

The Knowledge and Meta Knowledge Representation Framework for Context based Searching for Lyric Writing in Telugu

Ramakantha Rao Chakalakonda¹, Raghu B Korrapati²

¹Royalaseema University, Kurnool – 518 002, A.P., India

Abstract: Representation of knowledge and Meta knowledge plays a key role in context based searching. Recent advances in computational Linguistics and natural processing has brought linguistic formalism closer to the knowledge representation (KR). The techniques used in artificial Intelligence (AI) to represent knowledge, can be used successfully in computational linguistics in representing the grammar and other language and its literary requirements. Other areas where KR can be used in computational linguistics are common sense language, and default reasoning. In this paper the Knowledge representation related to the grammatical and “Chandasu” and “Alankara Sastra” related knowledge used in Telugu literature is explained. The paper discusses what kind of knowledge related information has to be represented, how the subject issues are represented, how they are related to each other and how reasoning is generated from this knowledge, using Meta knowledge related to this knowledge. In context based searching related to Telugu lyric the most important task is selecting the set of right words, the knowledge data base short lists these words, and aids in giving suggestions. Knowledge representation is one of the central and crucial parts for context based searching for word suggestion. The knowledge has to be organized based on purpose it will be used. Based on different types of knowledge forms, each form requires different kinds of representation. The Knowledge Representation models/mechanisms are often based on; logic, rules, Frames, Semantic Nets etc. Different types of knowledge require different kinds of reasoning. For context based searching used for lyric writing, the knowledge related to grammar and chandasu has to be represented for suggesting the right word suitable for the context. The lexical knowledge adds more connotations to the problem. The synonyms, thesaurus, the relationship of a word with other words, parts of speech etc., have to be taken into consideration to this problem of context based search. In this paper, we shall describe the knowledge and Meta knowledge representations done for context based searching.

Keywords: YATI- PRASA-ANTHYANU PRASA (These are rules of Chandasu in Telugu poetry), SHABDA ALANKARAS (sound related like rhythm), meaning related ARDHA ALANKARAS (based on meaning), and CHANDASU (rules governing the poetry writing), MATHRAS (unit of time pronounce a character).

1. Introduction

Representation of knowledge and Meta knowledge plays a key role in context based searching. Recent advances in computational Linguistics and natural processing has brought linguistic formalism closer to the knowledge representation (KR). Knowledge is the expertise or skill in a domain specific area or subject about facts and information in that domain area. In the situation of a problem or in a situation of complexity, Knowledge aids in decision making to resolve that problem [5]. The term knowledge is also used to mean the confident understanding of a subject with the ability to use it for a specific purpose if appropriate. Knowledge leads the decision maker from the state of uncertainty to the state of certainty. Knowledge can be classified, based on how that knowledge is going to be used, as Decision-oriented Knowledge, Procedural knowledge, Reasoning knowledge, Assimilative knowledge, Descriptive knowledge, Classification-based Knowledge [7]. Knowledge acquisition involves complex cognitive processes which are perception, learning, communication, association and reasoning. Knowledge representation should be so designed to help in the process of reasoning and inferencing (i.e. drawing conclusions) from knowledge. Meta-knowledge is the knowledge about a preselected knowledge. Meta-knowledge is a fundamental conceptual instrument in knowledge engineering, knowledge management.

"Knowledge representation (KR) is the study of how knowledge about the world can be represented and what

kinds of reasoning can be done with that knowledge. When we design a Knowledge Representation we have to make choices across a number of design spaces. Expressivity is the most important decision to be made in Knowledge Representation. Whereas propositional logic is less expressive KR, auto epistemic temporal modal logic is more expressive KR. Less expressive KRs may be both complete and consistent (formally less expressive than set theory). More expressive KRs may be neither complete nor consistent. The purpose of KR could be any - decision making, need to give description of other knowledge, need to do some procedure, need to reason etc.

