MATHEMATICAL PROPERTIES OF COMPLEX NETWORKS

Angel Garrido

ABSTRACT. Many researchers are attempting to create systems which mimic human thought, or understand speech, or beat to the best human chess-player [14]. Understanding intelligence and Creating intelligent artifacts both are the twin goals of *Artificial Intelligence (AI)*.In more recent times, the interest is focused on problems related with Complex Networks [3, 5,6, 19], in particular on questions such as clustering search and identification. We attempt, in this paper, a panoramic vision of such mathematical methods in AI.

KEYWORDS: *Searching and Representation Methods, Heuristics, Artificial Intelligence*

2000 *Mathematics Subject Classification*: 68R10, 68R05, 78M35.

1.Introduction

The origin of the ideas about thinking machines, the complex mechanism through work the human brain, the possibility of mimic its behavior, if we produce some computational structure similar to neuron, their synapsis or connections between neurons to produce Artificial Neural Networks [6, 14, 19]. All this can appear with resonances of a Science Fiction movie, but it is a real subject of study, and it is so from many years ago, and more in recent times [3].

The basic purpose of the *AI* is to create an admissible model of the human knowledge. Its subject is, therefore, "pure form". We try to emulate the way of reasoning of a human brain. This must be in successive, approximating steps, but the attempts proceed always in this sense. Initially, the work in *AI* was over idealizations of the real world, attempting the automatical proof of lemmas and theorems, modeling games, etc. Therefore, the fields were "formal worlds". Such search procedures were into the Space of States. This space integrate the set of all states, or nodes, in the case of representation by graphs, that we obtain when we apply all the disposable operators. Many early *AI* programs used the same basic algorithm. To achieve some goal (winning a game or proving a theorem), they proceeded step by step towards it (each time, a move or a deduction), i.e. as searching through a maze, backtracking whenever they reached a dead end. This paradigm is called *"reasoning as search"* [14].

The techniques for solving problems, in *AI,* can be of two different types, Declarative and Procedural.

Declarative: it permits the description of the known aspects of the problem. It is the *Heuristic Treatment.*

Procedural: itemizes the necessary paths to reach the solution of the problem. It is the *Algorithmic Treatment.*

But to pose problems is equivalent to constructs its solutions. This requires *an agent,* the system or program to execute a set of actions, which allow to reach such objectives; and *a procedure of election,* which allow to decide between distinct paths to reach its solution.

We can use [6, 14] a series of resources approaching problems in A I, as may be *Logic, Rules, Associative Nets, Frames, and Scripts.* The correct choice between these methods must be based in the own characteristics of the problem and our expectations about the type of solution. In many cases, we

take at a time two o more tools, as may be on the case of Frame Systems, frequently associated with Rules.

The *Inference in Systems Based Rules* (SBR in acronym) depends on establishing the certainty of some statement, from the disposable information. We have two methods, *going forward and going backward concatenation.* In the first case, we depart of Rules with verified affirmations in its antecedent, advancing through the affirmations which we find in their consequents. While in the second case, we depart of Rules verified in certain consequent (all the consequent must be also verified), and we turn back to the antecedent. This convert their affirmations in new *sub-objectives* for the proof, searching Rules where appear in their consequent, and so on. To hold up this process when we find the required affirmation, in the last consequent explored or the last antecedent, according the selected method.

The Rules [14] shows advantages on Classical Logic, where the reasoning was monotonic, with inferences without incurs in contradictions with pre-existing facts. While in the RBS we can delete or substitute facts of the Base of Facts, according to the new inferences. So, all may be provisional and modifiable. This produce a *Non-Monotonic Reasoning.*

2. Representation methods

We have some *Mechanisms of Control in RBS.* So, with Mechanism of Refractarity, by which we prevent to execute newly a Rule, once has been used, if do not exist more information which allow us or recommend such case. *Rule Sets.* It allows us to activate or neutralize Blocks of Rules. *Meta-Rules,* or Rules which treat (or reasoning) about other Rules. Such Meta-Rules can collaborate to obtain the Control of Reasoning, with the change or assignation of priorities to different Rules, according to the evolution of circumstances.

