सङ्गणककृत्रिमबुद्धिप्रविधिषु पाणिनितत्त्वानुप्रयोगः

Om Prakash Das
Ph.D. Research Scholar
Department of Vyakaran (P.G.)
Shree Jagannath Sanskrit University
Shree Vihar, Puri, Odisha, India
Email: omprakashdas20@gmail.com

भाषां सङ्गणकस्थपद्धतिपथे यन्त्राश्रितां साधिकां
विज्ञाने प्रविधौ च कृत्रिममतौ कूटप्रयोगे पुनः।
धृत्वा सर्वनयस्तुतां स भगवान् शब्दार्थतत्त्वप्रदो
बाभ्राष्टीह समस्तविज्ञविनुतो वैज्ञानिकः पाणिनिः।। (स्वरचितम्)

Abstract: यथा लोके विज्ञानविस्तृतेन नित्यं नुतनाः प्रवृत्तयः सम्पद्यन्ते, तथा तेषु कृत्रिमबुद्धिनामा (Artificial Intelligence) प्रविधिविशेषः संगणकतन्त्रसहितोऽद्यतनयुगे विश्वमानवजीवनस्य सर्वाङ्गेऽन्तवर्तित एव। भाषाप्रक्रियायाः (Natural Language Processing) नाम कृत्रिमबुद्धेः प्रमुखशाखायां, वाक्यविश्लेषणं, रूपनिर्माणं, यन्त्रानुवादः, ध्वनिपरिचयः, पाठसङ्ग्रहणम् इत्यादीनां कार्याणां संसाधनाय, एकस्मिन् सुव्यवस्थिते भाषानियमनतन्त्रेऽपेक्षत्वं संवर्तते। संस्कृतभाषा, प्राचीनतमविज्ञानपरम्परायाः धारिणी, तत्र पाणिनिप्रणीतम् अष्टाध्यायी-नामा व्याकरणमहाग्रन्थः, सूत्रलाघव-तार्किकसंरचना-व्यवस्थितसूत्रक्रमादिभिर्विश्वेऽस्मिन्नद्वितीयः। पाणिनीयव्याकरणं लघुत्वसूत्रसंयोजने परिपूर्णम् अस्ति, तथैव सङ्गणकीय-अल्गोरिद्मनिर्माणेऽपि प्रयोज्यमेव। अल्पाक्षरम्-असन्दिग्धम् इति सूत्रकारस्य नीतिरिव न्यूनतम-संकेतैरधिकतमवर्णनस्य सामर्थ्यं सन्दर्शयते, यत् संगणक-प्रोग्रामिंग-भाषासु प्रत्यक्षत्वेनाऽवधार्यते। अस्य शोधपत्रस्य प्रमुखलक्ष्यम्— पाणिनीयतत्त्वानाम् आधुनिकार्थे कृत्रिमबुद्धिप्रविधिषु, विशेषतो भाषा-प्रक्रियायां, अनुप्रयोगस्य मार्गाणां विशद्विमर्शः। अत्र सूत्रप्रणाली, धातुपाठ-गणपाठविन्यासः, कारकसिद्धान्तः, समासविधानम्, सन्धिनियमाश्च, Finite State Automata, Context-Free Grammar, Rule-based Modelling, Machine Learning इत्यादिषु प्रविधिषु कथं यथासम्भवम् अन्वयीभवितव्या इति विश्लेष्यते। अनुसन्धानपरिणामद्वयं विशेषत्वेन निर्दिश्यते— (१) पाणिनीयनियमाधारिताः सङ्गणकीय-अल्गोरिद्मतत्त्वानि बहुभाषाप्रक्रियायामधिकशुद्धतारूपं, कार्यकुशलत्वरूपञ्च प्रददाति। (२) संस्कृतभाषाया वैज्ञानिकसंरचना, सूत्रलाघवप्रणाली च, आधुनिककृत्रिमबुद्धिप्रविधिषु अन्वितत्वेन, भारतीय-ज्ञानपरम्पराया वैश्विकपुनर्मूल्याङ्कनाय मार्गं प्रशस्तीचर्करीति। अतोऽयं शोधः केवलं ऐतिहासिकव्याकरणानुसन्धानपर्यन्तं न स्यात्, अपि तु परम्परा-नवीनप्रवर्तनयोः सङ्गमबिन्दुरूपेण, प्राचीन-व्याकरण-तर्कस्य अद्यतन-सङ्गणकीयविज्ञानसङ्गतेश्च, अभिनवम् अद्वितीयञ्च योगदानं करिष्यति। दीर्घदृष्ट्या, विशेषतो भारतीयभाषासु, पाणिनीयव्याकरणाधारित-कृत्रिमबुद्धिप्रविधय आदर्शस्वरूपेण स्वीकर्तुं शक्या इति निश्चिन्तं परिचीयते।