The new research interests and advances in computational linguistics has opened up new areas in the study of commonsense knowledge, inheritance, default reasoning and even domain specific knowledge's, which were formerly thought impossible and beyond the scope of grammar and linguistics [3] It gave new thoughts in determining ' what is lexical knowledge' and what is outside the scope of such knowledge. Lexical knowledge is a term used for knowledge in the form mainly facts, figures, data and information in general taken from reliable sources. The scope of lexical knowledge is wide apart from some individual collections and productions. Narrowing the term lexical knowledge and to mention a few- "the synonyms, thesaurus, the relationship of a word with other words, Parts of speech, syntactic relationship and other taxonomies, nomenclatures, idioms and phrases, inflected and derived forms of words, associations with the word" fall in its scope. The language

and literature adds more such relationships. All these are decision factors in context based searching in determination of a word, based on context [4]. So, the knowledge representing all these facts will be very complex and have to be organized based on its use. We shall discuss these details in the following paragraphs.

2. The KR problem in Context based searching

Efficiency and interoperability are key issues in the problem of knowledge representation. According to Alex, in his paper on KR issues [Alex], the main issues in knowledge representation include, adequacy of representation, measure of knowledge role with respect to the goal that is trying to be achieved, measure of overall quality of knowledge within the knowledge representation, measure of knowledge uncertainty for the knowledge utilization by the autonomous system, measure of the consistency of knowledge. As per Rich et.al [8] in his famous book, on AI, the good System for KR should have the following properties:-

- Representational adequacy – ability to represent kind of knowledge.
- Inferential Adequacy – ability to manipulate representational structures.
- Inferential Efficiency – the ability to incorporate into Knowledge structure.
- Inquisitional Efficiency – the ability to acquire new information.

In summary efficiency of representation, ease of reasoning and use are the critical factors for knowledge representation for context based search related to current problem. The efficiency of designer lies in the fact of selecting the right knowledge representation mechanism, best suitable conditions in the light of Rich's four properties. Here are a few knowledge representation techniques considered in this work.

3. Rules

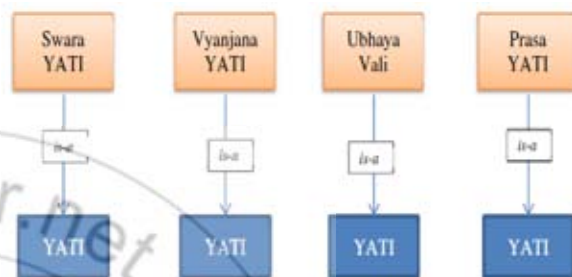
Rule based representation Contains rules, each rule a step in a problem-solving process. Human experts reason by applying "rules of thumb". Expert knowledge can often be seen naturally as rules of thumb Rules are persistent knowledge about the domain. Rule Contains facts about the world, Rules are independent of each other – new rules can be added or revised later [12]. The Frequent Rule Selection allows reasoner to continuously revise strategy, jumping from hypothesis to hypothesis multiple solutions is effectively pursued at one time. The following is an example of a rule, in this context:

If Letter(Haswa(oval)) → Then, Letter us Laghu,
 MatraCount = 1;
 If Letter(HaswaMixed(Consonant)) → Then, Letter is Laghu,
 MatraCount = 1;
 If Letter(Deergha(oval)) → Then, Letter us Guru,
 MatraCount = 2;
 If Letter(DeerghaMixed(Consonant)) → Then, Letter is Guru,
 MatraCount = 2;

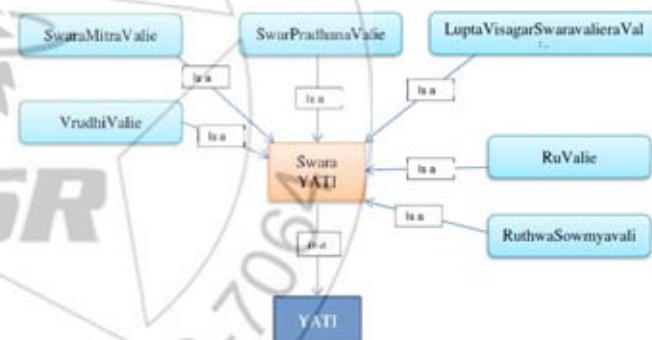
The applicable rules are selected by the process of Indexing; by searching through the rules, comparing each one's precondition to the current state and extracting all the ones that match.

