Global Scientific JOURNALS

GSJ: Volume 8, Issue 10, October 2020, Online: ISSN 2320-9186
www.globalscientificjournal.com

A NOVEL AGENT-BASED KNOWLEDGE DISCOVERY WITH A NEW EVALUATION FOR KNOWLEDGE QUALITY

Asmaa M. El-Saied

Dept. of communication and electronics engineering, Mansoura High Institute of Engineering and technology, Mansoura, Egypt.
asma.m.alsaid@gmail.com
amsaieed@mc.edu.eg

KeyWords
Intelligent Agents, Ontology, Q-Learning, Causal map, text extraction, Graph-based structure, Knowledge Exploration.

ABSTRACT
In the information technology era, there exists a huge amount of electronic data and information worldwide, therefore many great challenges arise. One of them is how to exploit these information and knowledge resources to turn them into useful knowledge available to concerned people, since the value of knowledge increases when people can share and capitalize on it. Thus approaches that can help researchers to benefit from existing hidden knowledge are needed and tools that can extract relevant and useful knowledge are required as well. By combining many information with many relations, more novel and useful information are found. Representing such combination requires a special structure (called information map). This paper introduces an efficient information map as a graph structure for diverse information and their relationships, called Oriented Directed Acyclic Graph (ODAG) algorithm. Moreover, formulating a new evaluation of the creative information, called Creative Information Quality (CIQ) algorithm by allowing agents to link diverse information dynamically. These diverse information are generated from an intelligent inference mechanism, which based on the principles of information associations; similarity, contrast and contiguity

Introduction

Due to the increasing availability of large amount of data in many fields of science, business environment and in many IT applications, most recent researches tend to use semantic knowledge. This semantic knowledge derived from domain ontologies to improve exploring new valuable knowledge (creative information) which hidden around data. Recently, there have become a large research community that shares the vision of sharing data and knowledge among humans and machines. So, the demands on delegation of many tasks to intelligent software agents are increasing [1]. Hence for any agent, to deal with a wide range of situations, it needs to have a sufficiently knowledge base and a flexible inference engine. So, efficient artificial intelligence techniques should be used in designing and implementing such agents [2, 3]. Basically, the creative information's generation process is a form of knowledge explored through experience, these acquired knowledge through experience is similar to the data mining concept (discovery of new knowledge through data analysis), but with intelligence that entails learning and creativity [4,5]. With regard to creativity, there are significant challenges to make agents capable of generating diverse creative informations dynamically. Furthermore, representing knowledge in a way that allow generating more valuable informations. These are achieved by using artificial intelligence learning technique, that based on semantic knowledge base (ontology and facts) and inference engine (reasoning) mechanism. The inference or reasoning process makes implicit semantics explicit, by deriving consequences from ontology [6]. The elementary theory of exploring new knowledge, is to integrate the learning process based on the three principles of information association (Similarity, Contiguity and Contrast)[7], with an efficient structure of representing the knowledge. Therefore, the fundamental goal of knowledge representation is to represent knowledge in a manner that facilitate the inferencing from knowledge, which based on the principles of information association as linking principle between informations[8]. These principles have an important behavior for generating diverse informations through the dynamic exchange of varied knowledge. By combining many informations with many relations, more novel and useful informations are found. Representing such combination requires a special structure (called informations map)[9,10].

This paper aims to introduce an efficient informations map as a graph structure for diverse informations and their relationships, called Oriented Directed Acyclic Graph (ODAG) algorithm. Moreover, formulating a new evaluation for the quality of the creative informations generated, called Creative Informations Quality (CIQ) algorithm by allowing agents to link diverse informations dynamically. These diverse informations generated from an intelligent inference mechanism based on the principles of information associations; similarity, contrast and contiguity. The remainder of this paper is organized as follow: Section П gives a brief overview on the related works in exploring new informations with relations, basic concepts and hypotheses was presented in section Ш, then the details about the proposed intelligent mechanism for exploring new creative informations with causal map is given in section IV with the experimental results in section V.

