❏ 124

A Tour Towards the Various Knowledge Representation Techniques for Cognitive Hybrid Sentence Modeling and Analyzer

Poonam Tanwar[1], T. V. Prasad[2], Kamlesh Dutt[3]
[1]Department of CSE, Manav Rachna International University, Faridabad, Haryana, India.
[2]Dean of Computing Sciences, Visvodaya Technical Academy, Kavali, Andhra Pradesh, India.
[3]CSE, National Institute of Technology Hamirpur, Himachal, India.

Article Info	ABSTRACT
	Knowledge Representation (KR) is a fascinating field across several areas of cognitive science and computer science. It is very hard to identify the requirement of a combination of many techniques and inference mechanism to achieve the accuracy for the problem domain. This research attempted to examine those techniques, and to apply them to implement a Cognitive Hybrid Sentence Modeling and Analyzer. The purpose of developing this system is to facilitate people who face the problem of using English language in daily life.

Corresponding Author:

Poonam Tanwar,
Department of CSE,
Manav Rachna International University,
Faridabad, Haryana, India.
Email: Poonamtanwar.fet@mriu.edu.in

1. INTRODUCTION

Artificial intelligence (AI) incorporates the intelligence of a human in the machine. Basically, AI is the branch of science which makes the machine exhibit intelligence as human beings for a particular domain. In 1950s, Alan Turing, the British Mathematician presented a paper on Computing Machinery and Intelligence that made an inquiry that if a machine could pass a certain test known as the Turing test, then system could be an intelligent. In this paper, Turing also considered a number of arguments for and the objections to the idea that computers could exhibit intelligence. McCarthy and Hayes (1969) state that a machine is intelligent if it solves/ perform/reason certain classes of problems requiring intelligence in humans. Other definitions for AI were also proposed like "AI is the part of computer science concerned/related with designing computer systems that exhibit the characteristics we associate with intelligence in human behavior". Charniak and McDennott (1985) states that "AI is the study of mental faculties through the use of computational models", whereas Yousheng Tian, et al (2011) states that it is the study of Cognitive Science. Mylopoulos (1983) presented a brief description on terminology and issues related to KR.

The research in this field is divided into two categories, KR and General (Learning, Planning, etc.). For making the computer or machine exhibit intelligence as human, the user will require two things, KR and Inference Mechanism. Development of an AI system has been a crucial task since at certain times incomplete information is available which can be ambiguous and uncertain. Hence, the solution to these problems is to build an effective knowledge base and an effective Inference Mechanism. Existing KR tools are used to represent either declarative knowledge or procedural knowledge, but not both. This has become the principal

motivation for designing and developing a system to represent both the types of knowledge for the purpose of reasoning, a QAS should be embedded with the system so that users can easily retrieve the facts.

The survey related to work is divided into the following categories:
1. Knowledge representation
2. Knowledge representation techniques
3. Hybrid knowledge representation system / languages
4. Other knowledge representation languages
5. Question answering system

2. KNOWLEDGE REPRESENTATION

Many of the problems in AI require extensive knowledge about the world. Objects, properties, categories and relations between objects, situations, events, states and time, causes and effects are the things that AI needs to represent. All the above defined things could be represented by the current KR System. In AI, for a specific domain system, it must have a knowledge base and various techniques for representing the knowledge by Brewster et al (2004). The KR performs the three major tasks: Acquisitions, Reasoning, Searching (Figure 1).
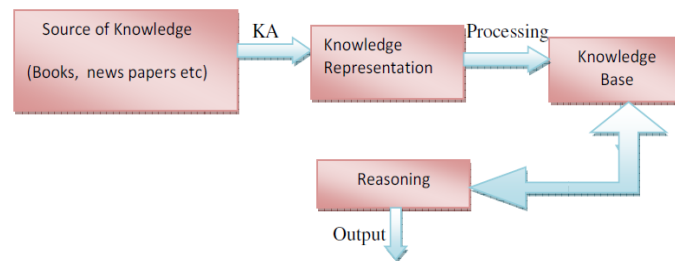

Figure 1. Knowledge representation model

2.1. Ontologies

Ontology is the study of Real World Entities and how these entities can be grouped / relates to each other and arrange them in a hierarchy shown in Figure 2 and then Categorize them according to the similarities and the differences presented by Davis et al (1993). Ontology explicitly represents the knowledge to make it easily accessible to machines. The terms which are not represented explicitly might be used with common sense. It binds the people/entities in a community based on the conceptualization. The complexity of conceptualization is proportionate to the size community. Ontology is helpful in the domain specific knowledge representation. Symbols significantly represent the concepts and their relations. Ontology comprises of statements, referred as an axiom. Ontology can be used for information integration, information retrieval, content management, architecture engineering and construction [1].

