International Journal of ChemTech Research
CODEN (USA): IJCRGG ISSN: 0974-4290
Vol.9, No.02 pp 242-248, 2016

An Overview of application of Graph theory

A.Prathik[1], K.Uma[2], J.Anuradha[3]

[1]School of Information Technology and Engineering, VIT University, Vellore-632014, Tamil Nadu, India.
[2]Applied Analysis Division, School of Advanced Sciences, VIT University, Vellore-632014, Tamil Nadu, India.
[3]School of Computing Sciences and Engineering,VIT University, Vellore-632014, Tamil Nadu, India.

Abstract: Graph theory is growing area as it is applied to areas of mathematics, science and technology. It is being actively used in fields of biochemistry, chemistry, communication networks and coding theory, computer science(algorithms and computaion) and operations research (scheduling) and also used in many application like coding theory, x-ray crystallography, radar, astronomy, circuit design, communication network addressing, data base management. This paper gives an overview of the applications of graph theory in heterogeneous fields to some extent, but mainly focuses on the computer science applications and chemistry that uses graph theoretical concepts. Various papers based on graph theory have been studied related to scheduling concepts, computer science applications and an overview has been presented here.
Keywords: Graphs, network, application of graphs, graph algorithms, bipartite graph, Chemistry.

Introduction

In mathematics computer science, chemistry graph theory is the study of graphs, which are mathematical structures used to model pair wise relations between objects. A "graph" in this context is made up of "vertices" or "nodes" and lines called edges that connect them. A graph may be undirected, meaning that there is no distinction between the two vertices associated with each edge, or its edges may be directed from one vertex to another; see graph (mathematics) for more detailed definitions and for other variations in the types of graph that are commonly considered. Graphs are one of the prime objects of study in discrete mathematics. In the most common sense of the term[1]a graph is an ordered pairG=(V,E) comprising a set V of vertices or nodes together with a set E of edges or lines, which are 2-element subsets of V (i.e., an edge is related with two vertices, and the relation is represented as an unordered pair of the vertices with respect to the particular edge). The origin of graph theory started with the problem of Koinsberg Bridge, in 1735. This problem lead to the concept of Eulerian Graph. Euler studied the problem of Koinsberg bridge and constructed a structure to solve the problem called Eulerian graph. In 1840, A.F Mobius gave the idea of complete graph and bipartite graph and Kuratowski's proved that they are planar by means of recreational problems. The concept of tree, (a connected graph without cycles[2]) was implemented by Gustav Kirchhoff in 1845, and he employed graph theoretical ideas in the calculation of currents in electrical networks or circuits. In 1852, Thomas Gutherie found the famous four colour problem. Then in 1856, Thomas. P. Kirkman and William R.Hamilton studied cycles on polyhydra and invented the concept called Hamiltonian graph by studying trips that visited certain

sites exactly once. In 1913, H.Dudeney mentioned a puzzle problem. Even though the four colour problem was invented it was solved only after a century by Kenneth Appel and Wolfgang Haken. This time is considered as the birth of Graph Theory.

Caley studied particular analytical forms from differential calculus to study the trees. This had many implications in theoretical chemistry. This lead to the invention of enumerative graph theory. Any how the term "Graph" was introduced by Sylvester in 1878 where he drew an analogy between "Quantic invariants" and co variants of algebra and molecular diagrams. In 1941, Ramsey worked on colorations which lead to the identification of another branch of graph theory called extremel graph theory. In 1969, the four colour problem was solved using Computers by Heinrich. The study of asymptotic graph connectivity gave rise to random graph theory[2, 3]. Combinatorial methods have important applications in chemical kinetics, For example, combinatorial and topological methods in nonlinear chemical kinetics.[4]some results of nonlinear chemical systems were derived[5]using graph theoretical methods. Graph theoretical models of finding the possible mechanisms for a given type of reaction[6]has shown that graph theory can be used to determine the dynamics of complex chemical reactions such as oscillatory reactions. The use of graph theory in chemical dynamics[7].

Graphs in Chemistry

All structural formulas of covalently bonded compounds are graphs; they are therefore called constitutional graphs. More than 90% are organic or contain organic ligands in whose constitutional formulas the lines (edges of the graph) symbolize covalent two-electron bonds. Constitutional graphs represent only one type of graphs that are of interest to chemists. This analysis will try to highlight the application of graphs in chemistry. Graph theory provides the basis for definition, enumeration, systematization, codification, nomenclature, correlation, and computer programming[8-14].