Related Works

The development of information technology takes giant leap in last decade. It provides fast, convenient powerful computing and communicating approaches which facilitate information gathering, sharing, analyzing and archiving. That is, used to generate new knowledge with creativity, where creativity is one of the key factors to explore new informations as knowledge [11]. Most informationtion research either implicitly or explicitly assumes Osborn's conjecture that if people generate more informations, then they will produce more good informations. Some studies have also reported that certain informationtion protocols can elevate both information quantity and information quality [12]. However, another work reported no relationship between information quality and information quantity [13]. Hence, most previous informationtion literatures were inconsistent in the arguments. Briggs and Reinig [14] provided a theoretical explanation (Bounded Informationtion Theory) to clarify the relationship between information quantity and information quality, and they recommended guidance for the development of informationtion techniques for improving the quality of informations. A good information was defined as one that is feasible to implement and would attain the goal. Reference [15] devised an automated decision agent called semantic informationtion learning agent (SILA), which constructed a map of generated informations that can learn and share their knowledge, that use an information generation protocol to construct a tree-like information map. Whereas, [7] focuses on the process of "linking" among distributed knowledge, clearing that linking principle has an extreme importance in creative thinking and problem solving. The outcomes of these researches pointed out that, the methodology of representing knowledge plays an important role in elevating both information quantity and information quality. There are many structures for representing knowledge tree-like, graph-like as a map, and so on. For instance, cognitive maps that follow personal construct theory, provides a basis for representing an individual's multiple perspectives. This theory has spawned many fields and has been used as a first step in generating cognitive maps. Huff [16] has identified five generic "families" of cognitive maps:

1. Maps that assess attention, association, and importance of concepts. With these maps, the map maker searches for frequent use of related concepts as indicators of the strategic emphasis of a particular decision maker or organization, for example, and looks for the association of these concepts with others to infer mental connection between important strategic themes. She also might make judgments about the complexity of these relationships or differences in the use of concepts.

2. Maps that show dimension of categories and cognitive taxonomies. Here, the map maker investigates more complex relationships among concepts. She might dichotomize concepts and construct hierarchical relationships among broad concepts and more spe-

cific subcategories. Maps of this type have been used to define the competitive environment and to explore the range and nature of choices perceived by decision makers in a given setting.

3. Maps that show influence, causality, and system dynamics (causal maps). These maps allow the map maker to focus on action; for example, how the respondent explains the current situation in terms of previous events and what changes she expects in the future. This kind of cognitive map is currently, has been, and is still the most popular mapping method.

4. Maps that show the structure of argument and conclusion. This type of map attempts to show the logic behind conclusions and decisions to act. Here, the map maker includes causal beliefs, but looks more broadly at the text as a whole to show the cumulative impact of varied evidence and the links between longer chains of reasoning.

5. Maps that specify frames and perceptual codes. This approach suggests that cognition is highly conditioned by previous experience and that experience is stored in memory as a set of structured expectations.

Causal maps lately have received much attention, where many researches had introduced various approaches to causal maps that based on simple inference mechanisms about the consequences of a causal maps. In the meanwhile, the definition of a precise semantic interpretation of qualitative causality has received very little attention. In [17] author used the Bayesian network approach to make inferences in causal maps. Thus, it used probability theory as is the case usually in artificial intelligence. However, its approach is applicable only in the acyclic case because circular relations or causal loops, common in causal maps, violate the acyclic graphical structure required in a Bayesian network.

Basic Concepts and Hypotheses

Basically, semantic knowledge (meaning, understanding) is a higher form of Information, which begins when facts and concepts (information) are connected, and structured as ontology, where the ontology is the formal description of a domain, using connected facts and concepts. So the Knowledge Base (KB) can be considered as a combination of ontologies and rules. In order to improve the knowledge base in the light of further information and facilitates learning from the experience of making mistakes, the knowledge base should be merged with inference procedures (for manipulating hidden knowledge in a knowledge base), if the content of the knowledge base is poor then the inferences will be correspondingly poor. Nevertheless it is vital to have a good inference engine to take full advantage of the knowledge base. Inference is based on forward chaining rules, that apply rules which cause more facts to be asserted until no more rules apply [1,6]. Knowledge-based systems are based on the methods and techniques of Artificial Intelligence. Their core components are the knowledge base and the inference mechanisms. All operations like modifications, enhancements and evaluations on the Knowledge Base (KB) , could be done on both object-level knowledge (knowledge about things) and meta-level knowledge (knowledge about knowledge) using intelligent agent as in fig(1) .

