European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 4, 2020

A Study on Knowledge Representation Models

S. Praveena Rachel Kamala[1] & Dr. S. Justus[2]
School of Computer Science and Engineering, VIT University, Chennai, India.
Email: praveena.rachel2013@vit.ac.in [1], justus.s@vit.ac.in [2]

Abstract:
Knowledge Representation has become the primary requirement for the budding smart world. The process of Representing Knowledge includes capturing data, representing information and retrieving accurate knowledge. It is very much evident that only if the representation model is simple and clear, then the role of capturing and retrieval can be effectively performed. Recently, an urgent need for a more efficient knowledge representation model has increased to cater to modern culture. This research topic focuses on the representation models designed over the century. The main objective of this survey is to summarize the work that has been carried out in the area of knowledge representation models and to discuss a range of approaches and methods for representing knowledge.

Keywords: Conceptual Graphs, Knowledge Acquisition, Knowledge Representation, Knowledge Retrieval, Semantic Webs.

1. Introduction:

Traditionally, over the years, skills are transferred from one generation to another. These skills are simple or complex information gained by an individual through experience. When it is not transferred correctly, then this results as lack of knowledge or inadequate knowledge. The facts which are clearly documented and transferable are classified as Explicit Knowledge; the data that remains in hand and cannot be easily expressed are known as Tacit Knowledge. Knowledge is something that is not readily available but has to be constructed. Knowledge includes identification of concepts, naming along with their relationship, association with the previous history, and conclude with an outcome.
Knowledge is how an individual understands the concept. The various types of knowledge are declarative knowledge, procedural knowledge, heuristic knowledge, meta-knowledge and structural knowledge. Suppose an assignment of identifying siblings in a party is given, then the different possibilities for solving the problem are:

- Familiarity or similarity in their attitude, speech, color, resemblance and so on,
- Awareness is complementary when we have a pre-knowledge about the family,
- Associating and clustering the group using justification,
- Description of the character and features with their relatives,
- Experience through handling such similar situations and
- Systematic study of the genes and other biological classification results in learning.

Therefore, Knowledge refers to familiarity, awareness about something, justified truth, description, experience, or learning. This term ‘knowledge’ cannot be fetched directly but has to undergo a process of refinement for accuracy. This fine-tuning procedure is known as ‘Knowledge Representation’. Representation describes the group and its function is to capture the required features related to the problem domain. It mainly deals with the vocabulary, constrains, access and grammar.

The unprocessed data is picked up from the open environment, sorted out using conditions; a proper method for storing is employed, word by word screening is done and placed in an appropriate position for future fetching, this is the working principle of knowledge representation.

In the present era, getting raw data is not a big deal, but to get the perfect correct information is the challenge! To confront this problem in this research paper, we are presenting the already existing approaches, methods, and knowledge representation models.

2. Background of Knowledge Representation

Starting from the 16th to early 18th century, a machine named 'computer' performed the work of a simple calculator. In 1837, Charles Babbage proposed the Analytical Engine comprising of Arithmetic Logic Unit, and Memory with the flow of control; his work was completed by Henry Babbage in 1910 and was able to perform basic calculations [1]. This was the beginning of storing data and doing tricks with the available resource from memory.

In 1930s the invention of the electro-mechanical binary programmable computer (i.e.) first functional modern computer and Turing machine were proposed by Konrad Zuse and Alan Turning. In 1956, Allen Newell and Herbert Simon contributed to mathematics and computer science resulting in logic theory and Artificial Intelligence [2]. From this period, onwards, data was manipulated and sensible information was delivered. Expert systems and reasoning interpretations were drastically developing.

Different data structures were used for placing the data for easy fetching. Graphical or the network structure was widely preferred for its straightforward approach in inserting the data—several methods for representing knowledge were planned to cater to different applications. From raw data to meaning added information, yields to reasonable knowledge and intelligent system. This leads to the reign of a knowledge based system.

3. Different Approaches to Knowledge Representation

The word "Knowledge" sounds simple but differs in definition and forms from one instance to another. Such differences in approach are listed below:

In *Philosophy*, the study of knowledge is known as epistemology. This reflects on the history of theses and theories.

Linguistics is the study of knowledge system of Human language, which is a mixture of complex abilities like speech, emotions, thoughts, desires, expression and so on [3]. Phonetics, phonology, morphology, syntax, semantics and pragmatics are parts of linguistics.