Keywords: पाणिनीयव्याकरणम्, कृत्रिमबुद्धिः, सङ्गणकतन्त्रज्ञानम्, सङ्गणकं, यन्त्रानुवादः, रूपनिर्माणं, रूपविश्लेषणं, भाषाविज्ञानम्।

पाणिनीयपद्धत्याः प्राकृतिकभाषासंसाधनस्य च विवेचनम् — Impact of paninian grammar on NLP (Natural Language Process)

आधुनिकसङ्गणकीयभाषाविज्ञाने प्राकृतिकभाषासंसाधनस्य क्षेत्रे यानि क्लेशजनकानि समस्या दृश्यन्ते, तासां समाधानार्थं पाणिनीयव्याकरणस्य सुशृङ्खलितपद्धतिरत्यन्तं उपयोगी भवति।

पाणिन्यष्टाध्यायी, यो लियोनार्ड् ब्लूम्फील्ड् इत्यनेन "मानवबुद्धेर्महदस्तम्भः" इति निरूपितः, नियमाधारितपद्धत्या आदर्शोऽस्ति। अस्य पद्धत्या विशिष्टा नियमरचनापद्धतिः, नियमेषु नियमानामपि समावेशः, अनेन अद्यतनभाषासंसाधनपद्धत्या विकासः साध्यते।[1]

पाणिनीयव्याकरणस्याधारभूमिः

पाणिनिः पञ्चमषष्ठशताब्द्योः ईसापूर्वे जीवित आसीत्, तेन रचितः 'अष्टाध्यायी' नामको व्याकरणग्रन्थो भाषाविज्ञानस्य सार्वत्रिकपद्धतिं प्रतिपादयति। पाणिनीयव्याकरणं प्रायः 4000 सूत्रैः संयुक्तम् अस्ति, यत् संस्कृतभाषायाः सम्पूर्णवर्णनञ्च करोति। अस्य व्याकरणस्य वैशिष्ट्यम् अस्ति यद् अत्र नियमेषु नियमरूपेण meta-rules अन्तर्भाविताः सन्ति, विशेषपरिभाषाभाषा (meta-language) च उपयुज्यते। अत्र पुनरावृत्तिनियमैः सीमितनियमैः अनन्तवाक्यरचनायाः सम्भावना दर्शिता, यत् generative grammar इत्यस्य तत्त्वं चोम्स्कीत्यादिभिः प्रतिपादितम्।

पाणिनीयव्यवस्थायां कृतं कारकतत्त्वं विशेषरूपेण गम्भीरार्थदर्शकञ्चाऽस्ति। अत्र षट् करकानि– अपादानं, कर्म, करणं, सम्प्रदानं, अधिकरणं, कर्ता इत्यादीनि प्रदत्तानि, यानि क्रियाया विभिन्नपक्षान् स्पष्टरूपेण व्यक्तीकुर्वन्ति, यद् अद्यतन natural language processing प्रणालीषु semantic role labelling इत्यत्र उपयुज्यते।[2]

भाषायाः स्पष्टार्थभावसमस्या (Ambiguity problem)