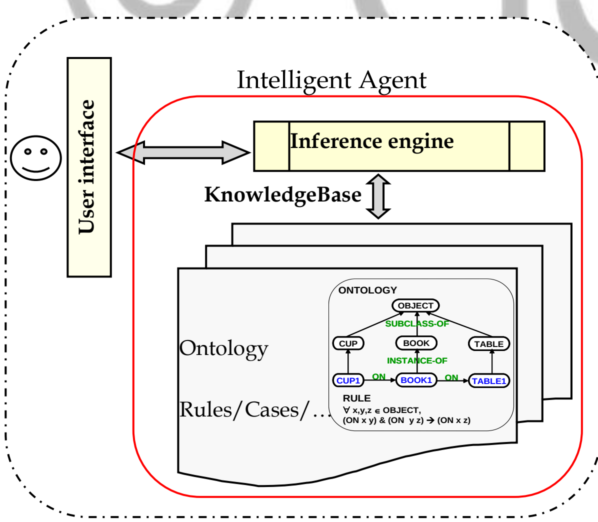


Fig (1): The traditional framework of intelligent agent to explore knowledge

In addition to employ reinforcement learning agents, which not only have a high reproduction capability of human-like behaviors (externally) but also human-like thinking (internally). Based on three fundamental human's association capabilities (similarity, contiguity, and contrast) during information generation [18], implementing these capabilities is done in an agent's inference mechanism. The three associations are accordingly introduced to the inference mechanism of informationtion agents in order to allow autonomous information generation. This can further enhance the number of informations generated by removing the limitations, as mentioned in the Bounded Informationtion Theory [14]. The basic architecture of an intelligent instructable agent, is working by informationtion protocol which is adopted to facilitate the progressive construction of informations map. In addition to the capability of generating informations. That is, could possibly be advanced in the following directions:

1. Learning capability for understanding the task. Moreover, the causality among the generated informations should be easy to characterize.
2. Continued adoption of additional external stimuli (KB).

So, this paper demonstrates how to construct an efficient ideas map as a graph structure for diverse ideas and their relationships, called *Oriented Directed Acyclic Graph (ODAG) algorithm.* Based on an automated decision agent called semantic ideation learning agent(SILA), in which the diverse ideas generated from this inference mechanism founded on the principles of idea associations; similarity, contrast and contiguity introduced in [15]. That considered as a novel ideation protocol which is adopted to facilitate the progressive construction of ideas map with elevating both idea quantity and idea quality. Moreover, formulating a new evaluation for the quality of the creative ideas generated, called Creative Ideas Quality (CIQ) algorithm by allowing agents to link diverse ideas dynamically.

An Intelligent Mechanism for Exploring new Creative information with Causal Map

The aim of this study is to provide an information protocol to facilitate the construction of information map through many rounds by intelligent agent. That elevates both information quantity and information quality to such an extent that unprecedented. This map provides an environment in which agents can learn and share their knowledge based on an inference mechanism for recommending the intelligent agent [15] to generate diverse informations that was adapted as in fig(2). The agent is equipped to experience learning capability by utilizing a reinforcement learning method based on Q-learning (design agent's inference engine) [19, 20, 21], together with the capability of semantic informationtion association of human thinking (similarity, contiguity and contrast). Information association sought to comprehend the emerging unity of reason and cause by means of linking. So the three association principles provide the effective information-linking strategies used for autonomous informations generation based on the causality among the generated informations, where information association plays an important role in linking and generating diverse creative informations. By linking the informations in a long-term memory internally with various participants knowledge externally, diverse and evolving creative informations are generated with their relations as a map structure.