Nets. Between them, the more recent studies [6] to deal with Bayesian Nets (BN in acronym), Belief Nets, or simply Networks. Before than these BNs appears, the purpose were simply to obtain useful systems for medical diagnosis, by classical statistical inferential techniques, such as can be the Bayes Rule, also called Formula or Theorem.

A *Bayesian Net* is represented as a pair *(G, D),* where G is a directed, acyclic and connected graph, and D will be a probability distribution, associated with the random variables. Such distribution must verify the *Property of Directional Separation,* according which the probability of a variable does not depends of their not descendant nodes.

The *Inference in Bayesian Networks (BNs)* consists in establish on the Net, for the known variables, their values, and for the unknown variables, their respective probabilities. The basic objective of BNs in Medicine is to find the probability of success with we can to give a determined diagnosis, known certain symptoms. We need to work with the subsequent *Hypotheses: Exclusivity, Exhaustivity, and Conditional Independence.* According the *Exclusivity,* two different diagnosis can not be right at time. With the *Exhaustivity,* we suppose at our disposition all the possible diagnosis. And by the *Conditional Independence,* the discoveries found must be mutually independents, to a certain diagnosis. The more usual problem with such hypotheses is their inadequacy to the real world. For this reason, it will be very necessary to introduce Bayesian Networks, Artificial Neural Networks, Probabilistic Graphical Models, and so on.

In the searching process, we have two options, without information of the domain *(Blind Search),* and with information about of the domain *(Heuristic Search).* In the first case, we can elect, according the type of problem, between *Search in extent* and *Search in depth.* There are other methods, obtained from the aforementioned, such as *Searching in Progressive Depth* and *Bidirectional Searching,* both with names sufficiently allusive to its nature. Also we can found another method, and in this case it is not derived, the *General Search in Graphs.* In such procedure, it is clear the possibility of translation to matrix expression, through their incidence matrices. All these methods appears joined, of course, to their corresponding algorithms.

Blind Search, or search without information of the domain. It appears with the initial attempts to solve, by idealizations of the real world, playing problems, or to obtain automatic proof of theorems. The searching process could be in state spaces.

Searching in extent. We advance in the graph [6, 15] through levels. So, we obtain the lesser cost solution, if there exists. Whereas, in the *Depth Searching,* we expand one link each time, from the root-node. If we reach a blind alley into the graph, we back until the nearest node and from this, we take one ramification in the graph. It is usual to establish an *exploration limit,* also called depth limit, fixing the maximal length of the path, from the root. We need to select the more promising trajectories. So, we can not obtain the best solution (optimum), but an efficient approach to her.

Another different procedure will be the *Gradient Method,* or *Climbing Search.* According to this method [6, 14], in the expansion of each node,

we must elect the link which connects with the node of the subsequent level where is greatest the value of *f*, supposing that reach its greatest value in the final node.

Heuristic Search, i.e. searching with knowledge of the domain. Initially, were usual to think that all the paths can be explored by the computer. But it is too optimist. Such exploration will be very difficult, because the well-known phenomenon of "combinatorial explosion" of branching, when we expand. Its spatial and temporal complexity can advise us against its realization. It is possible for a given search procedure, to be either heuristic or algorithmic, or both, or perhaps neither, with respect to some state-space problem. A search procedure oriented toward a problem said to embody heuristic information (i.e. an information which serves to discover). So, it will be called a *heuristic search procedure.* There are some *critics on Heuristic Search,* because its *unpredictability.* They found good solutions, but not necessary the best. This made very convenient to introduce the algorithm *A*,* with their properties, *completeness* and *admissibility.* According with this last property, if there is solution, find it. *A** is a particular case of searching procedure, *"first the best",* into the strategies of "alternative explorations". It belongs to *General Searching in Graph procedures.* In each step, we go revisiting the Open List. Initially, it is empty. But we introduce the root-node. If the Open List does not empty, or does not reach the final node, in each step, we go on with the process, expanding the subsequent node of the Open List. Our successive choices would be based in the previous assignment to each node of the value of *f* in it. The selection of each node is given according to the lesser value of the heuristic function on the nodes of its level, as a general rule. The comparison is carry out into the Open List, with clear independence of the original level of each node. Generally, we prefer the solution of lesser cost. All the visited nodes, then, pass to store in the Closed List. Such nodes remains inactive in the remaining process.