Ontology provides the following:
1. Semantically rich axiomatization of domain knowledge
2. An ontology captures domain knowledge
3. Reasoning about domain knowledge

Ontologies are divided into three main categories:
1. Top level ontologies are the fundamental ontologies and describe the abstract and general concept of domain specific knowledge
2. Domain ontologies and task ontologies represent knowledge within a specific domain of discourse, for example medicine or geography, or the knowledge about a particular task, such as diagnosing or configuring.
3. Application ontologies provide the specific vocabulary required to describe a certain task enactment in a particular application.
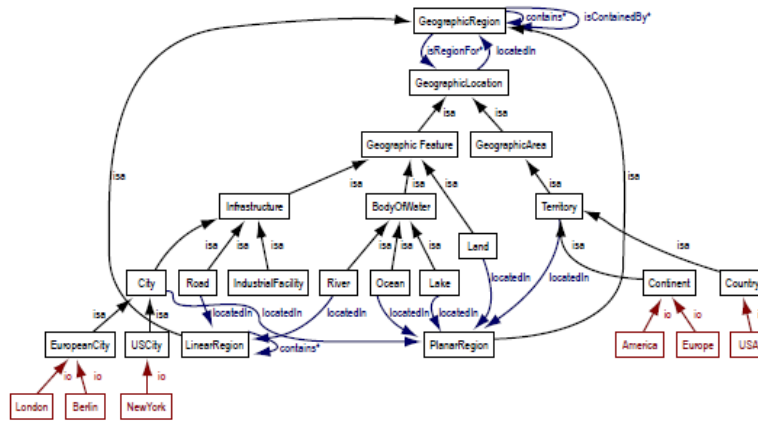
Figure 2. Knowledge Representation using ontology [2]

2.2. Description Logic

Description logic (DL) represents the knowledge in coherence with KR languages. DL is more expressive than propositional logic stated by Enrico (2004) [3]. In AI and KR, it could be put to utilization for reasoning for a particular domain and is mainly used for medical knowledge mentioned by Baader *et al,* (2003)[4] . The main difference between First order logic (FOL) and DL are: FOL represents class/objects and their properties (predicates) whereas DL represents the concepts and its role. DL consists of a Terminological box (T Box) and Assertional box. Terminological box represents the concepts and concept hierarchy and Assertional box (A Box) represents the relationship among concepts. As T Box represents the concepts, hence the complexity of the T Box can very much affect the performance of decision making in DL. YAK, KRIS, CRACK and KODIAK are description logic based systems [5].

2.3. Bayesian network

A Bayesian network represents and reason about an uncertain knowledge. The set of variables, X = X_1, X_2, X_3.......X_n, are represented by nodes and directed arcs/links connects the nodes. The connection between nodes represents the dependencies between variables shown in Figure 3. Conditional probability represents the strength of association between variables. The reasoning in Bayesian network can be diagnostic reasoning which is reasoning from symptoms to cause and predictive reasoning that is reasoning about causes to the new [6].
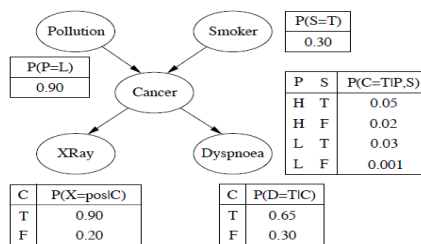

Figure 3. Bayesian network for domain lung cancer [7].

The nodes in Bayesian network include the following.
1. Propositions are represented by the Boolean nodes that can have binary value (True or False).
2. A node can have ordered value like speed (slow, medium, high).
3. A node can have integer values, for example age of a teenager may have possible value from 1 to 18.

2.4. Fuzzy Logic

In 1965, Lotfi Zadeh introduced the multi-valued logic, i.e. fuzzy logic, which extended the range of truth values to all real numbers in the interval between 0 and 1 whereas in case of crisp set the truth values are

defined as either 0 or 1. For example, the possibility that the sun is shining when there are some clouds in the sky might have a set to a value of 0.7. It is likely that the sun is shining.

Basically, fuzzy logic is the way to represent the expert knowledge that uses vague and ambiguous terms. Fuzzy logic is a set of mathematical principles for KR based on the theory of fuzzy sets, sets that calibrate vagueness and degrees of membership. Generally, the membership functions used to represent a fuzzy set are Sigmoid, Gaussian and pi presented by Zadeh (1965).

Fuzzy set theory in comparison with first order logic resembled the human reasoning in its use of approximate/estimated information and uncertainty to generate decisions. So, to depict the knowledge in a more precise way, fuzzy logic is the way which is designed to mathematically represent uncertainty and vagueness and it provide formalized tools for dealing with the imprecision built-in to many problems. Since knowledge can be expressed in a more natural way by using fuzzy sets, many business, Industrial, engineering and decision problems can be greatly simplified. In the fuzzy theory, fuzzy set A of universe X is defined by function $\mu_A(x)$ called the membership function of set A described by Zadeh (1965)

$$\mu_A(x): X \rightarrow \{0, 1\},$$

Where $\mu_A(x) = 1$ if x is totally in A;
$\mu_A(x) = 0$ if x is not in A;
$0 < \mu_A(x) < 1$ if x is partly in A.