The chemical information is associated with structural formulas and that structural formulas may be systematically and uniquely indexed and retrieved. One does translate chemical structures into words by means by nomenclature rules. One should note that in chemical information words usually come afterward. The importance of graph theory (GT) for chemistry stems mainly from the existence of the phenomenon of isomerism, which is rationalized by chemical structure theory. This theory accounts for all constitutional isomers by using purely graph-theoretical methods.

Defining and finding the constitutional isomers (which correspond to the same given molecular formula) are purely graph-theoretical problems. The problem of isomerism is the real crux of the documentation/retrieval problem. Although molecular formulas can be ordered for indexing purposes according to simple rules, It is here that more sophisticated techniques of graph theory (GT) may help.

According to definite rules, the essence of chemistry is the combinatorics of atoms. Until recently, most theoretical chemists viewed mathematics as a tool for professing (crunching) numerical data, but the present trend toward nonnumerical methods is noticeable, mainly due to the impact of graph theory (GT).

Constitutional and Steric Isomerism

In graph theory the numbers of lines meeting at a vertex are called vertex degree. Graphs with all equal vertex degrees are called regular graph. The mathematician Cayley proposed [15]128 years ago an algorithm for calculating the numbers of constitutional isomers for alkanes, C_nH_{2n+2}, and alkyl derivatives, e.g., $C_nH_{2n+1}Cl$.

Of course, the n vertices of degree 4 correspond to carbon and the 2n + 2 vertices of degree 1 to hydrogen atoms. Alternatively, one may ignore all hydrogen and look for the equivalent set of trees with n vertices of degrees 1-4; edges now correspond exclusively to C-C bonds, and together with the vertices they form the hydrogen-depleted graph, or skeleton graph, of the molecule.a0505

Data Mining

Data mining is the process of observing interesting information from a huge data sets by using methods from machine learning, artificial intelligence, statistics and database systems. A key idea is to reveal patterns in the large data set and in general includes the following

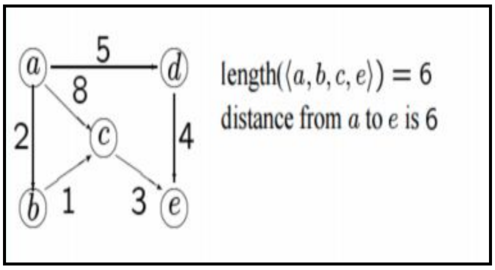
- Anomaly detection (detection of change/ deviation) which is the recognition of strange data records that might be interesting or has data errors.
- Learning of association rule or modelling of dependence which discovers relationship between variables.
- Clustering that groups the objects which are similar.
- Classification which infers known structure to apply to new data.
- Regression that reveal a function which groups available data with smaller error.
- Summarization which briefs about compress representation of data set including generation of report and visualization.

Clustering analysis is a important part of data mining. The aim of clustering algorithm is to group the objects based on the described information about them. Classes are considered by clusters which assign objects automatically to clustering process. As class label is not to be known ahead of time, clustering is often referred to as “unsupervised learning”. More number of clustering algorithms is available in data mining literature . In clustering of data are done based on the similarity of objects. Objects which are similar to each other are grouped into same clusters and objects which are dissimilar are grouped into different groups.

Algorithms and graph theory

1. Dijkstra’s Shortest Path Algorithm

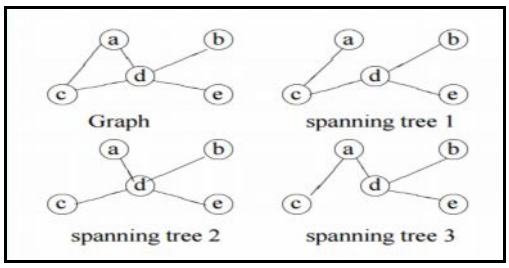
This algorithm is for graph with no negative weights. The idea of the algorithm is to maintain the list of vertices at every stage of graph that we have discovered. In terms of path weight, we process the vertex that is closest to the source. We keep track of the minimum path weight among all the paths from the source we’ve found so far. we explore every edge e = (v, u) that leaves v, when we process a vertex v and by using (v, u) as the last edge, consider the new path from source to u. we update the minimum path weight of vertex u, based on the weight of the new path.Let G= (V, E) be a weighted digraph ω= E→R mapping edges of real-valued weights. If e= (u, v) we write ω (u,v) for ω(e). The length of path p = ($v_0, v_1, \ldots, v_k$) is the sum of weight of its constituent edgesLength (p) = $\sum_{(i=1)}^{k} \omega(v_{(i-1)}, v_i)$ .The distance from u to v is denoted δ(u,v), is the length of minimum length path if there is a path from u to v and is ∞ otherwise[16]as shown in the graph 1.