4. Semantic Networks

Semantic Network is a *Functional Approach* to Knowledge Representation. A *semantic network* or *net* is a graphic notation for representing knowledge in patterns of interconnected nodes and arcs. In semantic net, information is represented as a set of nodes connected to each other by a set of labeled arcs, which represent relationships among the nodes. A Yati semantic net is shown below: -



Each of the above networks is a composition of more detailed sub-nets. For example, in the above semantic net example, say "Swara YATI" it is net of sub-nets as follows: -



network except each node represents prototypical concepts and/or situations [11]. Each node has several property slots whose values may be specified or inherited by default. The Yati and Prasa knowledge is maintained as frames; an example is given below,

SwaraMitraValie:

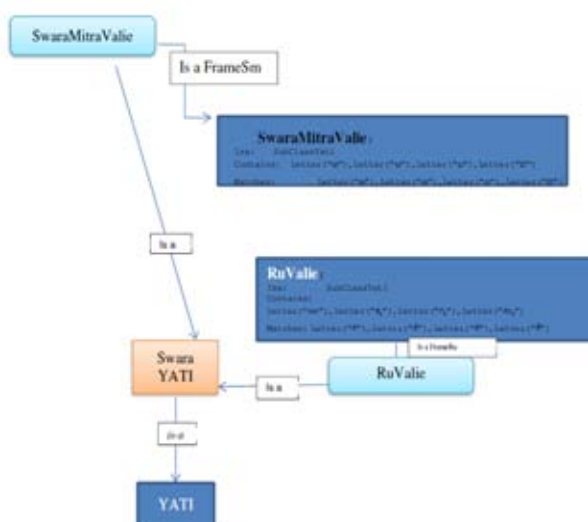
Isa: SubClassYati
 Contains: Letter("అ"), Letter("ఆ"), Letter("ఐ"), Letter("ఔ")
 Matches: Letter("అ"), Letter("ఆ"), Letter("ఐ"), Letter("ఔ")

RuValie:

Isa: SubClassYati
 Contains: Letter("ఋ"), Letter("ౠ"), Letter("ౡ"), Letter("ౢ")
 Matches: Letter("ౣ"), Letter("౤"), Letter("౥"), Letter("౦")

A. Hybrid Model (Semantic Net and Frame)

In this work, we considered a Hybrid model of semantic net and frame. The following example explains that.



Using the above mentioned rich knowledge representation model, a number of knowledge bases representing different aspects of Telugu grammar, linguistics, Chandasu and Alankara are maintained (*Raghu Korrapati, Prasa*). These knowledge areas include: -

- Laghu-Guru Determination.
- Yati
- Prasa.
- Ganamulu
- BhavanuBhavalu
- ShabdAlankara
- Arddhalankara
- Synonyms
- Antonyms
- Sandhi Rules
- Prakruthi-Vikruthi Kosamu
- Jantapada Kosamu
- Dwandwapada Kosamu
- Purana Nama Kosamu
- Vrushpadyardha Kosamu
- Rupanthara Kosamu.
- Vibhaktulu
- Verb-Noun-time rules

While a few of these knowledge bases are used in current work, extensions can be done to other knowledge areas in future works by researchers.

