Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 05, May -2018

REGIONAL LANGUAGE DRIVEN PRESENTATION TOOL

Sharayu Puranik 1st, Dr. D.M.Thakore 2nd

[1]*Department of Computer Engineering, Bharati Vidyapeeth (Deemed to be University), College of Engineering, Pune, INDIA*
[2]*Faculty of Department of Computer Engineering, Bharati Vidyapeeth (Deemed to be University), College of Engineering, Pune, INDIA*

Abstract – *India is a big research hub for Natural Language Processing area. There are 22regional languages in India, which can be worked upon for various aspects of processing. Natural Language Processing is an interdisciplinary research area at the border between linguistics and artificial intelligence. This paper aims at providing details about a regional language driven tool, which will take input in the form of Sanskrit Nyasa (from Mathematical Grantha Leelavati) and would convert that mathematical expression into Graphical format.*

***Key Words*:** *Regional, Leelavati ,Algebra ,Planetary motions Astronomy.*

I INTRODUCTION

Having 22 Regional Languages, India is always treated as big research hub for Natural Language Processing area. Aim of Natural Language Processing and Artificial Intelligence area is to develop computer programs capable of human-like behavior related to 'understand given texts or produce meaningful texts' in natural languages such as Sanskrit, Marathi, Hindi, English and many other regional languages.
The most important applications of natural language processing include Retrieval of Information, Organization of the collected information, Machine Translation, Automatic Summarization, Sentiment Analysis, Text Classification and many more.

As in any science, activities of researchers are manly concentrated on its internal art and craft. Many problems arise during analysis and generation of Natural Language texts. Researchers focus on the solution of these problems of Semantic and Syntactic analysis, compilation of dictionaries, language text and grammar ambiguities.

In this project, we map the Nyasa written in Sanskrit, into corresponding Algorithm steps and its Graphical representation. This would clearly show the mathematical steps mentioned in that specific method.

Being Emerging field in India, Natural Language Processing (NLP) has a very good potential for research. Extensive research can be done in below NLP levels;

1. 'Conversion from Speech to text' technology,
2. Understanding of the Natural Language Text (with context) and
3. Effective and Efficient management/organization of the knowledge

The main objective is to develop applications which are more relevant to those people speaking various regional languages.

Natural Language Processing (NLP) is a field of computer science, artificial intelligence and computational linguistics concerned with the interactions between computers and human (natural) languages. The idea of using a natural language for computer programming is to make it easier for people to talk to computers in their native languages. For many, it is tedious and painful to learn Computer friendly languages like assembly, C, C++, Java, LISP etc. Use of native languages for Computer programming relieves such pain of learning Computer languages.
Multiple languages are spoken in India, each with its own flavour. Being mother of all languages, Sanskrit is the perfect language for computer programming. This language is grammatically perfect and has huge treasure of knowledge from all the fields.

Among all the Natural Languages, Sanskrit in its style is identified to be the best language which has minimum deviation. The creator of Sanskrit grammar, Panini, formulated 3,949 rules. Sanskrit is said to be a Mother of all languages. It deals with multiple limitations of Artificial Intelligence like NLP, Semantic Net, Vibhakti, Dual Case, Inflection based Syntax etc. Sanskrit language fulfils almost all of the prerequisites of a Natural Language Processor.

Information retrieval and information organization are the most important applications of natural language processing.

Some other applications are Natural Language Interfaces, Machine Translation and many more. In Natural Language Processing, below eight Technical areas can be considered for both theoretical study and application development:–

- Information Retrieval and Text Clustering
- Morphology, Syntax, Named Entity Recognition
- Semantics
- Opinion, Emotions, Textual Entailment
- Text and Speech Generation
- Machine Translation
- Educational Applications
- Applications

We can apply NLP techniques to retrieve treasure of knowledge, written by our ancestors, in Sanskrit.