Also, there are adequate strategies designed for the treatment of *Searching problems with two adversaries.* In this case, the general purpose is to choice the more convenient walk, the lesser number of necessary steps to win the game. *Chess,* generally. For these, we may assume alternative moves. In each move, the ideal would be when the player knows his possibilities and realizes the more unfavorable move for its adversary. But is impossible control it completely, because the aforementioned "combinatorial explosion". So, we need to develop a tree of depth searching, with depth limited. Ever supposing

the move of more advantage, in each turn, for each player.

On A I, the problems may be classified according to its level. In a first level, the problems of *decision, learning, perception, planning and reasoning.* In a second level, the tasks of *classification, representation and search.*

When we formulate a problem [14], we depart of the statement, or the explanation of it, in natural language. Fundamentally, its treatment is based on the *"level of knowledge",* introduced by Newell, in 1981, as "abstract level of interpretation of systems, in A I". Also is basic the *"Rationality Principle"*, according to "if a system has the knowledge according to which one of its actions leads to one of its goals, then such action is carried out".

To formulate the knowledge of the domain, denoted by *D,* in a more effective and efficient way, as theory, we need at least three important characteristics, as *completeness, consistency, and tractability.* According the first of them, *Completeness,* each formula must be demonstrable into the theory. According the second of them, *Consistency,* the new contributions to the system not must admits inner contradictions with the previous asserts, or axioms. The *Tractability* must give us a moderate complexity. So, in Derived Calculus, manipulating laws/premises, by Inference Processes, must not result excessive their temporal and spatial complexities.

3. On complex systems

A *system* can be defined as *a set of components functioning together as a whole.* A systemic point of view [1, 3, 17, 19] allows us to isolate a part of the world, and so, we can focus on those aspect that interact more closely than others.

Network Science is a new scientific field that analyzes the interconnection among diverse networks, as for instance, on Physics, Engineering, Biology, Semantics, and so on. Between its developers, we may remember to Duncan Watts, with the *Small-World Network,* and Albert-László Barabasi, which developed the *Scale-Free Network.* About its work, *Barabási* has found that the websites that form the network (of the WWW) have certain mathematical properties. *Network Theory* is an quickly expanding area of Network and Computer Sciences, and also may be considered a part of Graph Theory.

Complex Networks are everywhere. Many phenomena in nature can be modeled as a network [19], as

- Brain structures. The brain as a network of neurons (as nodes), connected by synapses (their edges).

- Social interactions or the World Wide Web (WWW). All such systems can be represented in terms of nodes and edges.

- On Internet, the nodes represent routers, and the edges by wires, or physical connections between them.

- In transport networks, the nodes can represent the cities and the edges the roads that connect them. These edges can have weights.

These networks are not random. The topology of different networks are very close. They follow from the Power Law, with a scale free structure. How can very different systems have the same underlying topological features? Searching the hidden laws of these networks, modeling and characterizing them are the current lines of research.

As a previous concept, we need to introduce the *Clustering Coefficient,* in both versions, Global and Local.

The *Characteristic Path Length* defines the distance from every node to every other node. It will be calculated by the median of the shortest path from every node to every other node.

And the *Diameter* defines the maximum possible distance between all the pair of reachable nodes.

The *Global Clustering Coefficient* (denoted by C) is based on triplets of nodes, and can be obtained through the number of closed triplets over the total number of triplets, both closed and open. It is also possible to be defined as the mean of the clustering indices of all the nodes in the graph. Being computable through the neighbors of the node, and then finding the number of links amongst them. Because the ratio of the number of existing links to teh number of possible links gives the clustering index of the node. So,

$$C = \frac{3 \ card(triangles))}{card \ (connected \ triplets \ of \ nodes)}$$

Whereas the *Local Clustering Coefficient* is related to a particular node, i.It measures the closeness their neighbors are to being a clique, i.e. a complete graph. The more usual notation is C_i.