Graph. 1

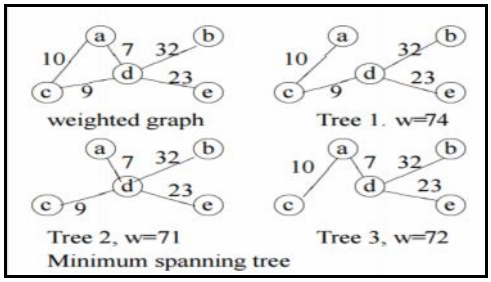
2. Minimum Spanning Tree

A spanning tree of that graph is a subgraph that is a tree and connects all the verticestogether, Given a connected, undirected graph**.** We can also assign a weight to each edge, which is a number representing how unfavourable it is, and use this to assign a weight to a spanning tree by computing the sum of the weights of the edges in that spanning tree**.** Spanning Tree is a subgraph T of undirected graph G= (V, E) is a spanning tree of G if is a tree and contains every vertex of G. A minimum spanning tree (MST) or minimum weight spanning tree is then a spanning tree with weight less than or equal to the weight of every other spanning treeas shown in the graph 2.



Graph.2

A Minimum Spanning Tree in an undirected connected weighted graph is a spanning tree of minimum weight (among all spanning trees)[17]as shown in the graph 3.



Graph.3

3. Finding graph planarity

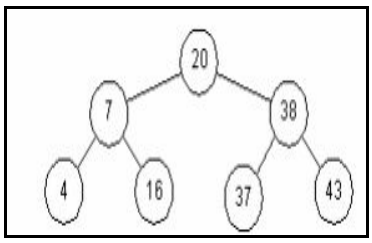
In graph theory, algorithmic of testing whether a given graph is planar graph(that it whether it can be in plane without edge intersections) or not is a problem of planarity testing. Most of these methods operate in O(n) time (linear time), where n is the number of edges (or vertices) in the graph, which is asymptotically optimal.Kuratowski's theoremis an algorithm for testing planarity in graphs, that a graph is planar if and only if it does not contain a sub graph that is a subdivision of K_5 (the complete graph on five vertices) or $K_{3,3}$ (the utility graph, a complete bipartite graph on six vertices, three of which connect to each of the other three)[18-19].

Planarity testing using Kuratowski's Theorem:

To test for K_5's subdivision, choose 5 vertices of G and check if all the 5 vertices re connected by $\binom{5}{4}$=10 distinct paths as K_5. To test for $K_{3,3}$'s subdivision choose 6 vertices of G. check if all vertices are connected by 3×3=9 distinct paths as $K_{3,3}$. Both are obviously exponential time algorithm.

4. Algorithms for searching an element in a data structure

A binary search tree is a tree where each node has a left and right child. Either child, or both children, may be missing. Assuming k represents the value of a given node, then a binary search tree also has the following property: all children to the left of the node have values smaller than k, and all children to the right of the node have values larger than k. The top of a tree is known as the root, and the exposed nodes at the bottom are known as leaves. In Figure below, the root is node 20 and the leaves are nodes 4, 16, 37, and 43. The height of a tree is the length of the longest path from root to leaf. For this example the tree height is 2 and shown in the following graph 4.



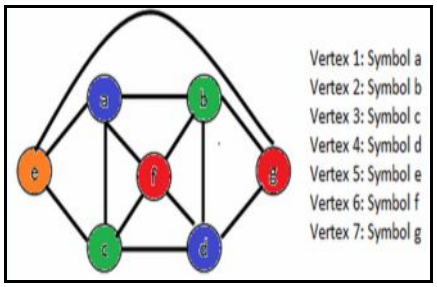
Graph. 4.