Similarity principle links informations with similar attributes; conversely, contrast principle links different informations based on their dissimilarity. The reasoning relationship between different informations can be linked using the contiguity principle. However, these principles can differ according to the context and type of informations that are exchanged [22].

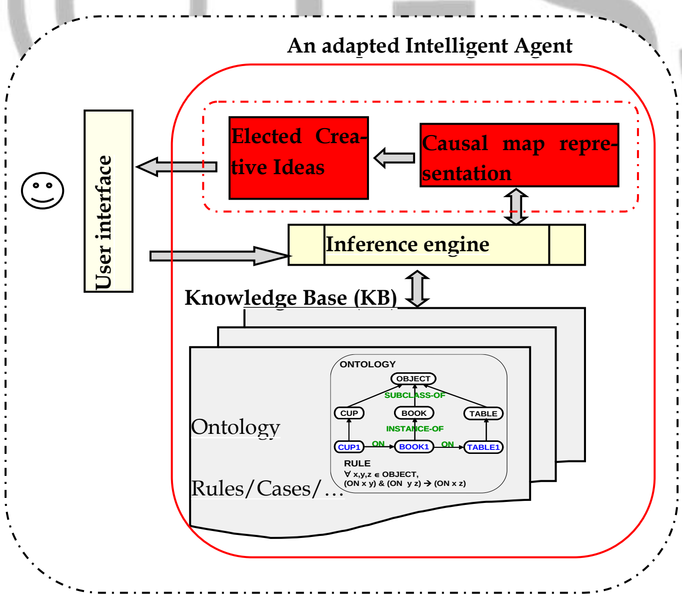


Fig (2): The adapted framework of intelligent agent to explore knowledge

The creative information generated with associative semantic relations between informations, are obtained as a result of analyzing the semantic structures of information association sources by using the causal map approach. Since a graph data structure is a good potential source of information about direction of causal flow, a novel Oriented Directed Acyclic Graph (ODAG) is introduced, for building an efficient information map that represent diverse information and their relationships. This map provides an environment, in which agents can learn and share their knowledge.

A. The proposed Oriented Directed Acyclic Graph (ODAG) algorithm

The basic concept of a graph theory is the many to many relationships between entities (each entity may have multiple predecessors or multiple successors). A graph may be undirected i.e. there is no direction (arrowhead) on any of the lines (known as edges) that connect two entities, or directed graph i.e. its lines (known here as arcs) may be directed from one entity to another [23, 24]. Whereas the decision graph has a new vision in the problem solving process for elevating both information quantity and information quality. Given the Vertices as a partial problem solving states (creative generated information) and the Arcs as the steps in a problem solving process (causal relationships), the Directed Acyclic Graph G=(V,A) would be represented as a collection of vertices (entities) and a collection of directed arcs that connect pairs of vertices with no path returning to the same vertex (acyclic), Where V is the set of vertices or states of the graph and A is the set of arcs between vertices. The ODAG algorithm in fig(3), based on two sets of data, the first set represents the vertices in a one dimensional array Vertex(V) where each item (creative information instance) in this array is annotated with the creative information label(CIL), round number(Rno), number of parents(Np), creativity value(Cv) and its weight(W). The second set represents the arcs in a two dimensional array Adjacency(V,V) where each item (creative information relation) is annotated with the link (0 or 1) and the relationship type (is-a or part-of).

```
Input
-The initial output of the inference algorithm (parent,
Instance idea, relationship).
 -Available number of rounds.
Output
ODAG structure as vertex vector and adjacent matrix.
Procedure
{ Initial round=1
While (round <= available number of rounds)
foreach Vertex in Vertex-Vector
If (Vertex.CIL != instance idea)
add Vertex(instance idea)
Vertex.Np +=1
Vertex.Rno=round
Add Arc()
Adjacent[parent ,instance idea].link =1
Adjacent [parent ,instance idea].relation= relationship
Else
Vertex.Np +=1
Add Arc()
Adjacent[parent ,instance idea].link =1
Adjacent [parent ,instance idea].relation= relationship
end if
End for
Round+=1
Foreach instance idea in vertex-vector
inference algorithm(instance idea, 0, 0)
endfor
End while}
```