Here, for any element x of universe X, membership, function µA (x) equals the degree to which x is an element of the set A. This degree, a value between 0 and 1, is used to represent the degree of membership, also called membership value, of element x in set A. The knowledge using fuzzy logic is represented by the production system by Esragh and Mamdani (1981)[8].

2.5. Semantic Web

KR could represent knowledge on web with semantic web technology. More specifically, it is the web of data. It is not a separate web whereas it is an extension of the current one, with well-defined meaning of information which enabled computers and people to work in cooperation presented by Heflin (2001). Original web interchanges the data, whereas semantic web records the data so that it relates to the real world objects. Semantic web depends on the ability to associate formal meaning with contents. KR is a medium to design the semantic web language that provides meaning for the data on the web. In case of semantic web, KR helps in data integration, management, conceptualization and retrieval. As shown in Figure 4.
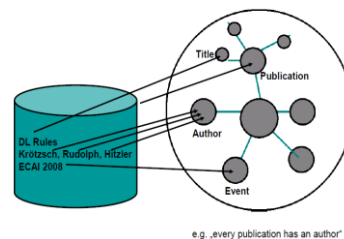

Figure 4. Knowledge to be represented using semantic web[9].

2.6. Conceptual Graph (CG)

In 1992, Sowa presented a graph known as a conceptual graph (CG). CG is a graph of logic based on the semantic networks. In addition to concept node the relations between concepts are also represented by the nodes known as relation node presented by Sowa (1992). The Concept can be concrete like human beings, animals, places, etc. or abstract like Feelings, senses, etc. CG can be finite, connected, bipartite graphs shown in Figure 5. Single CG can be corresponding to single proposition. Each node in CG must be unique and labeled by the type which is used to represent a class / object. All complex sentences were translated into small primitives before representation and then the each primitive CG joined to form the CG of the whole complex sentence mentioned by Guy *et al* (1993) mentioned that CG can be used for knowledge acquisition, reasoning, information retrieval, etc.
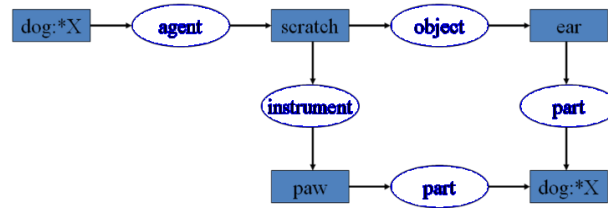
Figure 5. Concept graph representation of a sentence (Sowa, 1992) [8]

2.7. Hybrid knowledge representation techniques

Every KR techniques have their own demerits and merits, depending upon which type of knowledge user requires for representation. For adequate representation, the user needs different types of structures. To navigate the problem associated with the single KR technique, the hybrid KR has evolved. This section presents existing hybrid KR techniques.

2.7.1. Krypton

Ronald, Brachman and Richard (1983) developed Krypton, a HKR system in which KR was separated by two sections / two boxes called terminological box (T box) and assertion box (A box) mentioned by Ronald J. *et al* (1983) [10]-[12]. The TBox has the structure of KL-ONE in which terms are organized taxonomically, using frames an ABox has used the first-order logic sentences for those predicates which come from the TBox and a symbol Table maintaining the names of the TBox terms so that a user can refer to them. It is like a Tell-Ask module (Saffiottiand and Sebastiani, 1998). All interactions between a user and a Krypton knowledge base was mediated by TELL and ASK operations, Jacques and Indra (1994), shown in Figure 6. It was extensively used for representing the declarative knowledge [10], [11], [13].
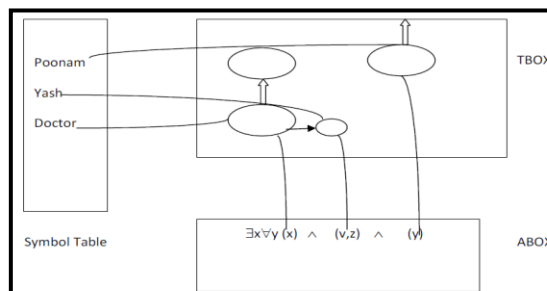

Figure 6. Overview of Krypton [10].

2.7.2. OBLOG 2

In 1987, Thomas F. Gordon proposed the idea of Oblog 2. The Oblog stands for Object-oriented Logic and is an experimental HKR and reasoning system. Oblog 2 is a hybrid of a terminological reasoning with a Prolog inference mechanism. The description of type and attribute taxonomies is supported by terminological component, whereas entities are instances of a set of types. Horn clause rules act as procedures, for determining the values of attributes are indexed by its type presented by Thomas (1987).

2.7.3. LOOM

Gregor (1987) implemented the LOOM HKR also belongs to the KL-ONE family was an intelligent system. The Loom is an application independent HKR system and is a classification based KR system. It is a frame based system and all the statements in LOOM were mapped into the predicate logic, Fikes and Kehler (1985). Loom provided an intelligent environment for the problem domain. Loom represents the declarative knowledge consisting of rules, facts, and default rules. A powerful reasoning mechanism is embedded by the system known as deductive engine / classifier presented by MacGregor and Robert [14]. Classifier in LOOM was based on the unification, forward chaining and object oriented concepts.