These measures was introduced by *Watts and Strogatz,* in 1998, to determine whether a graph is a Small-World network.

The degree of a node i, into a graph, is defined as the number of nodes comprised in its neighborhood. Usually, it is denoted as k_i. Such k_i is the total degree of the corresponding node, i, i.e. it may be calculated by adding in-degree of i and out-degree of i. So, we have

$$C_i = \frac{card(edges\ between\ nodes\ of\ their\ neighbor)}{k_i\ (k_i-1)}$$

Finally, the *Local Clustering Coefficient for the whole system,* denoted by **C**, will be

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n} C_i$$

therefore, it is the average of the clustering coefficient for each node. So, it is a property of a node into the network, either of the total network. It tells about the connectedness between the neighbors of the nodes. When the neighborhood is fully connected, the value of the clustering coefficient will be equal to one. And it will be a value next to zero, if there are sparse connections in the neighborhood. With all the scale of intermediate cases.

4. Structural Models

We may distinguish four of such structural models [3, 4, 5, 8, 11]. Thus, *Regular Networks, Random Networks, Small-World Networks* and *Scale-free Networks.*

In the *Regular Network,* each node is connected to all other nodes. I.e. they are fully connected. Because such type of structure, they have the *lowest path length* (L), and the *lowest diameter (D),* being $L = D = 1$. Also they have the *highest clustering coefficient.* So, it holds $C = 1$. Furthermore, the *highest possible number of edges*

$$card\ (E) = \frac{n\ (n-1)}{2} \sim n^2$$

About *Random Graphs (RGs),* we can say that each pair of nodes is connected with probability p. They have a *low average path length,* according to

$$L \approx \frac{\ln n}{\ln \langle k \rangle} \sim \ln n$$

for $n >> 1$. It is because the total network may be covered in $\langle k \rangle$ steps, from which

$$n \sim \langle k \rangle^L$$

Also they possesses a low clustering coefficient, when the graph is sparse. Thus,

$$C = p = \frac{\langle k \rangle}{n} << 1$$

given that the probability of being connected each pair of neighboring nodes is precisely equal to p.

The *Small-World effect* is observed on a network when it has a low average path length, i.e. $L << n$, *for* $n >> 1$. Recall the famous "*six degrees of separation*", also called the *small-world phenomenon.* This hypothesis was firsty formulated by the Hungarian thinker Frigyes Karinthy, in 1929. Then, it was tested by S. Milgram, in 1967. The subjacent idea [2, 8] is that two arbitrarily selected people may be connected by only six degrees of separation (in average, and it is not too larger than this value). Therefore, the diameter of the corresponding graph is not much larger than six. For instance, on social connections.

In the case of the *Watts-Strogatz* (in acronym, WS) *small-world model* (1998), it represent a hybrid case between a Random Graph and a Regular Lattice [6, 8, 17]. So, Small-World models comparts with RGs some common features, as may be Poisson or Binomial degree distribution, near to Uniform distribution; network size: it does not grow; each node has approximately the same number of edges, i.e. it shows a homogeneous nature.

WS-models shows the low average path length typical of Random Graphs,

$$L \sim \ln n, \; for \; n >> 1$$

And also such models the usual high clustering coefficient of Regular Lattices, giving

$$C \approx 0.75, \ for \ k >> 1$$

In consequence, WS-models have a small-world structure, being well clustered. Whereas the Random Graphs coincides on the small-world structure, but they are poorly clustered. This model (WS) has a peak degree distrution, of Poisson type.

In reference to the last model, called *Scale-Free Network,* it appears when the degree distribution follows a *Power-Law,* i.e.

$$P\left(k\right) \sim k^{-\gamma}$$

In such case, there exists a small number of highly connected nodes, called *hubs*, being the tail of the distribution. While, the great majority of the set of their nodes have few connections, representing the head of such distribution. Such model [3] was introduced by A.-L. Barabási and R. Albert, in 1999. Some of their *features* may be: non-homogeneous nature, in the sense that some (few) nodes have many edges from them, and the remaining nodes only have very few edges, or links; about the network size, it continuously grows; and about the connectivity distribution, it obeys a Power-Law distribution.

