पाणिनि : An Informatician

Amba Kulkarni

Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad

22^{nd} August, 2013
Scope of Advance Research Opportunities in Ancient Indian Sciences

Outline

- Analysis of Sanskrit Language
- Structure of Aṣṭādhyāyī
  - Syntax of Aṣṭādhyāyī
  - Ordering of the rules
  - Conflict Resolution
  - Modularity
  - अनुवृत्तिः
  - शिवसूत्राणि

Etymology:

Informatics: German Informatik
Fench Informatique

Information + automatic ⇒ informatic (Science of Automatic processing of Information)

A person who practices Informatics: Informatician

Pāṇini's Aṣṭādhyāyī

- Circa 500 B.C.E.
- Extant Grammar of the then prevalent Sanskrit Language
- Around 4000 sutras
- 8 chapters 4 sections each

One of the greatest monuments of human intelligence (Bloomfield) is only beginning to claim its rightful position in linguistics. Many of the insights of Panini's grammar still remain to be recaptured, but those that are already understood constitute a major theoretical contribution.

Prof. Paul Kiparsky
'The encyclopedia of Language and Linguistics', Asher, pp 2923.

Panini, then, was not an ancient and nebulous precursor of a science in which everything has since been done better, but a distant colleague of genius from whom linguists are still able to learn.

Prof. Fritz Staal
Pāṇini , 'Encyclopedia of Languge and Linguistics', Vol-6, Page no: 2917; Ed: Asher

Not only Panini was by far the first linguist in recorded history, but I claim he was the first informaticien, 24 centuries before computers came into existence.

– Prof. Gérard Huet, in the Inaugural Speech at the First International Sanskrit Computational Linguistics Symposium, Paris, 29th october, 2007

पाणिनि: An Information Scientist

Evident from

- The structure of Aṣṭādhyāyī
- His method of analysis of Sanskrit Language

Analysis of Sanskrit Language

Pāṇini paid utmost attention to the dynamics of Information flow while analysing Sanskrit.

We cite 3 sūtras to highlight this point.

- अनभिहिते (1.3.1) (*Where* is the information Coded)
- स्वतन्त्रः कर्ता (1.4.54) (*How much* information is coded)
- समानकर्तृकयोः पूर्वकाले (3.4.21) (*How* is the information coded)

In this talk I'll concentrate more on the second aspect of Pāṇini – the one related to the structure of his monumental work: Aṣṭādhyāyī

Structure of Aṣṭādhyāyī

A) Features of a Computer Programme

a) A Computer Programme is written in an Artificial Language
b) This artificial language has well-defined Syntax

Aṣṭādhyāyī is written in SANSKRIT
The syntax as well as the programme are intermixed in the same piece of work.

Syntax of Programming Languages

Backus Naur Form (BNF)
A BNF specification is a set of derivation rules, written as

<symbol> ::= Expression

- <symbol> is a nonterminal
- Expression consists of one or more sequences of symbols separated by |'
- Symbols that never appear on a left side are terminals
- Symbols that appear on a left side are non-terminals
- '::=' means that the symbol on the left is replaced with the expression on the right.

Syntax of Aṣṭādhyāyī

पदम् ::= सुबन्तम्
| तिङन्तम् (सुप्तिङन्तं पदं 1.4.14)
;
सुबन्तम् ::= प्रातिपदिकम् सुप्
प्रातिपदिकम्::= कृत् (कृत्तद्धितसमासाश्च 1.2.46)
| तद्धित
| समास
| terminal_प्रातिपदिकम्
;
समास::= अलौकिक_विग्रहः
;
अलौकिक_विग्रह:::= प्रातिपदिकम् सुप् प्रातिपदिकम् सुप् (सह सुपा 2.1.4)
;

Syntax of Aṣṭādhyāyī

Examples of Syntax:

- तस्मिन् इति निर्दिष्टे पूर्वस्य (1.1.65)
- तस्मात् इति उत्तरस्य (1.1.66)
- षष्ठी स्थानेयोगा (1.1.48)

Syntax of Aṣṭādhyāyī

Ingermann observed that the sūtras of Aṣṭādhyāyī have the same structure and suggested 'Pāṇini Backus form(1967 ACM Communications)'.