The capabilities of LOOM are:
1. Well defined semantics for the given language.
2. Expressive in nature
3. Contexts assertion (multiple images can be accessed at the same time.)

2.7.4. FRORL

Frame-and-Rule Oriented Requirement specification Language (FRORL) was developed by Jeffrey J. P. Tsai, Thomas Weigert and Hung-Chin Jang in 1992. FRORL is based on the concepts of frames and production rules which is designed for software requirement and specification analysis. There are two types of frames: object frame and activity frame. Object frames represent the real world entity not limited to physical entities. Frames in FRORL behave like a data structure. Each activity in FRORL is represented by activity frame to represent the changes in the world. Activity, precondition and action are reserved word not required in the specification. FRORL consist of Horn clause of predicate logic mentioned by Jeffrey *et al* (1998)[15].

2.7.5. RT-FRORL

RT-FRORL is the extension of FRORL proposed by Jeffrey J.P. Tsai, Mikio Aoyama, and Y.L. Chang in 1988. RT-FRORL inherits FRORL's basic structure, but also includes those language constructs needed to support the specification of real-time systems. The syntax of RT-FRORL is based on frames and production rules as in FRORL but the first order logic and temporal logic defines the semantic of RT-FRORL. RT-FRORL easily specifies the concurrent and absolute time properties of real time systems. A requirements model specified in RT-FRORL consists of two frame types: objects and activities. Each real world entity is modeled as an object. Changes taking place in the world are represented in the requirements model as activities. Each object and activity has certain properties, assumptions or constraints associated with it that is integrated into a frame representation. The syntax for an object frame is shown below (Jeffrey *et al,* 1991), [15][16].

Table 1 Frame representation

Object:	Object-name
Abstract relation:	Parent name
Attribute -name-1:	Value-l
Attribute -name-1:	Value-2
Attribute name-1:	Value-n

Each slot in the object frame corresponds to a particular property of the object and has a value set associated with it from which it can draw a value. The syntax for an activity frame is shown in Figure 2.

Table 2. Frame objects and their properties

Activity:	Activity _name(at tr _value-l ,.. . ,attr -value-n)
Abstract-relation:	Parent name
Part:	Attr-value-1: value-1 at t r _value -n : value _n
Preconditions:	Conjunction or disjunction of various activities or facts
Actions:	Sequence of activities or facts
alt_actions:	Sequence of activities or facts

In the above mentioned Table 1 and 2, the part, slot describes the objects participating in the activity, or their attributes. The precondition slot holds constraints to be satisfied prior to the execution of the actions in the next slot. If the preconditions are not satisfied, the actions outlined in the alt_actions slot will be performed. The RT-FRORL is developed to express the real time constraints. They illustrated examples on how RT-FRORL could be used to specify such systems. The temporal logic foundations used to reason about RT-FRORL were presented and typical assertions relevant to real-time systems were discussed by Jeffrey *et al* (1991)[16].

2.6. MANTRA

(MANTRA) stands for Modular Assertional, Semantic Network and Terminological Representation Approach. This work consists of four different KR techniques in designing a hybrid system integrating first order logic, terminological languages, semantic networks and production systems. The system architecture consists of three levels: (i) The Epistemological level, where the semantic primitives of the representation are defined, (ii) The Logical level, where the knowledge base management, the inference procedures and the interactions between the different epistemological methods are defined, and (iii) The Heuristic level, where ad hoc strategies could be introduced to improve the efficiency of execution by Calmet *et al* in1991 [17]. The semantics of the system has been formally defined using a model-theoretic formalism. The main contribution of this work is: (i) A methodology to define the semantics of KR methods. (ii) The integration in a multi-level architecture, of first-order logic, terminological reasoning, inheritance with exceptions and heuristic programming. The system has been implemented in Kyoto Common Lisp (KCL), a complete implementation of the standard Common Lisp, together with an object oriented extension called Common ORBIT stated by

Steele (1984) and Smedt (1987). The use of an object oriented programming tool increases the modularity of the system and makes it easy to modify. The interface has been developed using KYACC-KLEX, an interface between KCL and the compiler constructor YACCand LEX environment (Vigouroux, 1988), (Johnson, 1987), (Lesk, 1975) [18],[19]. The use of this environment makes the interface very easy to modify and to adapt to the changes in the syntax of the representation language described by Schneider (1986) [19].

The system presents two interface languages: one interactive interface based on menus and a programming interface allowing the use of the system primitives inside Lisp programs. The programming interface facilitates the integration of the system with other systems written in Common Lisp and to develop the not-yet-implemented a heuristic level of the architecture.