Natural Language Processing (NLP) मध्ये मुख्यसमस्या अस्ति भाषायाः ambiguity विभिन्नस्तरेषु– lexical, structural, semantic, pragmatic इत्यादिषु। उदाहरणद्वारा स्पष्टीकुर्वे यथा– "Stay away from the range," इत्यत्र 'range' शब्दस्य विभिन्नार्थाः सम्भवन्ति इति, यः lexical ambiguity इत्यस्य द्योतकोऽस्ति। पाणिनीयव्याकरणस्य exhaustive description तथा स्पष्ट syntactic-semantic mapping इत्यादि कारणेन एताः ambiguity इति समस्या न्यूनतामवाप, यत् NLP प्रणालीषु आदर्शरूपेण ग्राह्यम्।[3]

सङ्गणकसन्दर्भे पाणिनीयव्याकरणस्य प्रासङ्गिकता

आधुनिक NLP पद्धतिषु भाषाप्रक्रिया phonological, morphological, syntactic, semantic, pragmatic analysis इत्येतत् श्रेणीबद्धविधानेन व्यवह्रियते। पाणिनीयपद्धतिरपि अनेन श्रेणीबद्धविश्लेषणेन सह साहचर्यं स्थापयति, किन्तु अत्र meta-rules तथा karaka theory इत्यादिना अधिकप्रणालीबद्धताया दर्शनम् भवति।

पाणिनीयव्याकरणे वर्णनस्य श्रेणी– phonemes इत्यस्माद् आरभ्य पूर्णवाक्यपर्यन्तम्– अङ्गीकृताः सन्ति, यत् आधुनिक NLP पाइपलाइन्सु दृश्यते। अनेन efficient rule-based systems निर्माणे अपि साहाय्यं लभते ।[4]

पाणिनीयपद्धतिः knowledge representation frameworks मध्ये alignment दर्शयति, यथा augmented transition networks (ATNs), lexical functional grammars (LFG) इत्यादीनि, यानि अद्यतन NLP प्रणालीषु उपयुज्यन्ते। अस्यावस्थायां उल्लिखितं यत् semantic networks पाणिनीयवर्गीकरणस्य आधारं गृहित्वा agent-action-object सम्बन्धदर्शकविध्या विश्लेषणं साधयन्ति।

उदाहरणरूपेण– "At night you must not enter the monastery" इत्यस्य वाक्यस्य

पाणिनीयविश्लेषणे स्पष्टतया action, time, agent इत्यादीनि उपादेयानि सन्ति।[5]

NLP मध्ये पाणिनीयप्रणाल्या लाभाः

निर्दिश्यते यद् वर्तमाने AI-आधारिता NLP प्रणाली केवलं सीमितविषयक्षेत्रेषु सफलतां प्राप्नोति, कारणं ambiguity तथा contextual dependencies। पाणिनीयपद्धतिः exhaustive grammatical rules तथा semantic स्पष्टता कारणेन, जटिलभाषासंरचनानां systematic विश्लेषणाय विकल्परूपेण उपयुज्यते।

निरूप्यते यत् पाणिनीयव्याकरणेन structured paraphrasing सम्भवति, यत् AI knowledge representation पद्धत्या सह अनुरूपम् अस्ति। आधुनिक AI मध्ये यत् कार्यं क्रियते तत् पाणिनीयपद्धत्या पूर्वमेव निहितम् अस्ति।[6]

आधुनिकभाषासु पाणिनीयशैलीव्याकरणनिर्माणदिशा

Generative grammar विषये Chomsky इत्यादिनिर्देशोऽस्ति यत् universal grammar तत्त्वानि मानवेन्द्रियेषु नैसर्गिकरूपेण एव सन्ति, यत् पाणिनीयव्याकरणस्य रचनात्मकसिद्धान्तेन (Generative) सह मेलम् दर्शयति। अनेन आधुनिकभाषासु पाणिनीयशैलीव्याकरणस्य निर्माणं कृत्वा NLP मध्ये ambiguity resolution तथा context management सुगमं कर्तुं शक्यते।

