣa sa **r**	
7. ṅa ma nga Ṇa na **m**		14. ha **l**	

The objective of this arrangement is to concisely refer to a group. "Pratyāhāras" are nearest to abbreviation in English. It works this way. The last letter in each group is called the **ith** (इत्) Varṇa. (Ones marked in bold in the tables). They are not to be considered as the part of the rule but only mark the end of each rule. To utilize the list in a very efficient way, Pāṇini forms Pratyāhāras by using any of the alphabets (Other than ith varṇa) and combining it with the ith letter.

Example Ac (अच्) Pratyāhārā consists of a, i, u, ri, li, e, o, ai, au (अ इ उ ऋ ऌ ए ओ ऐ औ), nine letters and repeated use of these nine letters will make subject bulky and complex in the absence of

Pratyāhārā. However, Pāṇini smartly expresses these by uttering ac (अच्) [5]

(इकोयणचि) Ikoyaṇachi is a Sutra containing all Pratyāhārā which states ik (i u ri li) (इक् - इ उ ऋ ऌ) are replaced with yaṇ (ya va ra la) (यण् - य व र ल) when ac (a, i, u, ri, li, e,o,ai, au) (अच् - अ इ उ ऋ ऌ ए ओ ऐ औ) vowel follows.

Pāṇini's rules though enumerated in a definite order are classified into segments and Chapters, according to the topics and their functions. Pāṇinian Sutras are classified into different categories based on what function they fall into. On this basis they can be classified as

1. Sangnya Sutras (The defining sutras)
2. Adhikara sutras (Governing Sutras)
3. Paribhasha Sutras (Interpreting Sutras)
4. Vidhi Sutras (Operational Sutras)
5. Niyama Sutras (Restrictive Sutras)
6. Atidesha Sutras (Extension Sutras)

Another striking feature of Pāṇini sutra is the use of Adyāhāra – the omission of verbs in sutras. The finite verb asti (is) or bhavati (becomes) is taken to be present in each rule[6]. The omission of verbs *in* Pāṇini's grammar serves to enhance both the economy of expression and the clarity of communication. By omitting unnecessary elements that can be inferred from context, Sanskrit sentences can convey meaning efficiently.

C. Comparison with Western Grammar

The framework of Pāṇini grammar and western grammar differ significantly in their origins, purposes, methodologies and theoretical structures. Pāṇini's work came out of century's long effort of preserving the language from altered versions. Its main objective is to distinguish between साधु **शब्द** *(sādhu śabda)* and असाधु **शब्द** *(asādhu śabda)* and generate grammatically valid Sanskrit expressions or the correct usage of words [20].The *Aṣṭādhyāyī* is a sophisticated grammar tradition and can be considered as a devise to freeze language change and eclipsed all previous works. If not the first, it is the oldest such text surviving in its entirety. It emerges from its ancient Indian roots and is distinguished by formalized rules within Sanskrit. In contrast, Western grammar has evolved through diverse historical and theoretical developments influenced by Greek, Latin and other vernacular languages.

In Sanskrit each word's structure and meaning can typically be deciphered through its components (root, suffix, prefix, etc.). Hence Sanskrit's systematic grammar allows for precise understanding of words based on their constituent parts, thereby reducing ambiguity and enabling accurate interpretation. In Sanskrit grammar, the addition of prefixes (upasargās) to verb roots (dhātus) follows a highly systematic and rule-bound process. In contrast, while prefixes and particles exist in Western languages, their impact on verb meaning is often less rigidly defined and more context dependent.

Sutras: Pāṇini's Aṣṭādhyāyī consists of concise aphorisms (sutras) that systematically describe phonetics, morphology, and syntax. This makes the grammar highly formalized and focuses on precise rules governing word formation, sentence structure and meaning. Also, to help remember it, Pāṇini is said to have preferred brevity

---

[4] *Smith H (1992). "Brevity in Pāṇini". Journal of Indian Philosophy*, *20*(1), 133–147. http://www.jstor.org/stable/23445598

[5] Mohini Arya – "Structure of Paninian grammar (Ashtadhyayi)" International Multidisciplinary Research Journal - ISSN 2424-7073
[6] Bhate, Saroja, and SubhashKak. "PĀṆINI'S grammar and computer science." *Annals of the Bhandarkar Oriental Research Institute*, vol. 72/73, no. 1/4, 1991, pp. 79–94. *JSTOR*

over clarity. The sutra style and the brevity in representation remain unique to Sanskrit.