To facilitate the interconnection between the different methods a single data abstraction has been adopted mentioned by Rajeswari and Prasad (2012). This data abstraction consists of a set of directed graphs. Directed graphs subsume several of the most commonly used data structures and are also suitable to be used in an interactive system due to their inherent graphical character. The system Grasp, a graph manipulation package, has been adopted as the programming tool implementing this data abstraction stated by Bittencourt (1984). The base of all inference procedures implemented into the system is the unification function. A special unification package has been implemented in Common Lisp. The adopted algorithm is the almost linear algorithm mentioned by Martelli and Montanari (1982), Calmet *et al* (1991a and 1991b) [16] [20].

2.7. SOL

SOL (Smart Object Language) defines the smart object model for the design of complex knowledge based systems described by William et al (1995) [21]. The paradigm provides mechanisms for both HKR and multiple inference strategies.

Central to the paradigm is the concept of smart objects, engineered artifacts that combine a high level object structure with a rule based lower level language. The concept of smart objects was developed using criteria that evolved in the development of a large, complex knowledge based system. They present these criteria as desirable characteristics for KR s in general and use them to evaluate traditional knowledge and smart objects. An overview of a prototype knowledge base system (KBS) was implemented using the smart object paradigm and making its benefit concrete. In their work, they presented the overview of object based and smart object base KR and presented the hybrid KR technique with smart object and rule based system.

Smart objects are a tool for building systems in which inferential processes are an integral part of a broader design. Smart objects have an internal structure that partitions the knowledge contained in them and determines the interactions, the behavioral characteristics of a system of objects. The four elements of a smart object are a) Method, b) Interface, c) Attribute and d) Monitor. The smart object paradigm is successful in meeting the design criteria proposed for KRs used in the design of complex KBS'. It structures and captures domain and application knowledge in the form of smart objects: encapsulations of data and states enhanced with a production system like Rule Language. Smart objects have been conceptualized with an internal structure to aid in the design of complex KBS. The structure explicitly provides for meta-control of both the reasoning strategy and the structure and control flow of the implemented system.

The central methodological concept of the paradigm is the division of knowledge of a modeled environment into a domain component and an application component. The domain component is a potentially reusable base of knowledge that is common to a class of environments or problems. The domain component could represent behavior and structure common to all nuclear power plants governed by common Nuclear Regulatory Commission procedures and manufacturing assembly lines using common workflow architecture. The application component is the knowledge specific to an instance of the class.

SOL provides the reusability, up-to-date Figures, multiple inferencing, Mets-control, knowledge partitioning described by Kuechler and Vaishnavi (1995) and William *et al* (1995) [21].

2.8. AAANTS

Chirminda *et al* (2003) developed a hybrid KR system called AAANTS (Adaptive, Autonomous, Agent colony interactions with Network Transparent Services). The AAANTS model was a multi-agent system that conceptualizes and implements a colony of agents that actively interact with a collection of distributed services in order to give adaptive behavior. AAANTS was modeled on an intelligent environment related projects and built a prototype for an intelligent room that actively adapted the environmental conditions based on user behavior patterns. The KR was based on Frames that harmonize with continuous adaptation based on Reinforcement Learning techniques described by Singhe, *et al* (2003)[22]. AAANTS KR methodology has combined frame-based Uniframe and Accumulators to complement the learning achieved through Reinforcement Learning (LR) techniques. Agents keep frames representing the different states of activation. Each state relates to a value function that indicates the expected future rewards that initiates from this state. A correct mapping of a state signal from the environment would trigger an action of the highest expected reward.

AAANTS is a multi-agent system where each colony is exhibited, as a group of heterogeneous agents distinguished by their differences in ontology, behavior, knowledge and goals.

AAANTS is a general-purpose hybrid agent model that has the capability to interact with ubiquitous services embedded in the environment. AAANTS model has shown remarkable improvements over other functional monolithic gent models in terms of adaptability and knowledge component reusability. The core implementation is based on a component based distributed framework. The agent components in the AAANTS model are designed to interact with information sources from heterogeneous domains. They model the information sources as heterogeneous services that actively interface with the core implementation with the help of message based communication middleware.

In the AAANTS architecture, an adaptation layer acts as the sole communication medium in a typical insect colony such as the ants. The implementation has proven that the use of the adaptation layer functionality for interfacing has helped to overcome the conflicts faced with communication between agents and services. It was apparent from the implementation that the adaptation layer excludes the need for brokering and matchmaking services that are present in traditional deliberative architectures.

The AAANTS framework implementation is a distributed component based model that facilitates the well-being of a myriad of agent components. The framework was successful in providing services such as life-cycle management, agent reproduction, colony evolution, fault tolerance, load balancing and mobility to the agent components. Another advantage of the framework is the separation of common and redundant functionality of agent components to a single layer for common usage. This AAANTS model has succeeded in distributing knowledge and linear sequencing actions within an episode among several agents. The user ability to reward an individual action within an episode has enabled to properly adjust the value function of a state so that the sequence of actions adapts to the optimal pattern over a period. AAANTS model helped them to observe emergent behavior similar to that of a natural ant colony. These agents sense the environment and communicate with others using primitive message constructs to offer emergent adaptive behavior as a community (Christof, 1991)[23].