पाणिनीयपद्धत्या आधारितः दृष्टिकोणः आधुनिकभाषासंसाधनपद्धतिषु precision तथा systematic प्रक्रिया सुनिश्चितुं समर्थोऽस्ति। अस्य समर्थनं करोति यत् पाणिनीयतत्त्वसमन्वित NLP प्रणालीभिः knowledge representation तथा processing मध्ये स्पष्टता प्राप्तुं शक्यते।[7]

पाणिनेः Ashtadhyayi भाषाविज्ञानक्षेत्रे महत्त्वपूर्णं योगदानम् अदात्, यो जनरेटिव, नियमाधारित, पुनरावृत्तिसम्पन्नपद्धत्याः प्रतिपादनं चकार। एषा पद्धतिराधुनिक NLP मध्ये दृश्यमानैः क्लिष्टत्वैस्तथा लक्ष्यैः सह सुसम्बद्धितम् अस्ति। निर्दिश्यते यत् पाणिनीयसिद्धान्तानां NLP मध्ये समावेशनं systematic, precise तथा semantically clear language processing प्रणालीनिर्माणाय सक्षमम् अस्ति।

भविष्ये NLP तथा computational linguistics मध्ये अनुसन्धानं पाणिनीयप्रणाल्याः सम्भावनाया अन्वेषणाय निर्दिष्टं भवेत्, येन भाषासङ्केतनस्य जटिलत्वस्य स्पष्टत्वेम व्यवस्थापनं सम्भवेत्, तथा knowledge representation मध्ये precision स्थापनं साध्येत।[8]

सङ्गणकविज्ञानक्षेत्रे पाणिनीयव्याकरणम्

पाणिनिना कृतं व्याकरणशास्त्रं न केवलं भाषायाः शुद्धता-नियमकारणं किन्तु गणितीयचिन्तनस्य, सङ्गणकप्रणालीनिर्माणस्य च आदर्शरूपेण संप्रतिष्ठितम्। आधुनिकसङ्गणकविज्ञानम् अपि पाणिनीयव्यवस्थायाः कतिपय मौलिकतत्त्वानि स्वीकरोति।[9]

तस्माद् अयं ग्रन्थभागः पाणिनीयव्याकरणस्य तादृशं वैशिष्ट्यं विस्तरेण प्रकाशयति।

डेटा डेटासमूहसंरक्षणञ्च (Data and Data base)

आधुनिकसङ्गणकविज्ञाने डेटा इत्युक्ते यान्यप्यर्थहीनानि तथ्यानि प्रदत्तानि तानि व्यवस्थापूर्वकं नियोजितानि चेद् ज्ञानरूपेण परिणमन्ति। पाणिनिनाऽपि इदमेव विदधाति यत्र, अक्षरसामाम्नायः (Akṣarasāmāmnāya) – १४ शिवसूत्ररूपेण वर्णानाम् अनुशासनम्।

सूत्रपाठः (Sūtrapāṭha) – ४००० सूत्ररूपेण नियमनिर्देशनम्।

धातुपाठः (Dhātupāṭha) – १९६७ धातुरूपेण क्रियापदानां मूलाधाररचना।

घणपाठः (Ghaṇapāṭha) – अव्यय, प्रातिपदिकनिर्माणस्य साधनम्।

चत्वार इमे डेटाबेसरूपेण पाणिनीयव्यवस्थायाम् उपयुज्यन्ते, यथा आधुनिक डेटाबेस् सङ्गणकप्रणालीषु।[10]

पुनरावृत्तिः (Recursion)

पाणिनिना पुनरावृत्तिपद्धतिः (Recursion) स्पष्टत्वेन प्रतिपादिता। पूर्वसिद्धसूत्रेषु विहितपदानि, परसूत्रेषु व्याप्तरूपेण स्वीकृतानि, पुनः पुनः प्रयुक्तानि च भवन्ति । उदाहरणरूपेण: —

आधुनिकसङ्गणकविज्ञाने पुनरावृत्तेः उपयोगः यथा Factorial Function वा Fibonacci Series इत्यस्य निर्माणे क्रियते, तथैव पाणिनीयपद्धत्यामपि क्रियते ।[11]