Fig (3): The proposed Oriented Directed Acyclic Graph (ODAG) algorithm

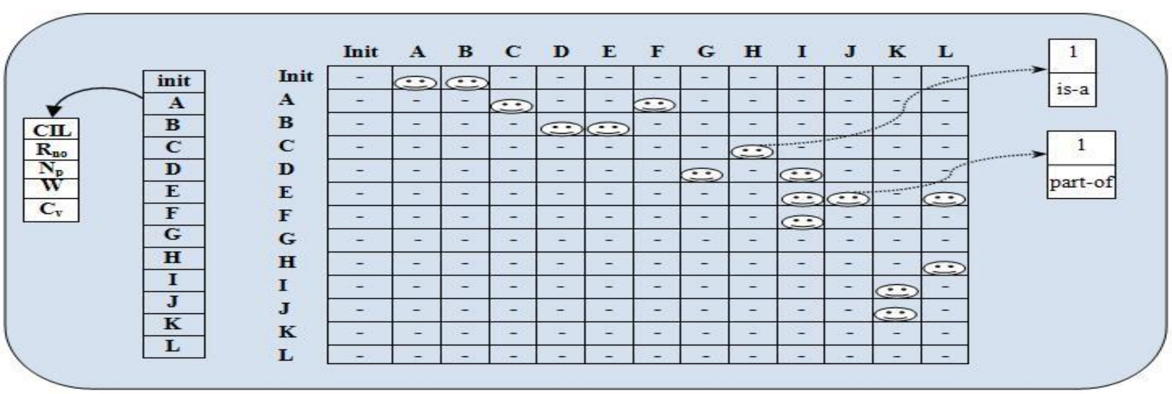
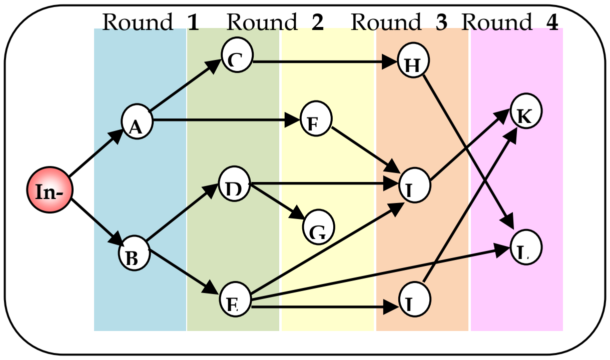


Fig (4): An example of allocating the generated creative ideas in *Vertex(v)* and their relationships in *Adjacency(V,V)*

In the initial state, there is no item in both arrays, and then the algorithm will be fed with the subsequence informations from the inference algorithm and the available number of rounds. At every round, once the inference algorithm delivers the parent, the generated creative information, its creativity value (number of sibling from the ontology) and their relationship to the ODAG algorithm, then the ODAG algorithm spontaneously allocate this generated creative information instance in Vertex(V) vector if it is not already exist then set round number and ignoring weight, simultaneously it allocates their relationships to its parent in Adjacency(V,V) matrix, an illustration example is shown in fig(4). Subsequently the vertices with highest creativity value(Cv) in the vertex(v) of all the previous rounds, would be the initial states to the next rounds until reach the limited number of rounds, and complete the structure of graph, as illustrated in fig(5).



Fig(5): An example of a complete graph structure after five rounds

B. New evaluation of Creative Informations Quality (CIQ) algorithm

After complete the information map construction process as vertex(v) vector and adjacent(v,v) matrix, a Creative Informations Quality (CIQ) algorithm is introduced to collect an efficient informations from the map. Whereas each generated creative information instance in Vertex(V) is annotated with CIL, Rno, Np, Cv and its W, then the Creative Informations Quality (CIQ) algorithm used to compute it's weight(W) based on Rno and Np , as illustrated in fig(6). .Furthermore, it filters the top n (a user-specified number) of generated creative informations as regard as its W of highest value.