W{5} W{6} W{7} ::= W{5} W{1} W{7}

इक्{6} अच्{7} ::= यण्{1} अच् {7} (6.1.74)
(Bhate Kak)

Ordering of the rules

- Sañjñā triggers the rules
  समासान्ताः, द्वन्द्वे, अव्ययीभावे , उत्तरपदे, संहितायाम् , ..
- पूर्वत्र असिद्धम् (8.2.1)
- असिद्धवत् अत्राभात् (6.4.22)
- षत्वतुकोरसिद्धः (6.1.83)

Ordering of the rules

असिद्धवत् अत्राभात्

- हुझल्भ्योः हेर्धिः (6.4.101)
- शा हौ (6.435)

6.4.101: शास् + हि → शास् धि
6.4.35: शास् + हि → शा + हि

Application of one rule blocks the application of the other.
Both the rules need to be applied.

Ordering of the rules

Instead of a rule as
R: a b $\rightarrow$ c d

Pāṇini states it as a combination of two rules:

R_1: a b $\rightarrow$ c b
R_2: a b $\rightarrow$ a d

Implications:

- Brevity
  Instead of n_1 * n_2 rules, only $n_1 + n_2$ rules are needed.
- Parallel Computing

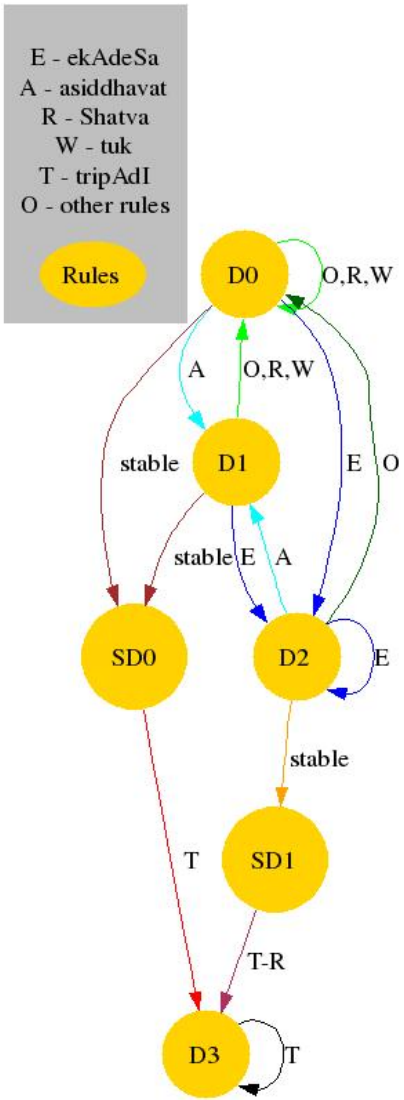
Conflict Resolution

Programming Languages

- Longest match
- Precedence Rules

Aṣṭādhyāyī

- विप्रतिषेधे परं कार्यम् (1.4.2)
- आ कडारात् एका सञ्ज्ञा (1.4.1)
- पर–नित्य–अन्तरङ्ग–अपवादानां–उत्तरोत्तरं–बलीयः


E - ekAdeSa
A - asiddhavat
R - Shatva
W - tuk
T - tripAdI
O - other rules
Rules
D0
O,R,W
A
O,R,W
stable
D1
E
O
stable
E
A
SD0
D2
E
stable
T
SD1
T-R
D3
T

Modularity

The rules related to a particular task are grouped together.
For example,
consider the following sūtras which identify sounds used as markers (*anubandha*).

- उपदेशे अच् अनुनासिक इत् 1.3.2
- हल् अन्त्यम् 1.3.3
- न विभक्तौ तुस्माः 1.3.4
- आदिः ञिटुडवः 1.3.5
- षः प्रत्ययस्य 1.3.6
- चुटू 1.3.7
- लशकु अतद्धिते 1.3.8

Modularity

If we take into account the recurrence (*anuvṛtti*) of terms from preceding sūtras, the rules may be rewritten (indicating the anuvṛtti by indentation) as

- उपदेशे इत्
  - अच् अनुनासिक 1.3.2
  - हल् अन्त्यम् 1.3.3
    - न विभक्तौ तुस्माः 1.3.4
  - आदिः
    - ञिटुडवः 1.3.5
    - प्रत्ययस्य
      1. षः 1.3.6
      2. चुटू 1.3.7
      3. लशकु अतद्धिते 1.3.8

Translation of this set of rules into a simple algorithm will show the parallel between Pāṇini's sūtras and a computer algorithm.

Features of Aṣṭādhyāyī .. contd

Object Oriented Programming:

Encapsulation of data with the (markers to the) functions
Bhaj + (gh)a(ñ): In the presence of gh, j → g

Inheritance:

Multiple inheritance → arranged as a linear inheritance
Taddhita pratyaya
(Ashwini Deo 2007)

अनुवृत्तिः

Factorisation in Mathematics:
$2 * (3 + 4) = 2 * 3 + 2 * 4$

Factorisation in Languages:

रामः गृहं गच्छति
रामः पाठं पठति

रामः गृहं गच्छति पाठं च पठति

Or simply,
रामः गृहं गच्छति,
पाठं पठति.