आव्यूहः (Array)

शिवसूत्राणि यद् महेश्वरसूत्राणि यथानिर्दिष्टं वर्णानां क्रमबद्धरचनारूपेण उपयुज्यन्ते। संरचनमद आधुनिकसङ्गणके Array वा List इत्यादिकल्पनया साकं तुल्यम्।

प्रत्याहारेषु अन्त्यस्वराणां लोप आधुनिक '\0' Null Character तुल्यः।

वर्णसमूहस्य सङ्गठनं आव्यूहात्मकं यथासङ्ख्यं व्यवस्था अपि दृश्यते।[12]

उत्तराधिकारः (Inheritance)

पाणिनीयनियमा न्यूनस्थानव्यवस्थितत्वं पालयन्ति। पूर्वसूत्रनिर्दिष्टनियमः परसूत्रेऽपि स्वीक्रियत, एवं स्थानव्यवस्थायाः संरक्षणेन लघुत्तमत्वं प्राप्तम्। एषा पद्धतिराधुनिक- OOP (Object Oriented Programming) भाषायाम् Multiple Inheritance कल्पनायास्तुल्या ।[13]

बहुरूपत्वम् (Polymorphism)

एकस्य नाम्नो बहुविधार्थत्वं पाणिनिना प्रदर्शितम्। यथा दक्षिणम् इति शब्दः शिवस्य च, दिशायाश्च बोधकः। इयं पद्धतिः आधुनिकसङ्गणकविज्ञाने run time polymorphism इति नाम्ना प्रसिद्धा। उदाहरणानि यथा:

मृगवाहनः – रुद्रो वा पशुधारी पुरुषः।

पीताम्बरः – विष्णुर्वा पीतवस्त्रधारी पुरुषः।

गजाननः – गणेशो वा गजमुखयुक्तः।[14]

निष्कर्षः

पाणिनिः Backus-Norm Form इत्यस्य स्थाने Panini-Backus Form इत्यपि नामकरणं Kruth (1964) इति महाशयेन समर्थितम्। NASA अपि संस्कृतव्याकरणस्य Morphological Clarity इत्यमुं संवीक्ष्य संगणकभाषारूपेण परीक्षणं कुर्वन्नस्ति। पाणिनीयनियमा उपनियमाश्च आधुनिक NLP (Natural Language Processing) प्रणालीषु उपयोगी। पाणिनीयव्याकरणमेव आधुनिकसङ्गणकविज्ञानस्य, NLP प्रणाल्याः, संगणकभाषायाश्च आदर्शरूपेण संस्थापितम्। पाणिनियुक्ता नियमा यथा Recursion, Array Structures, Inheritance, Polymorphism इत्यादीनि, आधुनिकतन्त्रज्ञानस्य आधारीभूतसिद्धान्तेषु सम्मिलितानि।

Endnotes

1. Kak, S. C. (1987). The Paninian approach to natural language processing. International Journal of Approximate Reasoning, 1, 117–130. Elsevier Science Publishing Co., Inc.
2. Ibid
3. Ibid
4. Ibid
5. Ibid
6. Ibid
7. Ibid
8. Ibid
9. Parul Saxena, Kuldeep Pandey, Vinay Saxena (2011). Panini's Grammar in Computer Science, Recent Research in Science and Technology
10. Ibid
11. Ibid
12. Ibid
13. Ibid
14. Ibid