Western grammatical traditions are also systematic but developed within different linguistic contexts. It focuses on descriptive and analytical methods than the sutra-based approach of Pāṇini.

Pāṇini's work however has a great influence on modern linguistics and many of the foundational ideas proposed by Ferdinand de Saussure, professor of Sanskrit, who is widely considered the father of modern structural linguistics. The founding father of American structuralism, Leonard Bloomfield, wrote a 1927 paper titled "On some rules of Pāṇini" and described Pāṇini's grammar as 'one of the greatest monuments of human intelligence'.

III. APPLICATIONS IN COMPUTER SCIENCE

A. Linguistic Theories and Computational Linguistics

Through his Magnum Opus, The Aṣṭādhyāyī, Pāṇini has made a profound contribution to the field of linguistics and is continuing to influence linguistic theories even today. He has set benchmarks as to how languages could be methodically described and analyzed through his systematic and rule-based approach. Pāṇini provided structured framework for understanding language structure and usage. He has covered all aspects of language structure including phonetics, morphology (word formation) and syntax (Sentence structure). He provides detailed rules for forming words from roots (dhātu), inflecting them with case endings (Vibhakti) and constructing sentences by combining the affixes (Pratyaya).

Leonard Bloomfield an American linguist who led the development of structural linguistics in the United States during the 1930s and the 1940s has made significant contributions to Indo-European historical linguistics. Bloomfield's synchronic grammatical works were heavily influenced by Pāṇini. Word formation, compounds, form-classes, and generality and specificity in Bloomfield's work are correlated with their counterparts in Pāṇini's grammar. Selections from a manuscript of Bloomfield's translation and annotation of the Kāśikā (Kāśikā-vṛtti, commonly known as Kāśikā, is a renowned commentary on Panini's Aṣṭādhyāyī,) provide concrete evidence for the influence of Pāṇini on Bloomfield [7].

The influence of Bloomfield on structural linguistics declined in the late 1950s and 1960s as the theory of Generative Grammar developed by Noam Chomsky came to predominate[8]. Chomsky is cardinal for a new way in linguistics changing the focus from traditional grammar design to transformational-generative grammar design. He admits that great grammarians of the past shared and contributed to his inspiration. He is still revered in those circles today. All the "great" spirits that he cites as his pioneers all had one thing in common: they studied Sanskrit and admired Pāṇini the Hindu grammarian. And this is where this article originated[9].

Computational Linguistics is a field of study that combines principles and methods from linguistics and computer Science to understand and process natural language using computational techniques. It is an emerging field in the 20th Century. Despite the vast gap from Pāṇinian period till now, there has been profound influence of Pāṇinian Grammar on computational linguistics. As we understand that Pāṇinian Grammar is the earliest attempt to formalize grammar in a systematic and rule based manner. Such formalization became the foundation to computational linguistics to demonstrate how language structure can be described using precise rules analogous to algorithms in computational models. Developing algorithms and models helps in better understanding of natural Language. Hence Pāṇinian grammar provides foundational principles in the development of NLP.

B. Natural Language Processing (NLP)

Pāṇini's sutras or the system of rules systematically explain how linguistic elements should be combined and how they get transformed in different contexts. This approach can be applied to parsing in NLP where algorithms are used to analyze the syntactic structure of the sentences. NLP parsers parse sentences by applying rules sequentially to breakdown the sentences into grammatical forms like nouns, verbs, preposition etc. based on syntactic categories.

This is inspired by Pāṇinian approach of morphological analyses. Every word formation can be analyzed by identifying the root form and applying rules for inflection changes thus generating grammatically precise words. Such precision is utilized in NLP for tasks like machine translations where algorithms generate coherent translations based on grammatical rules. Also analyzing word formation from root and inflectional changes helps in transformational grammar of NLP which helps in generating syntactic transformations like passive voice or question formation.