2.9. Extended Semantic Network

Extended Semantic Network (ESN) by Reena et al (2006) was the proposed HKR model to overcome the difficulties faced in the application area of information recovery and categorization in the current era of information overflow. The ESN prototype is a new proposed method for KR for easy ontology construction which can be employed in new generation search algorithm to facilitate information management, retrieval and sharing. This prototype enables easy construction of conceptual networks. Unlike NLP techniques, no heavy computations are required.

In order to develop new networks, a need for a set of documents related to that particular topic is required. These documents are needed to be for that only input into the proximal network (PN) program. Then automatically a network of nodes has been developed called the word network. This network contains all the different words that can be found in the input data related to the domain. Thus forming a recall process network, the network is then combined with the semantic network and the structure got restricted to 50 nodes.

Semantic network is essentially the precision network where the nodes are placed in the network with the help of expert knowledge. For constructing the extended semantic network they extend the precision model by adding nodes from the recall model at all possible and required positions. ESN was the hybrid of two networks to provide the extended semantic network.

1. First, is a PN model which involves three phases of processing, first the pre-treatment process where the documents related to the domain are analyzed in two stages and an output of word document matrix is obtained. This matrix is then passed on to the intermediate process and is analyzed by the data mining and clustering algorithms, namely k-means clustering, principal component analysis and word association obtain an output of word pair matrix with a value between each word pair. This value is the proximity between the word pair in the projected space depending on their occurrence in the contents of the documents processed. This data are further subjected to post-treatment process where partial stemming is carried on the word pair matrix depending on the case based requirement.

2. Secondly the semantic network which is based on the KL-ONE model, with the domain being the center of the network, which is expatiated by the domain components which in turn define concepts using the instance and inheritance relations. They follow the scheme process where minimum required information on a domain is precisely represented using the semantic relations they defined. The model is built based on the same set of documents used in PN and the 50 most important concepts is chosen with the help of a domain expert and is put into the semantic model. Each relational link used, namely the compositional, instantiation and inheritance links are given a predefined unit during calculation. This model is then stored and can be visualized using graph editor. The objective is to introduce the semantic based relation into our mathematically modeled PN. The network thus developed is then analyzed and merged to obtain one single semantic network for that domain. This process is repeated on different lists of concepts concerning to

various domains to obtain one network for each domain. The result obtained is fed into the MySQL database along with the relational links they share. This network is then readily combined with the PN, thus forming the extended semantic network that mainly works on the lines of precision (Reena *et al,*1987)[23][24].

3.   COGNITIVE HYBRID SENTENCE MODELING AND ANALYZER (CHSMA)

As per the 2005 Indian Human Development Survey, the surveyed households reported that the majority of the men as close to 72 % did not speak English, 28 % spoke limited English and 5% spoke good English. Considering the large population having limited knowledge or very less knowledge of English, it is important to provide a platform which can help them to understand the English language easily with the help of graphical presentation so as to visualize the representation with limited English knowledge.

The existing systems are domain specific and mainly extract the knowledge from standard KB like KL ONE, KL TWO, and LOOM whereas few HKR systems work with symbols. The main aim of this effort was to develop such resources that can allow the user to visualize the knowledge for a given input immediately. CHSMA represent the English sentences graphically. CHSMA is not only used to represent the concept, but also the events with step-by-step procedures visualized.

Following are the key factors for considering semantic net and script for Hybrid Knowledge Representation (HKR):
1.   Semantic net is the proven KR technique for various applications (KL-ONE family, Krypton, semantic web), for representing declarative knowledge graphically.
2.   Script is used to represent the events in a systematic way.

Initially, the KR /HKR system was implemented in Lisp and Prolog language. Prolog and Lisp provided the interface which were not interactive and user had to be familiar with the programming language to use the system. These languages were slow as compared to the Object Oriented Programming (OOP) languages and other traditional languages currently available.

This section of paper presents the design and methodology of CHSMA. CHSMA acts as an intermediate between the computer and user who is not at ease with the English language.

The system takes input as an English sentence/ paragraph/story and then translates it into a hierarchical graphical representation. The user is allowed to save the representation so that next time all the steps of the processes need not be executed again. The users are also allowed to print the representation. For longer stories, the users are also allowed to input query in an easy and simple way. The system was capable of updating the knowledge base and database at the time of execution just going through simple commands.

3.2. Mathematical Model of Chsma

Mathematically the CHSMA architecture / Model could be represented by the five Tuples {N, P, U, A, R} here N represents set of Entity nodes and concept nodes, P was the parent node, A was the association rule applied between nodes and R is the relation between nodes.
1.   N is a set of general class /Object/attribute < class/object and concepts>
Example of class <Person / Poonam >
Whereas concepts can be an event, state, behavior, etc and example of them is
Ringing, solid and good.
2.   P was a parent node used to like the semantic net with script representation as shown in Figure 5.7 Poonam and baby are the parent nodes. P $\subseteq$ N. i.e. parent node was the subset of N and it can be a single node or multiple.
3.   'U' was the universal set that consists both N and P.
4.   A is a set of association, as a composite sentence/ composite preposition can be formed with the help of association / connectors. Three different types of association were defined to form the composite network of a given sentence.
5.   R was the relations between the nodes can be represented by a labeled arc for example for the statement 'Poonam is a girl' than the relation was represented by the arc with the label is a.