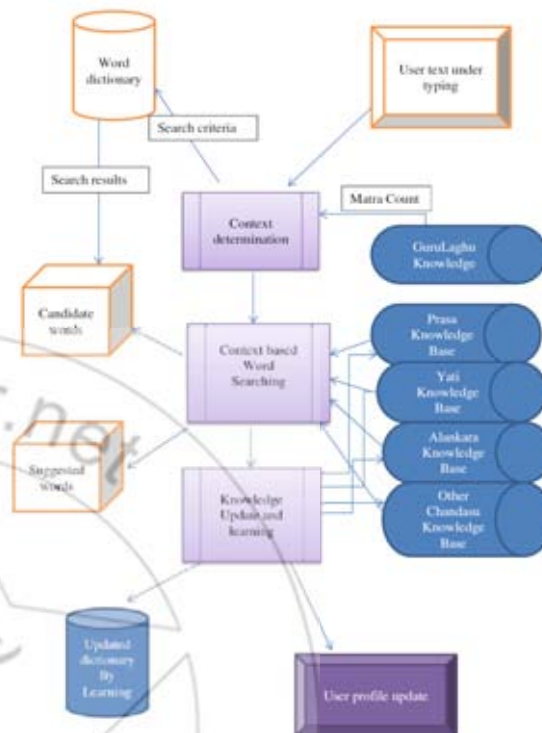
5. Knowledge bases required

In the current system, there are a few areas of knowledge that were required to maintain by the system and have to be updated for the problem of context based search. The process in which the context based searching is done is as follows:

- From the information typed by the user, the context is determined and from the dictionary a set of candidate words is selected. The GuruLaghu Rule base is used in this process.
- Once the candidate words are determined, a series of decisions are taken on this candidate word sets, using Prasa, Yati, Alankara and other chandasu knowledge bases.

- The result of this process is establishing the set of suggested words.
- Based on user selection, the word selected for the context and the context is applied for learning process and the dictionary is updated.
- The user profile is preserved for later use.

This process is shown diagrammatically below: -

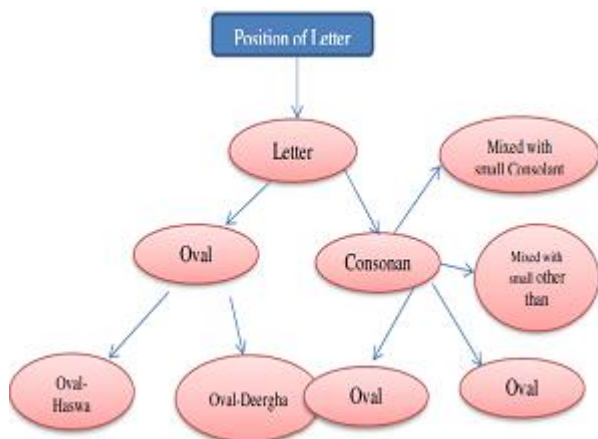


6. Details of Knowledge bases used

We have mentioned in section 2, the knowledge bases are required for the system to completely generate suggestion words, best suitable for the context. Due to limitations, only a few from the following list are considered in this work, future research workers can use the rest and develop excellent models for context bases searching. Out of the list the following are considered here, are the details of each of them are explained below:-

A. Laghu-Guru Determination

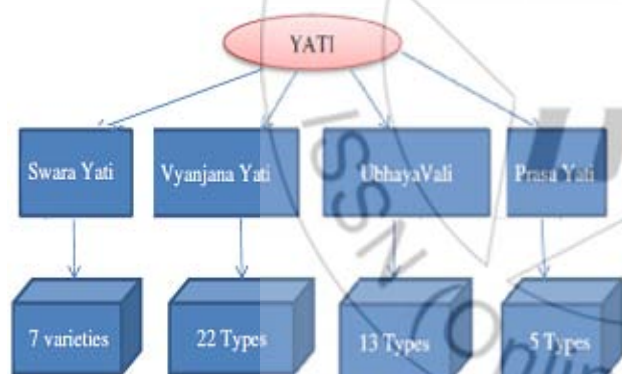
There are a number of rules that determine, whether a letter or letter combination in a word is whether a Laghuvu (taking one matra i.e. one unit of time), or Guruvu (taking more than one unit of time). This timing is very important to determine the correct place of the coming new word, in the lyric, and based on that Yati or Prasa or other Chandasu rules have to be applied. This is very important to make the song suitable for music composition, the Raga determination and swara for the lyric. The complete set of rules is stored in the Rule base of the system. There are a number of rules that govern this; we shall diagrammatically explain a few of them:-



This is a partial example; the decision tree expands in this fashion to determine the matra count of the letter under determination in the lyric.