अनुवृत्तिः

Consider the following sūtras:

- उपदेशे अच् अनुनासिक इत् 1.3.2
- हल् अन्त्यम् 1.3.3
- न विभक्तौ तुस्माः 1.3.4
- आदिः ञिटुडवः 1.3.5
- षः प्रत्ययस्य 1.3.6
- चुटू 1.3.7
- लशकु अतद्धिते 1.3.8

अनुवृत्तिः

- उपदेशे (a) (= इत्) (c)
  - अच् अनुनासिक (b)
  - हल् अन्त्यम् (d)
    - न विभक्तौ तुस्माः (=इत्) (e)
  - आदिः (f)
    - ञिटुडवः (=इत्) (g)
    - प्रत्ययस्य (h)
    1. षः (=इत्) (i)
    2. चुटू (=इत्) (j)
    3. लशकु (=इत्) अतद्धिते (k)

अनुवृत्तिः

- a c
  - b
  - d
    - e
  - f
    - g
    - h
      1. i
      2. j
      3. k

a (b + de + f [g + h { i + j + k }]) c

अनुवृत्तिः

No Proper Nesting; मण्डूक–प्लुति

- 1.1.1 वृद्धिः आदैच्
- 1.1.2 अदेङ् गुणः
- 1.1.3 इको गुणवृद्धी (वृद्धिः गुणः)

अनुवृत्तिः

Maximum advantage of features of Natural Language:

How are the complete phrases reconstructed?

आकाङ्क्षा (Expectancy): Major role in deciding the अनुवृत्ति (Bhate)

अनुवृत्तिः

Example of borrowing from as many as 11 sūtras

Original sūtra: 3-3-65 क्वणः वीणायां च
After अनुवृत्ति: 3-3-65: क्वणः वीणायां च (प्रत्ययः परः च आद्युदात्तः च धातोः कृत् क्रियायां क्रियार्थायाम् भावे अकर्तरि च कारके सञ्ज्ञायाम्) अप् उपसर्गे वा नौ

अनुवृत्तिः

Some Statistics:

Total sūtras	(3984) 4000
Total Words (with sandhi)	(7007) 7000
Total Sandhi split words	9843
Total words after repeating the words with anuvṛtti	40,000
Compression because of anuvṛtti	1/6
In terms of byte size, compression	1/3

Significant from Oral Tradition.
The time to memorise the sūtras grows exponentially.
With anuvṛtti, a student can memorise the complete Aṣṭādhyāyī in about only 6 months!

शिवसूत्राणि

Normal Arrangement of Alphabet

अ आ इ ई उ ऊ ऋ ॠ ऌ ए ऐ ओ औ अं अः
क ख ग घ ङ
च छ ज झ ञ
ट ठ ड ढ ण
त थ द ध न
प फ ब भ म
य र ल व
श ष स ह

शिवसूत्राणि

Pāṇini required several(42) subsets of this alphabet to describe various operations.

शिवसूत्राणि

Some of these subsets:

All vowels
All consonants
All vowels + semivowels + Anunāsikas (अम्)
The voiceless stops (खय्)
stops and sibilants (झर्)
Any of ब ग ड द (बश्)
Any of भ घ ढ ध (भष्)
Any of य व र ल ञ म ङ ण न झ भ (यञ्)
....

शिवसूत्राणि

It is not advisable to give 42 names to these sets.
It will be difficult to memorize the association.

These are Partially ordered sets.

Pāṇini arranged them linearly in the form of 14 ShivasUtras.

शिवसूत्राणि

अ इ उ (ण्)
ऋ ऌ (क्)
ए ओ (ङ्)
ऐ औ (च्)
ह य व र (ट्)
ल (ण्)
ञ म ङ ण न (म्)
झ भ (ञ्)
घ ढ ध (ष्)
ज ब ग ड द (श्)
ख फ छ ठ थ च ट त (व्)
क प (य्)
श ष स (र्)
ह (ल्)

शिवसूत्राणि

Optimality of these sūtras is proved independently by

Kiparsky (linguistically)
and Petersen (mathematically)

शिवसूत्राणि

Given a set of Partially Ordered Sets,

Now it is possible to tell

Whether the elements are

Śivasūtra encodable or not.

Ref: Petersen(2008)

धन्यवादः