System Analysis is a problem-solving process which includes several factors and concepts. The knowledge and methods of modern science and technology combine with concepts of goals, to find the essentials, and suitable reflection on a superior framework is planned to overcome uncertainties [4]. The eight major areas that cover the knowledge aspect of an institution are organizational structure, organizational behavior, organizational politics, functional requirements, systems development, information technology, individual behavior and corporate characteristics [5].

Artificial Intelligence includes knowledge representation and reasoning for placing information of the real-world scenario in the understandable computer format; for tackling the situation.

Connectionist Theories give the theory of mind implying the way human knowledge acquires and represents information.

Constructivist Theories argue the knowledge and meaning generated by humans with experience and ideas.

Educational Psychology is termed as awareness of something and understanding its importance, adapted by experience or learning.

Cognitive Science is an interdisciplinary study of science that brings psychology, linguistics, philosophy, computer modeling in rising theories about human thinking, perception and learning.

4. Methods of Knowledge Representation

The various methods for knowledge representation are as below;
List (hierarchical) scheme processes word by word. The sequence of action (knowledge) is stored in the repository (Knowledge Base) for performing the necessary action (reasoning). When new instances are encountered, the Knowledge Base gets updated.

Procedure (functions/control flow) scheme contains a set of instructions to be followed for solving the problem. This is a processed form of information.

Network (Trees /graphs) scheme use the nodes to represent the concepts or objects and the edges represent their relationships. These data structures make it more evident that humans make it easier to store and manipulate the resource. Decision Tree and Semantic Network (nodes & links) resolve the big problem by formulating concept networks for the intelligent task.

Structured (stereotype) schemes are extended network representation with a complex data structure for each node in a graph. Frames (slots) and Scripts contain additional procedures within each node in a network, which assists in obtaining the goal[17].

Rule-Based Representation (special problem-solving contexts) schemes contain inference rules and predictions. This method of representation is for solving the problem using a set of production rules. The condition – action pair is defined as If – then (situation – action) with the conditions as Initial State, Goal State, Legal Operators and Operator Restrictions.

Logic-Based Representation is a scheme using mathematical notations and symbols. The problem is expressed in terms of symbols and notations for understanding clearly by both computers as well as humans. Since the syntax and semantics are correctly fed, the output received is unambiguous. The various types for solving and reasoning are done using Facts & Premises, Rules of Propositional Logic (Boolean), Rules of Predicate Calculus (additional information), and Measures of Certainty (certainty factors) for Diagnosis, Expert Estimation, Statistical Data, Bayesian Probability, Fuzzy Logic and so on.

Ontology scheme is a domain-based. The representation is done using formal semantics, for explicitly accessible by machine, easily sharable, viewed conceptually and field specific[16].

Hybrid scheme is a combination of different methods.

5. Existing Knowledge Representation Techniques

Late 1940s, knowledge representation was based on words following syntax or a procedure. Edward Tolman proposed Cognitive Maps/ Mind Maps (1948) as a type of intellectual representation that provides a person to receive, interpret, remember, think over and reciprocate information about the location and use it [8]. It is data and connected data. Akers provided a graphical solution for the two-job m-machine job shop using Production Rules (1956) [9]. This was imagined as a sufficient form of representation for procedure knowledge. In a clearly scheduled job, production rules play a vital part in supporting the task. Semantic Pattern Matching (1967) refers to the representation of knowledge in patterns of interconnected nodes and arcs. This was developed for implementing artificial intelligence and machine translation in the computer system [7]. John F Sowa proposed Conceptual Graphs (1976) as a Graphical Representation of Information. Two types of nodes for storing concepts and relations and arcs to connect are deployed. The theory of conceptual graphs has four operations, such as copying, restricting, joining and simplifying. Alain Colmerauer and Philippe Roussel invented Prolog as an abbreviation for "PROgrammation en LOGique" [10] used the concept of First Order Logic (FOL) (1972) for representing the knowledge.

In the early 1980s, KR was based on concepts consequential semantic structure emerged. Gardner explored the noticeable differences between an individuals abilities and proposed theory of multiple intelligences. This gave rise to Concept Maps (1984) to help students learn how to learn by Novak and Gowin [11]. This provides structured data experts systems for Knowledge Based System. Judea Pearl proposed Bayesian Networks (1985) combining cognitive science, artificial intelligence and directed graphical model [12]. Semantic network (1987) captures knowledge as a graphical structure. The nodes denote concepts, objects, events, time and so on. The arcs represent the relationships like 'is a', 'is an instance of', 'has', 'part of', 'value', logical (and, or, not) and linguistic (likes, owns...) relations. Sowa proposed an extended work with object oriented concept on Semantic Net (1992). This lead to ontology based development in Artificial Neural Networks. Kosko coined Fuzzy Cognitive Maps (1994) as an expansion of cognitive maps for casual relationships between nodes and dynamic system feedback [13].
After a score of years, KR became narrative yielding to pragmatic models. Tim Berners-Lee suggested Semantic Web along with Hendler and Lassila (2001) as a common framework in a wired network allowing Information for Sharing and Reusing [14]. Ontology Web Language (OWL) created by OWL working group in 2004 to supply the needs of complex knowledge about things in ontology.