B. YATI Determination

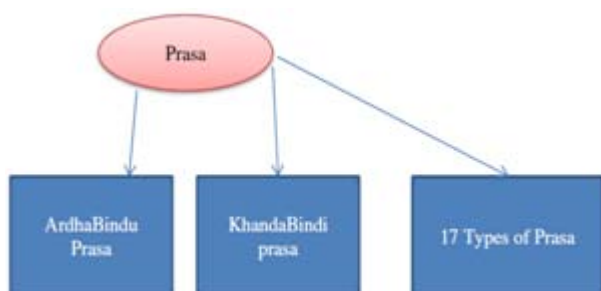
YATI is the most important requirement in Telugu poetic literature, and in particular to poems and songs, if Yati requirements are satisfied, there is an automatic rhythm comes in the song, that makes it musical. There are basically four classes of Yati, and they are in turn, a set of its sub classes, each has its own requirements. The way all these main, subclasses are represented in knowledge base is already explained earlier in KR representation section (Chando Swarupalu).



All the knowledge related to Yati is represented using semantic nets and Frames, a Hybrid approach. There are 16 varieties of prasa in Telugu literature.

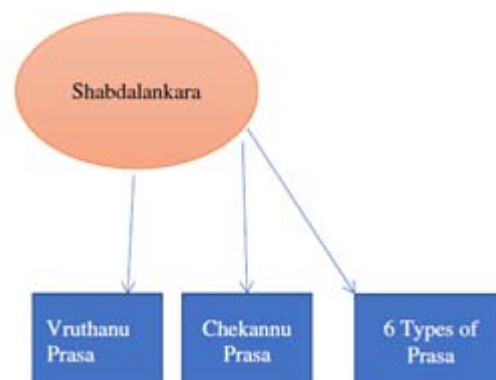
C. Prasa Determination

The Prasa adds musicales to the rhythm of the lyric, there are 17 varieties of prasa each has its own requirements, they are represented as Hybrid using Semantic net and Frame [9].



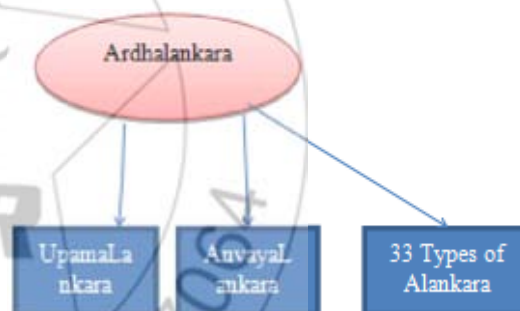
D. Shabdalanakara

These are important ornaments for a lyric which add musical sweet sound to the lyric that makes the lyric a great song to listen. There are six kinds of it, if we use them in our lyric; it makes the song very easy to compose. The diagram explains some details: -



E. Arddhalankara

There are 32 varieties of Arddhalankara, which are the ornaments that add beauty to the lyric, though they do not add any value to the musical nature of lyric, but add a lot of value to the contents of the lyric. A poet's expertise is evaluated based on value of these. They fall in the category simile and metaphor etc. often described in poetic literature.



F. Synonyms

Synonyms are those words which mean the same, but different words. They are very useful for a poet, in a lyric. The poet may be looking for a word which means the same, word he is having, but suitable with respect to usage for the context based on number of matra, gana, chandasu and Alankara requirements. This semi knowledge base maintains these words which are N X N in the knowledge base.

G. Antonyms

Just like Synonyms. Antonyms are equally important for a poet. These are words which mean the opposite to the word under consideration. They are very useful for a poet, in a lyric, to generate opposite meaning. The poet may be looking for a word which means different for what he wants, but suitable with respect to usage for the context based on number of matra, gana, chandasu and Alankara requirements. This knowledge base maintains these words which are (N X N) knowledge representation.