Input
- ODAG structure as vertex vector and adjacent matrix
- n (a user-specified number of ideas)
Output
- The wanted quantity of valued ideas
Procedure
{ Foreach *Vertex* in *Vertex-Vector*
j= *Vertex.index*
For i=0 to Vertex.count
***Vertex .W+= Vertex*.Rno * Adjacent [i,j].link**
Endfor
Foreach *Vertex* in *Vertex-Vector* sorted by *W*
and stored into *stored-ideas[]*
While(valued ideas[]<n)
Valued-ideas[]← *Stored-ideas[]*
End While
End for
}

Fig(6): The proposed Creative Ideas Quality (CIQ) algorithm

Experimental Results

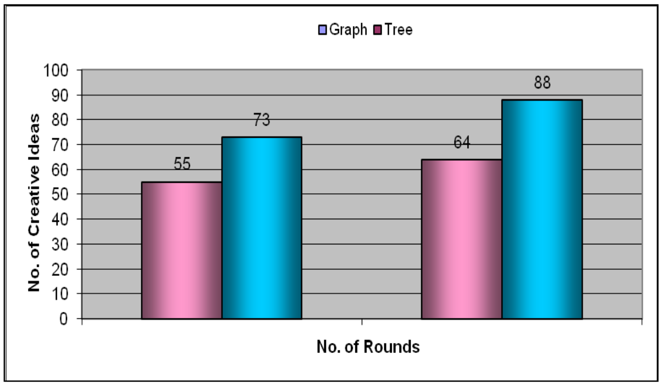
Validation of the proposed algorithms is done via a prototype system in order to clarify the relationships between informations quality and informations quantity. This relationship come out in term of the number of generated informations with higher Cv value versus the number of rounds, using .NET technology to achieve our study and analysis.

Case Study

In fig(4) and fig(5) An example of a complete graph structure after five rounds has been recognized by our system. Subsequently the vertices with highest creativity value(Cv) in the vertex(v),where Vertex(V) is annotated with CIL, Rno, Np, Cv and its W, where W is computed as follow:

$$W= \sum (\text{Vertex.Rno} * \text{Adjacent } [i,j].\text{link})$$

The creative information quality depends on its W, fig(7) shows a comparison on number of creative informations with high quality representing in tree structure and in graph structure. So the decision graph (informations) has higher quality and quantity than the decision tree.



Fig(7): A comparison on number of creative ideas with high quality between graph structure and tree structure

Conclusion

In this paper, an efficient informations map as a graph structure for diverse informations and their relationships, called Oriented Directed Acyclic Graph(ODAG) algorithm was introduced. Moreover, a new evaluation of Creative Informations Quality (CIQ) algorithm by allowing agents to link diverse informations dynamically. Whereas, the creative informations generated with associative semantic relations between informations, are obtained as a result of analyzing the semantic structures of information association sources by using the causal map approach. Since a graph data structure is a good potential source of information about direction of causal flow.