C. Formal Language Theory

Formal language theory deals with examining the characteristics of the language like structure, generative rules, and classification. This theoretical framework provided by the language is used in automata to provide computational models to recognize and generate these languages. The objective of the formal language theory is to define and study formal languages through grammar where production rules define how words or sentences can be generated. This is in line with Pāṇinian grammar's systematic and rule-based nature that lay the groundwork for computational models in formal language theory. The hierarchy in classification of rules in Pāṇini grammar is central to formal language theory which classifies languages into Regular, Context free and Context sensitive. Pushdown automata recognize context free grammar which are more expressive than regular grammar and governed by context free rules. This parallels Pāṇini grammar's use of recursive rules (anuvritti) and hierarchical structure for generation and analysis of sentences.

IV. SUCCESS STORIES

Pāṇini's grammar is studied in the realm of linguistics and Sanskrit studies and has profound influence on theoretical framework of computer science. Direct practical application of Pāṇini's grammar is more prominently seen in its conceptual influence on fields such as formal language theory, computational linguistics and AI. Besides natural language processing (NLP), Pāṇini's grammar and its principles have influenced or have potential applications in several other areas like

- **Programming Languages** - The concept of formal grammars, influenced by Pāṇini's work, is fundamental in computer science and programming language design. Semantics in programming is about whether or not the instruction has a valid meaning. Ambiguity in syntax often leads to ambiguity in meaning. By defining clear rules for sentence construction and word usage, Pāṇini's grammar helps in minimizing ambiguity and promoting semantic clarity. This systematic nature of Pāṇinian grammar promotes linguistic precision in expression hence users can convey their intended meaning accurately.

---

[7] David E. Rogers - The influence of Panini on Leonard Bloomfield - historiographia linguistica, Volume 14, Issue 1-2, Jan 1987, p. 89 - 138

[8] Yale Linguistics – Leonard Bloomfield [Online] Available: Leonard Bloomfield | Yale Linguistics

[9] Koot van Wyk "Evidence of Hindu religion on the theory of Chomsky's transformational grammar" International Journal of Language and Linguistics Vol. 2, No. 5; November 2015

- **Compiler Design**: Software that translates high level language to machine instructions relies on formal grammar and parsing techniques. In Pāṇinian grammar, parsing can be achieved with remarkable efficiency due to its systematic and structured approach. Each rule addresses a specific linguistic element like Prātipadika (Noun), Kriya pada (Verb), vibhakti (Declensions), samāsa (Compounds) etc. This ensures that each segment of the text is analyzed according to specific guidelines. Also, as soon as the pattern matches a rule, it can be processed immediately without backtracking which enhances the parsing efficiency and minimizes redundant checks.

V. FUTURE DIRECTIONS

A. Current Research and Challenges

While Pāṇinian grammar was developed for Sanskrit, researchers are exploring the possibility of its application to modern languages. The challenges include integrating Pāṇinian principles to languages with different syntactic structures and morphological complexities. Current NLP systems use formal representations (like syntactic structures, dependency trees, etc.) to encode and process language. Pāṇini's grammatical theories have implications for linguistic theories and cognitive science, providing insights into how language structures are processed and understood by humans and machines alike. The design of rule-based systems in artificial intelligence and expert systems, where rules govern reasoning and decision-making processes is an example of Pāṇinian grammar's influence on current AI systems.

B. Potential Applications

Pāṇinian grammar provides a structured framework for describing syntax and morphology of the language. This aligns with AI systems that rely on rules and algorithms for processing and understanding the language. This integrated system can be applied to tasks like Parsing and grammar checking in NLP. It can produce valid sentences which is a foundational concept in AI to understand the structure of sentences in Natural language. AI researchers are exploring the possibility to adapt Pāṇinian style grammar to modern languages enhancing the accuracy and efficiency of AI systems in multilingual contexts.