7. Meta Knowledge representation

Meta-knowledge is knowledge about knowledge. For example, knowledge about how a system makes decisions would be considered meta-knowledge. The Meta Knowledge could be, Declarative Knowledge, Procedural Knowledge. The meta-knowledge is useful for generalizing and improving efficiency of the system. The context based system under consideration heavily depends on Meta knowledge, to select what kind of knowledge base has to be used, what kind of inferencing technique has to be applied, and how to extract the information required from the results. Meta-knowledge is highly used in the learning process, which consists in transforming information into knowledge by attributing values to knowledge from other domains: truth, usefulness, importance, knowledge priority.

8. Meta-Analysis and Conclusion

Knowledge representation is a key factor in context based searching. Correct representation and an efficient technique used on Knowledge base makes the search very efficient in determining the right word for the context. We used in this work, a number of representation methods and inferencing techniques to make the search efficient and fast. This area of research bridges three areas of knowledge: -

- Computer science, computational linguistics and artificial intelligence.
- Telugu Literature, grammar and Chandasu.
- The science of music and composition of music.

The integration of knowledge in these three areas of knowledge is determined by meta-analysis (*Raghu Korrapati*). The results of this analysis bring a close relationship among these areas. The results will be discussed in our paper being written. The hybrid knowledge representation chosen in this work, has generated efficiency of representation and inferencing.

References

- [1] Raghu Korrapati, COMPUTATIONAL LINGUISTICS: SPECIAL REFERENCE TO APPLICATIONS IN TELUGU LANGUAGE USING INTERDISCIPLINARY META ANALYSIS APPROACH.... 77
- [2] Raghu Korrapati, Ramakantha Rao Chakalakonda, Rayalaseema University, - A CONTEXT BASED SEARCHING ALGORITHM FOR LYRIC WRITING IN TELUGU- Volume 14, Number 1 Allied Academies, International Conference New Orleans, LA, April 14-16, 2010, Academy of Information and Management Sciences, PROCEEDINGS, Copyright 2010 by the DreamCatcher Group, LLC, Cullowhee, NC, USA, Volume 14.
- [3] Raghu Korrapati, Ramakantha Rao Chakalakonda, Rayalaseema University, , Page.58 - THE DESIGN ASPECTS OF CONTEXT BASED SEARCHING ALGORITHM FOR LYRIC WRITING IN TELUGU – AN INTELLIGENT APPROACH USING COMPUTATIONAL LINGUISTIC TECHNIQUES Allied Academies International Internet Conference

- 2010 Allied Academies Internet Conference Proceedings Affiliates
- [4] Raghu Korrapati, TVVV Prasad, Rayalaseema University - A NATURAL LANGUAGE APPLICATION TO DETERMINE 'CHANDASSU' - (GRAMMAR) & PREDICT THE RIGHT WORD IN TELUGU POETRY... 70
- [5] Alex Abramovich, On Knowledge Representation issues, www.isd.mel.nist.gov/research_areas/research.
- [6] Abramovich A. Domain knowledge representation. — Wuhan: Wuhan University, 2009.
- [7] Pejman Makhfi, Introduction to Knowledge Modeling- http://www.makhfi.com/KCM_intro.htm
- [8] Rich, Elaine, Kevin, Knight – Artificial Intelligence, McGraw-Hill Inc.
- [9] Malladi Krishna Kumar, Telugu Grammar –, Victory Publications.
- [10] Chando Swarupalu, Telugu University – – MA course Material, Potti Sree Ramulu.
- [11] Charniak, Drew McDermott – Introduction to AI – Addison Wesley Publishing Company.
- [12] Patrick Henry Winston – AI – Pearson Education.

Author Profile



Ramakantha Rao Chakalakonda received Masters degrees in computer science. He is B.Sc., M.Sc. M. Tech(IT), M.Phil., MCA and MBA in finance. He has worked in USA and India and has more than 25 years of experience in implementing computer based solutions. He is now a Research Scholar at Rayalaseema University.