To search a tree for a given value, we start at the root and work down. For example, to search for 16, we first note that 16 < 20 and we traverse to the left child. The second comparison finds that 16 > 7, so we traverse to the right child[20].

Graph Colouring

Graph Coloring is an assignment of colors (or any distinct marks) to the vertices of a graph. Strictly speaking, a coloring is a proper coloring if no two adjacent vertices have the same color. A (vertex) coloring of a graph G is a mapping c: V(G)→S. The elements of S are called colors. If |S|=k, we say that c is a k-coloring. Coloring is proper if all adjacent vertices have different colors.

A graph is k-colourable if it has a proper k-colouring.The chromatic number χ (G) is the least k such that G is k-colourable. Obviously, χ (G) exists as assigning distinct colours to vertices yields a proper |V (G)|-colouring. An optimal colouring of G is a χ (G)-colouring. A graph G is k-chromatic if χ (G) = k. In a proper colouring, each colour class is a stable set. Hence a k-colouring may also be seen as a partition of the vertex set of G into k disjoint stable sets Si = {v | c (v) = i} for $1 \leq i \leq k$ as given in the following graph 5.



Graph. 5

Aircraft scheduling

Aircraft scheduling could roughly be described as:

• Input: a set of aircraft + a set of flights.

• Output: individual rosters for each aircraft so that all flights are assigned.

The timetable of the flights is fixed (i.e. departure and arrival times) when starting the aircraft scheduling, but the type of aircraft is not yet decided. The choices of aircraft type for each flight depend on the expected number of passengers, and on the distance that has to be flown. For example a Boeing 767 can take twice as many passengers as a DC9, and can also fly a longer distance (e.g. across the Atlantic Ocean). Typically we would reserve our 767: s to fly our long-haul flights, and let the DC9:s fly the short haul. But for short haul flights with a high passenger load we would also need a 767.So the aircraft types have to be considered already when making the timetable, but it is still not fixed; only restricted. It is in the aircraft scheduling that the aircraft type of each flight will be fully defined.

Job scheduling

In computing, scheduling is the method by which work specified by some means is assigned to resources that complete the work. The resources may be virtual computation elements such as threads, processes or data flows, which are in turn scheduled onto hardware resources such as processors, network links or expansion cards. Schedulers are often implemented so they keep all compute resources busy (as in load balancing), allow multiple users to share system resources effectively, or to achieve a target quality of service. Scheduling is fundamental to computation itself, and an intrinsic part of the execution model of a language. A scheduler is what makes it possible to have multitasking as a way of executing more than one process at a time on a single processor[21]

Scheduling involves in any one of goals for examples, throughput maximization (work completed per time unit), response minimizing (until first point, the time from work become enabled it, gives execution on resources), latency minimizing (the time between work becoming enables and its subsequent completion, fairness maximizing (each process having equal CPU time. In general, these goals often conflict (throughput versus latency).

Graph Theory Applications

In computer science, graphs are used to represent networks of communication, data organization, computational devices, the flow of computation, etc. For instance, the link structure of a website can be represented by a directed graph, in which the vertices represent web pages and directed edges represent links from one page to another. In condensed matter physics, the three-dimensional structure of complicated simulated atomic structures can be studied quantitatively by gathering statistics on graph-theoretic properties related to the topology of the atoms. In chemistry a graph makes a natural model for a molecule, where vertices represent atoms and edges bonds. This approach is especially used in computer processing of molecular structures, ranging from chemical editors to database searching.

In statistical physics, graphs can represent local connections between interacting parts of a system, as well as the dynamics of a physical process on such systems. Graphs are also used to represent the micro-scale channels of porous media, in which the vertices represent the pores and the edges represent the smaller channels connecting the pores.Likewise, graph theory is useful in biology and conservation efforts where a vertex can represent regions where certain species exist (or habitats) and the edges represent migration paths, or movement between the regions. This information is important when looking at breeding patterns or tracking the spread of disease, parasites or how changes to the movement can affect other species[22].