Pāṇini's systematic approach to analyzing grammar shares fundamental principles with modern formal language theory. Formal language theory has applications in various fields, including

- Compiler Design - Formal grammars are used to define the syntax of programming languages, and automata are used in lexical and syntactic analysis
- Natural Language Processing (NLP) - Formal language theory principles are applied in tasks such as parsing, syntactic analysis, and language generation.
- Cryptography

Hence advancements in the realm of formal language theory under the influence of Pāṇinian grammar can benefit advancements in these fields as well

VI. CONCLUSION

The system devised by Pāṇini is, therefore, looked upon as a Great Science and, is always at the centre of vast and varied traditions of Sanskrit Grammar. Pāṇini's grammar is distinguished above all similar works of other countries, partly by its thoroughly exhaustive investigation of the roots of the language and the formation of words; partly by its sharp precision of expression, which indicates with brevity whether forms come under the same or different rules , Pāṇini's grammar stands out for its systematic formalization, generative approach, specificity to Sanskrit, concise aphoristic style, meta-linguistic analysis, and computational aspects, setting it apart from other grammatical traditions of its time and influencing later developments in linguistics and language theory.

It exhibits several characteristics that can be seen as forerunner to computational concepts and structures. His work remains foundational to understanding language structure and continues to shape research and development in computational approaches to language processing and understanding.

Walter Eugene Clark, who was Professor of Sanskrit at Harvard University, and who translated Aryabhatta's Aryabhatiyam into English, wrote that "Panini's grammar is the earliest scientific grammar in the world, and one of the greatest." He said the "Indian study of language was as objective as the dissection of the body by an anatomist."
The principles and methodologies derived from Pāṇini's grammar continue to inspire research in computational linguistics, historical linguistics, and the development of language technologies.

Pāṇini's influence on computational linguistics lies in his pioneering efforts to formalize grammar, develop generative grammar theories, and establish a systematic approach to linguistic analysis. His work remains foundational to understanding language structure and continues to shape research and development in computational approaches to language processing and understanding.

Pāṇini's contributions to linguistic theories are foundational and enduring. His systematic approach to grammar, formalization of linguistic analysis, and emphasis on structural principles has shaped the study of language across different linguistic traditions and continue to inspire research and theory development in linguistics today. By building on Pāṇini's foundational insights, AI researchers can continue to push the boundaries of language understanding and generation in computational linguistics. As technology evolves, leveraging Pāṇini's structured, and rule-based approach promises to enhance our understanding and utilization of language in various computational contexts.

Pāṇinian grammar has influenced subsequent development of linguistic theory, formal grammar and computational linguistics. His magnum opus, "the greatest monument of human intelligence" the Aṣṭādhyāyī, remains a foundational text in the history of linguistic and formal language theory.

REFERENCES

1. Dr. Arindam Mondal- "The development of Sanskrit grammar on pre-Paninian stage" ISSN: 2394-7519 IJSR 2020; 6(6): 154-158
2. T.R.N. Rao - The Pāṇini-Backus Form in Syntax of Formal Languages Subhash C Kak - "The Pāṇinian Approach to Natural Language Processing"
3. Sreenivasarao, "Who was Dhirgatamas?" [Online] Available: Who was Dhirgatamas …? – Part One | Sreenivasa Rao's blogs (sreenivasaraos.com)

4. *Smith H (1992). "Brevity in Pāṇini". Journal of Indian Philosophy*, 20(1), 133–147. http://www.jstor.org/stable/23445598

5. Mohini Arya – "Structure of Paninian grammar (Ashtadhyayi)" International Multidisciplinary Research Journal - ISSN 2424-7073

6. Bhate, Saroja, and Subhash Kak. "PĀṆINI'S grammar and computer science." *Annals of the Bhandarkar Oriental Research Institute*, vol. 72/73, no. 1/4, 1991, pp. 79–94. *JSTOR*

7. David E. Rogers - The influence of Panini on Leonard Bloomfield - historiographia linguistica, Volume 14, Issue 1-2, Jan 1987, p. 89 - 138

8. Yale Linguistics – Leonard Bloomfield [Online] Available: Leonard Bloomfield | Yale Linguistics

9. Koot van Wyk "Evidence of Hindu religion on the theory of Chomsky's transformational grammar" International Journal of Language and Linguistics Vol. 2, No. 5; November 2015