A very notable example of Scale-Free Network [2, 8] may be the *World Wide Web* (WWW, in acronym). As we known, it is a collection of many sub-networks. In such case, the nodes corresponding to their pages, whereas the edges to hyperlinks. About the Web graph characteristics, we notice as more important the *Scale Invariance* and its structure of *Small-World Network.*

5. Conclusions

This panoramic over view of the different types of searching methods must be adequately enlarged [4, 7,10-13, 16]. Surely also the description of Complex Networks, and the analysis of the problems involved into the search of clustering. But our initial purpose only was to explain how it appears, and its more efficient use. We hope at least a partial success in such objective.

References

[1] R. Albert, and A.-L. Barabási (2002). Statistical Mechanics of Complex Networks. *Reviews of Modern Physics.* **74:** 47-.

[2] A.-L. Barabási, and Bonabeau (2003): Scale-Free Networks. *Scient. Am.* May 2003, pp. 50-59.

[3] A.-L. Barabási (2004). *Linked: How Everything is Connected to Everything Else.* Plume Publ.

[4] A. Barrat et al. (2008). *Dynamical processes in Complex Networks.* Cambridge University Press.

[5] S. Bocaletti et al. (2006): Complex Networks: Structure and Dynamics. *Phys. Rep.* **424**: 175-308.

[6] S. Bornholdt, and H. G. Schuster, editors (2003): *Handbook of Graphs and Networks: From the Genome to the Internet.*

[7] A. Broder et al. Graph structure in the web. *Proc. 9th. WWW:* 309-320.

[8] G. Calderelli (2007). *Scale-Free Networks.* Oxford University Press.

[9] R. Cohen et al. (2000). Resilience of the Internet to random breakdown. *Phys. Rev. Letter,* **85:** 4626-.

[10] S. Dill et al. (2002). Self-similarity in the web. *ACM Trans. Intern. Techn.:* 205-223.

[11] S. N. Dorogotsev, and J. F. F. Mendes (2002). Evolution of Networks. *Adv. Phys.* **51**: 1079-.

[12] S. N. Dorogotsev, and J. F.F. Mendes (2003). *Evolution of Networks: From biological networks to the Internet and WWW.* Oxford University Press.

[13] S. N. Dorogotsev, A. V. Goltsev, and J. F. F. Mendes (2008). Critical phenomena in Complex networks. *Rev.Mod. Phys.,* **80**:1275-.

[14] A. Garrido (2004). Logical Foundations of Artificial Intelligence. *Proc. "FotFS V" Intern. Conf.,* Institut für Angewändte Mathematik, University of Bonn.

[15] F. Harary. *Graph Theory.* Addison-Wesley, 1975.

[16] P. R. Kumar et al. (2000). Stochastic models for the web graph. *Proc. 41st. FOCS:* 57-65.

[17] M. Newman (2003). The structure and function of Complex Networks. SIAM Review, **45**: 167-256.

[18] G. Shafer (1976). *A Mathematical Theory of Evidence.* Princeton Univ. Press. New Jersey.

[19] R. Solé (2009). *Redes Complejas. Del genoma a Internet.* Colección Metatemas. Tusquets Editores, Barcelona.

[20] A. M. Turing (1950). Computing machinery and intelligence. *Mind,* **59:** 433-60.

[21] N. Wiener (1947). *Cybernetics.* MIT Press and John Wiley. New York.

[22] L. A. Zadeh (1965). Fuzzy Sets. *Information and Control,* **8:** 338-55.

Angel Garrido
Departamento de Matemáticas Fundamentales
Facultad de Ciencias de la UNED
Senda del Rey, 9
28040-Madrid
Spain
e-mail: agarrido@mat.uned.es; algbmv@telefonica.net