Each node in CHSMA must be unique as used by Ali, (1994). A node could be a Class /Object and it could be a concept. A node in CHSMA is a labeled node as in case of Semantic network and conceptual graph. The label of the node is the name of the node and each node has been processed /used further by that name only. The name can be the particular value like 'Poonam' or it can be variable like 'x'. In case of a node with label x substitution mechanism has been used to substitute the exact value of that node.

Each node in CHSMA must be unique. For example the word "Poonam" occurred n times in a given paragraph / story than instead of n nodes corresponding to word Poonam the structure consist a single node with label Poonam i.e. the node representation in CHSMA must follow the property of uniqueness.

3.2.1 Substitution in CHSMA

Inheritance in the semantic net binds the general entity / concept about the problem domain to a specific entity/ concept. An example is illustrated in Figure 7.
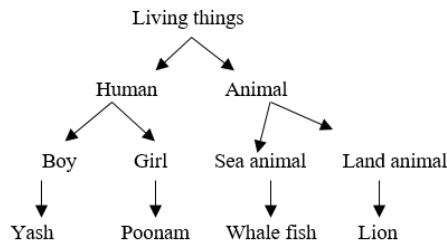

Figure 7. Substitution from General Class to Specific Class/Object.

Following substitution rules were applied for inference and reasoning:

Rule 1 *Direct substitution:* If n_1 and $n_2 \in N$ and where $n_1 \in$ general class and $n_2 \in$ specific class than apply value/n_2. For example, Poonam/ n_2 i.e. assign a value Poonam to n_2.

Rule 2 *Indirect substitution*: If $n_1, n_2, n_3, n_4, ..., n_n \in$ general class and $m_1, m_2, m_3, m_4, ..., m_n \in$ specific class, where $n_1, n_2, ..., n_n$ and $m_1, m_2, ..., m_n \in N$ then apply the association rule and check the association $m_1/ n_1/ n_2/ n_3, ..., n_n$.

3.2.2. Substitution procedure

Substitution in CHSMA was used to unify the value of each variable which can be a concept/ class. Substitution worked recursively generating the possible substitution for all the variable.

Substitution (m, n) in ({N, P, U, A, R}

1 n must have m

2 reduction of (m, n) where m and n $\in$ N / P and there exist a relation r1,r2, r3 from n to m in the structure S then.

For all r in S do the following

1.   If m = m than return m.
2.   If m≠n then

Follow associations $\forall a$ in set A and relation $\forall r$ in set R between n to m i.e. n1 to n2 to n3 and so on till nm. and returns m.

3.2.3. Information processing in CHSMA

Let P and S be the paragraph and story respectively chosen for representation which consists n number of sentences (s) then each sentence in the paragraph /story was processed according to the equation given below.

$$(W, P) = (W_1 s, t + W_2 s, t + \text{-------} W_i s, t) \qquad (1)$$
$$n$$

$$= \sum_{i=1} W s, t. \qquad (2)$$

Where n is the total number of sentences in P and $1 \le i \le n$.

Matching technique was applied during each sentence processing which takes a pair of nodes (j, k) those were associated/ related to each other and return the substitution for (j_i, k_i) for sentence.

4. CONCLUSIONS

It has been found that many HKR systems were developed in different domains and used for different applications. So it made sense to take the advantage of technology to serve the people so that they can learn and explore the world. The cognitive hybrid sentence modeling and analyzer (CHSMA) for the English language successfully processed the sentences, paragraphs and stories and recognized them with the average

---

accuracy of about 80 percent. To increase the accuracy of the system, soft computing approach could be considered.

REFERENCES

[1] Costa Ruben, 2014. Knowledge representations with ontology support for collaborative engineering in architecture engineering and construction, *ITcon*, Volume. 19, pp. 434-461, available at http://www.itcon.org/2014/26.

[2] Steven J. Fenves, Mary, Lou Maher, Duvvuru Sriram and Penas Anselmo, 2007. Information Retrieval Baselines for the ResPubliQA Task. In Cross Language Evaluation Forum, 2007. Article available at file:///C:/Users/Birbal/Downloads/bse-pe-003_1985_9_J-29_a_004_d.pdf.(last visited 14 July 2014).

[3] Brewster Christopher, Hara Kieron, Fuller Steve, Wilks Yorick, Franconi Enrico, Musen Mark, A., Ellman Jeremy, Buckingham Shum and Simon, 2004. Knowledge Representation with Ontologies: The Present and Future, *IEEE Intelligent Systems*, ISSN 1541-1672.