II LITERATURE SURVEY

Sr No	Paper Name	Year of Publication	Methodology Used	Description
1.	Relational Learning of Pattern-Match Rules for Information Extraction	1998	Machine Learning	Information extraction is a form of shallow text processing that locates a specified set of relevant items in a natural-language document. Systems for this task require significant domain-specific knowledge and are time-consuming and difficult to build by hand, making them a good application for machine learning.
2.	Semantic Lexicon Acquisition for Learning NL Interfaces	1998	Artificial Intelligence	Describes WOLFIE (WOrd Learning From Interpreted Examples) system that acquires a semantic lexicon from a corpus of sentences paired with representations of their meaning.
3.	Journal of Machine Learning Research Homepage	2016	Machine Learning	A tool for measuring String Similarity
4.	White Paper on Natural Language Processing	2007	Natural Language Processing	Semantics analysis in Natural Language Processing
5.	**Leelavati**	2013	Hindi Translation of Leelavati Grantha	All the mathematical methods described in Original Leelavati Grantha are translated in Hindi language and given in Nyasa form.
6.	NLP (Natural Language Processing) for NLP (Natural Language Programming)	https://link.springer.com/chapter/10.1007/11671299_34-2006	Natural Language Processing	This paper tries to address gap, between Programming Language and NLP,by proposing a system that attempts to convert natural language text into computer programs. Starting with an English text, we show how a natural language programming system can automatically identify steps, loops, and comments, and convert them into a program.

III PROPOSE SYSTEM

Bhaskaracharya wrote Siddhantashiromani at the age of 36, Lilavati is the first part of it. The main Grantha Siddhantashiromani consists of four parts namely (Bhaskaracharya: 1144 – 1223 AD).

1) Lilavati (लीलावती)

2) Algebra (बीजगणित)

3) Planetary motions (ग्रहगणित)

4) Astronomy (गोलाध्याय).

Lilavati, the first 'prakarana' of Siddhantashiromani deals with 'Pati-Ganit' i.e. 'VyaktaGanit' or Arithmetic in today's Mathematical Term. It contains 278 verses. Being a Kavi also, Bhaskaracharya has written these verses in Poetic form (Shlokas) in Sanskrit language. There are certain verses which deal with Mensuration (measurement of various Geometrical Objects), Volume of Pyramids, Cylinders, heaps of grains etc., wood cutting, shadow and trigonometric relationship. Also on certain elements of Algebra such as finding an unknown quantity subject to certain constraints with the help of supposition method.

The Lilavati consists of 279 verses of rules and examples. The main contents are:

- Basic arithmetic operations including square roots and cube roots calculation for numbers, fractions, and the effect of text encryption.
- The rule of three, rule of five and so on
- Bartering, buying and selling
- Permutations and combinations
- Progressions and series
- Geometrical operations
- Solutions to indeterminate equations

In proposed system, we consider the Nyasa written in Lilavati Grantha. These Nyasa represent specific Mathematical formulae written for specific Mathematical Methods like Addition, Subtraction and so on.

Proposed system takes Sanskrit Nyasa as an input from front end GUI. Maps these Nyasa tokens into corresponding English words (Mathematical operations) and prepare and algorithm (in English) for the given method. The algorithm then also would be represented in the form of Flowchart. This would help user to clearly visualize the mathematical steps mentioned in that specific Sanskrit Nyasa.