References

[1] Aas, T. H. "Implementing a value assessment tool for service innovation ideas". International Journal of Innovation Management, 14(6), 1149-1167, 2010.
[2] W.-K.Linsey, J. S., Wood, K. L., & Markman, A. B. "Modality and representation in analogy". Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 22(2), 85-100, 2008.
[3] R. Reddy, "Three open problems in AI," JACM 50(1): 83-86, 2003.
[4] .Nov, O., & Jones, M. "Ordering creativity? Knowledge, creativity, and idea generation in the advertising industry". International Journal of Product Development, 3(2), 252-262, 2006.
[5] Baregheh, A., Rowley, J., & Sambrook, S. "Towards a multidisciplinary definition of innovation". Management Decision, 47(8), 1323-1339, 2009.
[6] Sun Yu, Xu Tianwei, Li Zhiping, "Intelligent Search Agents for Web-Based Intelligent Tutoring Systems," csse, vol. 5, pp.1148-1151, International Conference on Computer Science and Software Engineering, 2008.
[7] I.C. Lai, K.T. Kao and T.W. Chang, Role playing for linking ideas in the idea association process. In: A. Bhatt, Editor, Proceedings of Tenth Conference on Computer Aided Architectural Design Research in Asia pp. 119–129, 2005.
[8] I.C. Lai, Mapmaker: linking plays among ideas, cases and links. In: A. Bhatt, Editor, Proceedings of Tenth Conference on Computer Aided Architectural Design Research pp. 222–227, 2005.
[9] I.C. Lai, Dynamic idea maps: a framework for linking ideas with cases during brainstorming, International Journal of Architectural Computing Vol 3 (No 4), pp. 429–447,2005.
[10] Liuqing Chen, Pan Wang , Hao Dong, Feng Shi, Ji Han, Yike Guo, Peter R.N. Childs, Jun Xiao, Chao Wuc. "An artificial intelligence based data-driven approach for design ideation "J. Vis. Commun. Image R. 61, 10–22, 2019.
[11] H.T. Lin. A Multifunctional system for supporting collaborating works and decision making. Information Technology Journal 8(1) , 49-56, 2009.
[12] Pierry Teza, Viviane Brandão Miguez, Roberto Fabiano Fernandes, Gertrudes Aparecida Dandolini, João Artur de Souza. "Ideas for innovation: a systematic mapping of the literature", Gest. Prod., São Carlos, v. 23, n. 1, p. 60-83, 2016.
[13] M. Aiken, M. Vanjani, and J. Paolillo, "A Comparison of Two Electronic Idea Generation Techniques," Information and Management, vol. 30, pp. 91-99, 1996.
[14] R.O. Briggs and B.A. Reinig, "Bounded Ideation Theory: A New Model of the Relationship between Idea-Quantity and Idea-Quality During Ideation," Proc. 40th Ann. Hawaii Int'l Conf. System Sciences (HICSS '07), 2007.
[15] S.T. yuan, and Y.C. Chen. Semantic ideation learning for agent-based E-brainstorming. IEEE Trans. Knowledge Data Eng. , vol. 20, no 2, pp. 261-275, February -2008.
[16] Brahim Chaib-draa, Causal Maps: Theory, Implementation, and Practical Applications in Multiagent Environments. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 14, NO. 6, NOVEMBER/DECEMBER 2002.

[17] S. Nadkarni and P.P. Shenoy, “A Bayesian Network Approach to Making Inferences in Cognitive Maps,” European J. Operational Research, vol. 128, no. 3, 2001.

[18] A.F. Osborn, “Applied Imagination: Principles and Procedures of Creative Problem-Solving,” Creative Education Foundation, third revised ed., 1993.

[19] Takadama, K., Kawai, T., and Koyama, Y. “Can Agents Acquire Human-like Behaviors in a Sequential Bargaining Game? – Comparison of Roth’s and Q-learning agents,” The Seventh International Workshop on Multi-Agent-Based Simulation (MABS’06), pp.153–166. 2006.

[20] Watkins, C. J. C. H. and Dayan, P. (1992) “Technical Note: Q-Learning,” Machine Learning, Vol. 8, pp. 55–68.

[21] Keiki Takadama, Tetsuro Kawai, Yuhsuke Koyama, Micro- and Macro-level Validation in Agent-based simulation: Reproduction of Human-like Behaviors and Thinking in a Sequential Bargaining Game, Third International Model-to-Model Workshop (M2M), 2007.

[22] Ih-Cheng Lai and Teng-Wen Chang, A distributed linking system for supporting idea association during the conceptual design stage, Science Direct, Volume 27, Issue 6, November 2006, Pages 685-710 .

[23] Vahid Rezaei Tabar, Hamid Zareifard, Selva Salimi & Dariusz Plewczynski,"Learning directed acyclic graphs by determination of candidate causes for discrete variables", Pages 1957-1970, 14 Apr 2019.

[24] DAVID J. PEARCE and PAUL H. J. KELLY, A Dynamic Topological Sort Algorithm for Directed Acyclic Graphs, ACM Journal of Experimental Algorithmics, Vol. 11, Article No. 1.7, Pages 1–24, 2006.