Graph based approach for finger print classification

These approaches are based on the observation that relational graphs are independent of the fingerprint image rotations. Therefore, the relational graph appears to be suitable for overcoming the rotation issue in fingerprint representations. However, each graph node needs to be enriched with a set of features which are usually dependent on fingerprint rotation, thus the problem is only partially solved. The first ideas for obtaining graph-based representations of fingerprints are based on the segmentation of the orientation field into regions having homogeneous directions. A node of the graph is associated to each region and an edge joins two nodes associated to adjacent regions.

This idea has been taken up by Lumini et al., who perform inexact graph matching with a template graph for each class to determine the best match for the final classification[23].Due to the high variability of the segmentations obtained, it is very difficult to find a set of graph prototypes. A large number of prototypes are needed to represent as many variations as possible, but the high computational complexity of graph matching algorithms only increases classification time.

Following the same idea as Lumini et al., Cappelli et al. found a set of prototype-segmentations for each class, and designed an algorithm to ”guide” the orientation field segmentation in order to produce a class-dependent segmentation[23]. A cost is associated to each segmentation. The orientation field of an input fingerprint is segmented according to five dynamic masks. The least cost segmentation corresponds to the “structure” that best represents a given orientation field, and the relative class is associated to the fingerprint.

However, detected prototypes are not optimal for effectively discriminating between fingerprint classes, because of the small between-class separation.

Conclusion

In this manuscript we reviewed applications of graph theory and research challenges have also been surveyed. The topics we reviewed include biochemistry, chemistry, communication networks, coding theory, algorithms, computation, operations research, x-ray crystallography, radar, astronomy, circuit design, communication network addressing and data base management.There are many problems in this area which are yet to be solved. It is hoped that this review would attract many new investigators into graph theory.

References

1. Iyanaga and Kawada, 69 J, p. 234 or Biggs, p. 4.
2. Narasingh Deo, "Graph theory with applications to engineering and computer science", Prentice Hall of India, 1990.
3. Shrinivas.S.G, Vetrivel.S, Elango.N.M. International Journal of Engineering Science and Technology Vol. 2(9), 2010, 4610-4621.
4. Glass, L. J. Chem. Phys. 1975, 63, 1325.
5. Beratta, E.Vetrano, F.Solimano, F.Lazzari, C. Bull Math.Bid.1979, 41, 641.
6. King, R. B. Theor. Chim. Acta 1980, 56, 269
7. Clarke, B. L. Stud. Phys. Theor. Chem. 1983,28, 322
8. Balaban, A. T., Ed. "Chemical Applications of Graph Theory"; AcademicPress: London, 1967.
9. Balaban, A.T. Rouvray, D. H. In"Applications of Graph Theory";Wilson. R. J.: Beinecke. L. W.Eds.:Academic Press: London. 1976.
10. Rouvray, D. H. Chem.Soc. Rev**.** 1974, 3,355
11. Balaban. A. T. Math. Chem. 1975. 1. 33.
12. King, R: B., Ed. "Chemical Applications of Topology and GraphTheory". Stud. Phys. Theor. Chem. 1983, 23.
13. Biggs, N. L.; Lloyd, E. K.; Wilson, R. J. "Graph Theory 1736-1936.
14. Cayley, A. Philos. Mag.1857, 13 (l), 172.
15. http://www.cse.ust.hk/~dekai/271/notes/L10/L10.pdf .
16. https://www.cse.ust.hk/~dekai/271/notes/L07/L07.pdf.
17. http://www.tcs.tifr.res.in/~workshop/nitp_igga/slides/shreesh-planarity-patna.pdf
18. https://en.wikipedia.org/wiki/Graph_theory#cite_ref-1.
19. http://epaperpress.com.
20. Lumini,A. Maio,D. and Maltoni,D. Inexact graph matching for Fingerprint Classification, Machine Graphics and Vision , 1999, 8 (2) 241-248.
21. Cappelli,R.Lumini,A. Maio,D. and Maltoni,D. Fingerprint Classification by Directional Image Partitioning, IEEE Trans. on PAMI, 1999, 21 (5) 402-421.
22. Fayyad,U. Shapiro,G.P. and Smyth,P. "From data mining to knowledge discovery in Databases," AI Magazine, 1996, 37-53.
23. Jain,A.K. Murty, M.N.and P. J. Flynn, "Data clustering: A review," *ACM Computing Surveys (CSUR)*, 1999,31(3),264-323.

*****