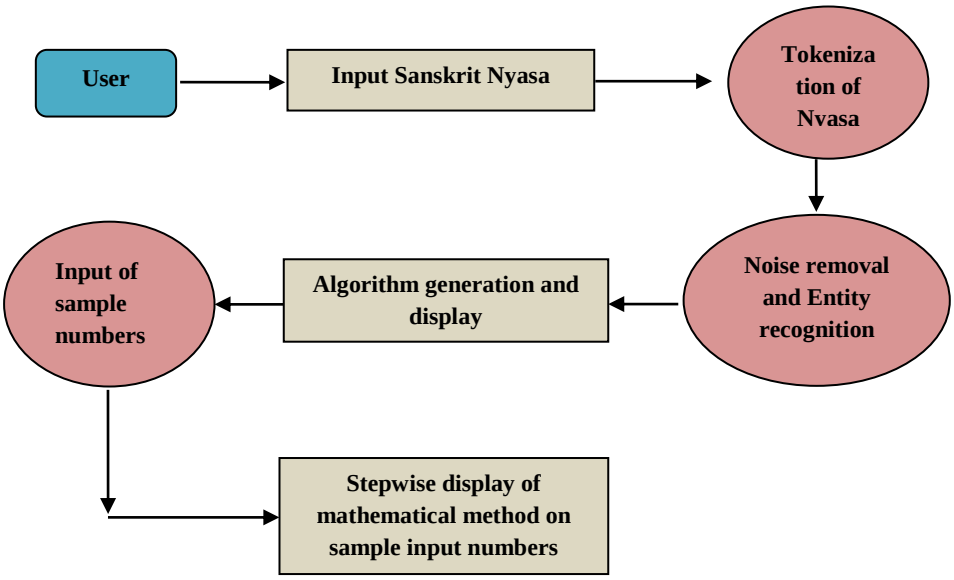

Fig. System Architecture

IV METHODOLOGY

APIs are developed to process the Sanskrit (Regional Language) statements for Tokenization, Lexical Normalization and Entity Recognition. These can be used for the other Regional language, provided with the database for that specific language.
For the generation of English language mathematical steps mentioned in Sanskrit Nyasa, Divide and Conquer algorithm is used. Once the mathematical operator is found from the Nyasa, the Nyasa is 'Divided' into left and right side of the operator. The operations to be performed on left and right side operands are formed stepwise. The final method of calculation is extracted and that operation is performed as 'Combine' process.

Modular Design of a System:

The first module is used to take input in form of Sanskrit Nyasa from user. User selects the type of mathematical expression for which Algorithm would be generated.
Types of mathematical formulas are

- गुणने करणसूत्रम् for Multiplication methods,
- भागाहारे करणसूत्रम् for Division methods,
- वर्गे करणसूत्रम् to find Square of a number,
- वर्गमूले करणसूत्रम् for finding Square root of a number,
- घने करणसूत्रम् for finding Cube and
- घनमूले करणसूत्रम् to find Cube root of a number.

Second and third module deals with aspects of Natural Language Processing namely

1. Removal of Noise from Input String (Removal of unwanted tokens)
2. Lexical Normalization (extract the exact word which may have multiple representations)

Entities are defined as the most important chunks of a sentence. Next module will extract entities from given Nyasa. Using look-up database table [Table 1], the module will find out mathematical operators from Nyasa.

Module 5 and 6 deals with preparation of algorithm and execution of the method on sample input numbers.

V CONCLUSION

The paper proposes a Regional language tool (Which is a Sanskrit Language) which would accept input Nyasa from Lilavati Grantha and would represent the mathematical methods mentioned in it graphically. This tool can be used by Learners of Sanskrit Language and can be extended for other Indian Languages in which the mathematical formulae are mentioned. The same can further be extended for other learning areas like Science, Economics, Politics and many other where ancient Indian literature is present in Regional Languages.

VI REFERENCES

[1] Mary Elaine Califf and Raymond J. Mooney, Relational Learning of Pattern-Match Rules for Information Extraction
[2] Cynthia A. Thompson and Raymond J. Mooney TR AI98-273, Artificial Intelligence Lab, University of Texas at Austin, May 1998. Semantic Lexicon Acquisition for Learning Natural Language Interfaces.
[3] KonradRieck, Christian Wressnegger; 17(9):1−5, 2016, Journal of Machine Learning Research Homepage Harry: A Tool for Measuring String Similarity.
[4] Ralph Weiscbedel, Chairperson BBN Systems and Technologies Corporation ,Jaime Carbonell Carnegie-Mellon University Barbara Grosz Harvard University ,Wendy Lehnert University of Massachusetts, Amherst Mitchell Marcus University, of Pennsylvania Raymond Perrault SRI International ,Robert Wilensky University of California, Berkeley, White Paper ; Natural Language Processing.
[5] http://www.omg.org/news/meetings/workshops/SOA_MDA_WS_Workshop_CD/00-T-03_Amsden-Part3.pdf, Business Driven Development.
[6] Jochen M. K¨uster, Jana Koehler, and KseniaRyndina IBM Zurich Research Laboratory 8803 R¨uschlikon, Switzerland Improving Business Process Models with Reference Models in Business-Driven Development.
[7] M.D. Pandit, 2013, English translation of Sanskrit Grantha Lilavati. Rada Mihalcea, Hugo Liu, Henry Lieberman, 2006, Natuaral Language Processing for Natural Language Programming