[4] Baader Franz, Calvanese Diego, Deborah L. McGuinness, Nardi Daniele and Peter F.Patel-Schneider, 2003. The Description Logic Handbook: Theory, Implementation, and Applications, *Cambridge University Press*.

[5] Acar Erman, 2014, Computing Subjective Expected Utility using Probabilistic Description Logics, Frontiers in Artificial Intelligence and Applications STAIRS 2014,pp 21-30. Available at https://ub-madoc.bib.uni-mannheim.de/36878/1/AcarSTAIRS14.pdf.

[6] Zhang N.L. and Guo H, 2006, chapter on Introduction to Bayesian Networks, *Science Press*, Beijing.

[7] Kevin B. Korb and Ann E.Nicholson, 2010. Chapter on Introducing Bayesian Networks, from Bayesian Artificial Intelligence, Second Edition. [Online] Available at www.csse.monash.edu.au/bai/book/BAI_Chapter2.pdf.

[8] John F. Sowa, 1992. Encyclopedia of Artificial Intelligence, Wiley, 2nd edition.

[9] Hitzler, 2011, Presentation on Introduction to Semantic Web, Available at http://www.semantic-web-book.org/w/images/8/87/W2011-01-introduction.pdf, [Accessed 7 June 2014].

[10] Ronald J. Brachman, Richard E. Fikes and Hector J. Levesque, 1983. KRYPTON: Integrating Terminology and Assertion, [Online] Available at: http://www.aaaipress.org/Papers/AAAI/1983/AAAI83-005.pdf [accessed 28 May 2014].

[11] James G. Schmolze, Thomas A. Lipkis, 1983. *Classification in the KL-ONE Knowledge Representation System*. Proceedings of the Eighth international joint conference on Artificial intelligence (IJCAI'83), Volume 1, Pages 330-332, source IEEE Xplore.

[12] Ronald J. Brachman and James G. Schmolze, 1985. An overview of the KL-ONE Knowledge Representation System, Cognitive Science, Volume 9, Issue 2, *Elsevier Inc*., April-June 1985, pp 171-216.

[13] Tanwar Poonam, Prasad, T. V. and Dutta Kamlesh, 2012. Hybrid technique for effective knowledge representation & a comparative study, *International Journal of Computer Science & Engineering Survey*, Volume 3, No.4, August 2012, pp 43-57.

[14] MacGregor and Robert, 1991. Using a description classifier to enhance knowledge representation, source IEEE Explore.

[15] Jeffrey J.P. Tsai, Mikio Aoyama and Change, Y.L., 1998. Rapid Prototyping Using FRORL Language, Department of Electrical Engineering and Computer Science, University of Illinois at Chicago, Chicago, source IEEE Xplore.

[16] Jeffrey J.P. Tsai, Hung-C'hinJaiig and Karen J. Schellinger, 1991. RT-FRORL: A Formal Requirements Specification Language for Specifying Real-Time Systems, Dept. of Electrical Engg. and Computer Science, University of Illinois, Chicago, source IEEE Xplore

[17] Calmet, J., Tjandra I.A and Bittencourt, G., 1991b. MANTRA: A shell for hybrid knowledge representation. 3rd International Conference on Knowledge Representation with Ontologies: The present and future, November 10-13, 1991.

[18] Johnson, L. and Johnson, N.E., 1987. Knowledge Elicitation Involving Teachback Interviewing in Kidd, Knowledge Acquisition for Expert Systems: A Practical Handbook, 1987. [Online] Available at: http://ksi.cpsc.ucalgary.ca/articles/KBS/KSS. [Accessed 18 May 2014].

[19] Patel-Schneider P.F, 1986. A four-Valued Semantics for Frame-Based Description Languages, Proceeding of American Association for Artificial Intelligence -86, 1986.

[20] Jacques Calmet and Indra A.Tjandra, 1994. Buildinga bridges between knowledge representation and algebraic specification, Methodologies for Intelligent Systems, Lecture Notes in Computer Science, Volume 869, 1994, [Online] Available at: http://www.springerlink.com/content/70051573m2513005 [Accessed 8 June 2014].

[21] William L. Kuechler, Jr.Nina Lim and Vaishnavi Vijay K., 1995. *A Smart Object Approach To Hybrid Knowledge Representation and Reasoning Strategies*. Proceeding of the 28th Annual Hawaii International Conference on System Sciences.

[22] Singhe, R.A., RanaChirminda and Ajilh, P.MadurApperuma, 2003. Enhanced Frame-based Knowledge Representation for an Intelligent Environment, *IEEE, KIMAS* 2003 Boston, USA.

[23] Shetty Reena T. N., Pierre-Michel Riccio and Joël Quinqueton, 2006. *Hybrid Model for knowledge representation*, ICHIT '06 Proceeding of the Inernational Conference on Hybrid Information Technology, Volume 01, 2006, pp 355 – 361.

[24] Shetty Reena T.N., Pierre Michel Riccio and Joel Quinqueton, 1987. Extended Semantic, Network for Knowledge Representation, Information Reuse and Integration, France, *IEEE-IRI*, vol1, 2009.