Table 1 – Overview of Knowledge Representation Models

Proposed by	**Year**	**Model**	**Description**
Gentzen	1935	Propositional Logic	Logical conclusion is derived from one or more propositions
Akers and Friedman	1955	Production Rules	First proposed the job shop problem [9]
Lotfi Zadeh	1965	Fuzzy Logic	Multi valued logic giving a variable between 0 and 1.
Quillian	1967	Semantic Net	To store information in computer memory [6].
John Henry Holland	1970	Genetic Algorithm	Used to find the heuristic solution for solving the problem.
Newell's & Simon's	1972	Procedural Knowledge (Hierarchical Representation)	Rule based systems; defines word by word procedural Knowledge model
Jerry Fodors	1975	Language of Thought Hypothesis (LOTH)	Thoughts are expressed as symbols and logics
Minsky	1975	Frames	Represents a stereotype situation mixes declarative knowledge and structured procedural knowledge.
Shank	1975	Scripts	Used to model networks
Shank	1982	Dynamic Memory Model	Explain higher aspects of cognition
John Sowa	1984	Conceptual Graphs	Connects concepts and their relations
Lipp	1984	Fuzzy Petri nets	Proposed for using fuzzy logic in place or transition network
Shastri	1988	Neural Networks	Widely used in Data Mining, Pattern Recognition, Manufacturing System,

			and Scheduling
Pachulski & al	2000	Knowledge Webs and Networks	Deals with mapping information with different concepts.
Tim Berners Lee	2001	Semantic Web	Provides common framework for reusing data
Patel-Schneider	2005	Description Logic	Describes definitions and rules for reasoning.
Gauch	2007	Ontology	Categorizing the group and their relationships.
Freebase	2007	MetaWeb	Overcoming the problem of XML and metadata
Calaris	2008	OpenCalais	Semantic metadata
MS	2009	Satori	Version of digital assistance
CMU	2010	NELL	Basic set of fundamental semantic properties
MSRA	2011	Probase	Entity-Relation Graph
Google (Freebase)	2012	Knowledge Graph	Information to knowledge and Graph Database
Facebook	2013	Graph Search	User's network of friends search engine
Google	2015	LPG (Labeled Property Graph)	KG with visualization capabilites
Cambridge Semantics	2017	AnzoGraph	Bridging the gap between graph properties using RDF
Stardog	2017	GraphQL	Supports in adding more databases
W3C	2019	Neo4j	Standardize the graph data

Over a decade, Knowledge Graph (2012) is a knowledge base for enhanced searching used by Google. Semantic search of Information gathers data from a vast source to Knowledge and Graph Databases [15]. The entities are represented along with their properties and linked with each other. Association of entities with similar properties are made simpler. Simialarly, Facebook has adopted graph search technology for asoociating the friends of different users. Thus enabling the user to locate their friends and establish new connections. Related properties are also obtained and based on their interest recommendations are also prossible [18].

GraphQL and Ontotext worked togather in reducing the complexity of representation of each entity in a graph structure. Due to dynamic change of information and since each element is represented as an entity along with its properties, it becomes more complex in putting as well as getting information. AnzoGraph developed by Cambridge Semantics is mainly developed for bridging such gap and making the process simpler by using RDF* [19].

W3C has proposed for standardizing the graph data in 2019, aiming in associating the graph community together like Neo4j, Oracle, Ontotext and ArangoDB. The focus on developing strong graph databases.

6. Discussions

Representing knowledge has been advancing over the century in various dimensions. The primary reason behind this growth is the rapid generation of data and information. From the comparative study of knowledge representation, it is evident that initially, representation was required for only a few words. Hence symbols were used for representing one or two things using logic.

In a while, rule-based representation came into existence. These models used data in a minimum quantity and gave results for a particular domain. They were domain-specific models restricted to perform only one or two different tasks.