Bibliography

- वाक्यपदीयम् - प्रो. हरिनारायण तिवारी, CHOWKHAMBA SANSKRIT SERIES OFFICE, 2015
- महाभाष्यम् - डॉ. हरिनारायण तिवारी, CHAUKHAMBHA VIDYA BHAWAN, 2014
- काशिका - श्री वामन और जयादित्य, TARA PRINTING WORKS PUBLISHER, VARANASI, 2016
- अष्टाध्यायी - पदवाक्यप्रमाणज्ञ श्री पं० ब्रह्मदत्त जिज्ञासु, सम्पादक, मन्त्री- रामलाल कपूर ट्रस्ट, राधाप्रेस, कैलाश नगर, दिल्ली, 2010
- वाक्यपदीयम् - प्रो. हरिनारायण तिवारी, CHOWKHAMBA SANSKRIT SERIES OFFICE, 2015
- पाणिनीयशिक्षा - शिवराज आचार्य कौंडिण्न्यायन, CHAUKHAMBHA VIDYA BHAWAN, 2012
- भाषाविज्ञान- भोलानाथ तिवारी, किताब महल प्राइवेट लिमिटेड, इलाहाबाद, 1884.
- भाषा विज्ञान एवं भाषा शास्त्र- कपिलदेव द्विवेदी, विश्वविद्यालय प्रकाशन, वाराणसी, 2021.
- अर्थविज्ञान और व्याकरणदर्शन, डा.कपिलदेवद्विवेदी, विश्वविद्यालयप्रकाशन, वाराणसी, २०२१
- संस्कृतं युरोपीया भाषाश्च- तुलनात्मकमध्ययनम्, डा.भवानीशङ्करत्रिवेदी, आर्यभारती, १९८५
- Greenberg, Joseph H. On Language: Selected Writings of Joseph H. Greenberg. Edited by Keith Denning and Suzanne Kemmer, Stanford University Press, 1990.
- Language, Leonard Bloomfield, George Allen & Unwin Ltd., London, first published in Great Britain 1923; reprinted 1940, 1955, and 1957.
- Selected Writings I: Phonological Studies, Roman Jakobson, Mouton & Co., 1962.

- Aspects of the Theory of Syntax, Noam Chomsky, The MIT Press, 1965.
- The Science of Ritual, Frits Staal, Bhandarkar Oriental Research Institute, 1982.
- Ramopākhyāna: The Story of Rāma in the Mahābhārata, Peter M. Scharf, RoutledgeCurzon, 2003.
- Pāṇini: A Survey of Research, George Cardona, Mouton Publishers, 1976.
- A Reader on the Sanskrit Grammarians, J. F. Staal (ed.), MIT Press, 1972.
- संस्कृत व्याकरण शास्त्र का इतिहास, युधिष्ठिर मीमांसक, रामलाल कपूर ट्रस्ट, बहालगढ़ (हरियाणा), 1984
- Panini: His Work and Its Traditions, Volume One – George Cardona, Motilal Banarsidass, 1988.
- The Philosophy of Word and Meaning, Gaurinath Sastri, Sanskrit College, Calcutta, 1950.
- Philology: The Forgotten Origins of the Modern Humanities, Turner, James, Princeton, NJ: Princeton University Press, 2014.

Journal

- Kiparsky, Paul. On the Architecture of Pāṇini's Grammar. Paper presented at the Conference on the Architecture of Grammar, CIEFL, Hyderabad, January 15–17, 2002. https://web.stanford.edu/~kiparsky/Papers/hyderabad.pdf
- Saxena, Parul, Pandey, Kuldeep, and Saxena, Vinay. "Panini's Grammar in Computer Science." Recent Research in Science and Technology, vol. 3, no. 7, 2011, pp. 109–111. ISSN: 2076-5061. Published by Scholar Journals. Available at: https://www.researchgate.net/publication/216868723
- Kak, Subhash C. The Paninian Approach to Natural Language Processing. International Journal of Approximate Reasoning, Vol. 1, 1987, pp. 117–130. Elsevier Science Publishing Co., Inc., 52 Vanderbilt Avenue, New York, NY 10017, USA. ISSN: 0888-613X. https://doi.org/10.1016/0888-613X(87)90007-7
- Dr. Prem Lata Chandra and Abhay Kumar Mishra. Sanskrit's affiliation to European languages and competing models of 'Similarity': Sir William Jones, Gaston-Laurent Coeurdoux, and Friedrich Schlegel. Int. J. Arts Humanit. Social Stud. 2024;6(1):22-29. https://doi.org/10.33545/26648652.2024.v6.i1a.84

——