Along with mining data and performing a small task, predicting knowledge about a situation using fuzzy logic came into reality. It was interlinking the data using the graph slowly popped into the knowledge

world. Various technologies based on graph demonstrated their excellence in the field. This influenced the viral growth of sharing information to the world around us.
Innovation grew up from word to paragraph representation and sharing the content to the society through the internet. Now the people are well-nourished with information and the data are available in surplus but interconnected knowledge about a scenario is scarcely presented.
In this era, the system is trying its best to put concept-based representation to overcome this problem of inadequate accessibility of knowledge about a concept. Ontology helps in combining a few similar Ontologies of the same domain interest. With all these advancements, the web world is marching towards the conceptualization of sharing knowledge.

7. Proposing Future Knowledge Representation Model

We have proposed a model for knowledge representation based on Description Language and Conceptual Graph, "Towards MORK using Description Language and Conceptual Graph".[20] This model is based on concept representation.
In this model, the captured data is placed in a CR graph denoting the concept and relation. The reasoning and retrieval of knowledge from the graph are done by using the description logic.
The advantages of MORK over other existing Models are

1. Easy logical entailment
2. Resulting in an accurate answer to the posted query
3. This is not word searching but logical node pulling
4. Principal followed is magnetic ball effect
5. In this related information can be fetched from the neighboring nodes
6. Similar to human thinking and reasoning

8. Conclusion

In this paper, we have studied the growth of knowledge representation model over the century. The approaches, methods and technologies used varied over the period to meet the requirement of the user. We have also given an outline of the techniques used in our model, which will overcome the problems of the present-day circumstances and will give a version that will satisfy the requests of the future generation.

9. Reference

[1] http://www.computerhope.com/issues/ch000984.htm
[2] https://en.wikipedia.org/wiki/Logic_Theorist
[3] http://linguistics.ucsc.edu/about/what-is-linguistics.html
[4] http://www.iiasa.ac.at/web/home/about/whatisiiasa/whatissystemsanalysis/what_is_systems_analysis.html
[5] Nicholas P. Vitalari, "Knowledge as a Basis for Expertise in Systems Analysis: An Empirical Study", MIS Quarterly, Vol. 9, No. 3, Sep., 1985, pp. 221-241.
[6] Allan M. Collins and M. Ross Quillian, "Retrieval Time from Semantic Memory", Journal of Verbal Learning and Verbal Behavior, 1969, Vol 8, pp 240 – 247.
[7] http://aitopics.org/publication/semantic-networks, http://aitopics.org/topic/semantic-networks
[8] Edward C. Tolman. "Congitive Maps in Rats and Men", in 'Image and Environment: Cognitive Mapping and Spatial Behavior', Chapter 2, pp27.
[9] Jacek Bta~ewicz, Wolfgang Domschke, Erwin Pesch, "The job shop scheduling problem: Conventional and new solution techniques", European Journal of Operational Research 93, 1996, 1-33.
[10] Alain Colmerauer1 and Philippe Roussel, "The birth of Prolog", November 1992.
[11] Joseph D. Novak, "Learning, Creating, and Using Knowledge: Concept maps as facilitative tools in schools and corporations", Journal of e-Learning and Knowledge Society, Vol. 6, n. 3, September 2010, pp. 21 - 30.
[12] Judea Pearl, Stuart Russell, "Bayesian Networks", Lecture notes, University of California.
[13] M Shamim Khan,"Fuzzy Cognitive Maps and Intelligent Decision Support – a Review"

1. Berners-Lee, Tim, James Hendler, Ora Lassila, "The Semantic Web". Scientific American Magazine". Retrieved March 26, 2008.
2. Singhal, Amit , "Introducing the Knowledge Graph: Things, Not Strings". Official Blog (of Google). Retrieved May 18, 2012.
3. Stephan Grimm, Pascal Hitzler, Andreas Abecker, "Knowledge Representation and Ontologies: Logic, Ontologies and Semantic Web Languages", Semantic Web Services, 2007, pp 37 – 87.
4. Carlos Ramirez and Benjamin Valdes, "A General Knowledge Representation Model of Concepts", www.intechopen.com.
5. Senthil K. Chandrasegaran et al. "The evolution, challenges, and future of knowledge representation in product design systems", Computer-Aided Design Vol 45, pp 204-228, 2013.
6. Meenakshi Malhotra and T R Gopalakrishnan Nair, "Evolution of Knowledge Representation and Retrieval Techniques", International Journal of Intelligent Systems and Applications, Vol 07, pp.8-28, 2015.
7. S. Praveena Rachel Kamala, S. Justus, "Towards MORK: Model for Representing Knowledge", International Journal of Modern Education and Computer Science(IJMECS), Vol.8, No.3, pp.45-53, 2016.DOI: 10.5815/ijmecs.2016.03.06