Improving Tools in Artificial Intelligence

Angel Garrido
Faculty of Sciences, National University of Distance Education, Madrid, Spain
Paseo Senda del Rey, 9, 28040, Madrid, Spain
algbmv@telefonica.net

Abstract

The historical origin of the Artificial Intelligence (AI) is usually established in the Dartmouth Conference, of 1956. But we can find many more arcane origins [1]. Also, we can consider, in more recent times, very great thinkers, as Janos Neumann (then, John von Neumann, arrived in USA), Norbert Wiener, Alan Mathison Turing, or Lofti Zadeh, for instance [8, 9]. Frequently AI requires Logic. But its Classical version shows too many insufficiencies. So, it was necessary to introduce more sophisticated tools, as Fuzzy Logic, Modal Logic, Non-Monotonic Logic and so on [1, 2]. Among the things that AI needs to represent are *categories*, *objects*, *properties*, *relations between objects*, *situations*, *states*, *time*, *events*, *causes and effects*, *knowledge about knowledge*, *and so on.* The problems in AI can be classified in two general types [3, 5], *search problems* and *representation problems.* On this last "peak", there exist different ways to reach their summit. So, we have [4] *Logics, Rules, Frames, Associative Nets, Scripts*, and so on, many times connected among them. We attempt, in this paper, a panoramic vision of the scope of application of such representation methods in AI. The two more disputable questions of both modern philosophy of mind and AI will be perhaps the *Turing Test* and the *Chinese Room Argument.* To elucidate these very difficult questions, see our final note.

Keywords: Knowledge Representation, Heuristics, Rough and Fuzzy Sets, Bayesian Networks, AI in Medicine.

1. Representation Problems

We can use a series of resources [4, 5] to approach the problems in A I, *Logic, Rules, Associative Nets, Frames* and *Scripts.* The election between these methods depends on the particular characteristics of the problem and also on our expectation about the type of solution [3]. In many cases, we take two o more tools at a time, as in the case of the Frame System, with the participation of Rules, and so on.

2. Rules

Concerning the usual way of appearance of Rules [3, 4, 5], as ***RBS*** (acronym of *Rule Based Systems*), we need four elements: *Interface of Usuary (IU)*, which will be very useful for the interaction with the usuary; *Motor of Inference (MI)*, responsible for the control of the flow of information between the modules; *Base of Facts* or *Base of Affirmations (BF or BA)*, that contains the initially known facts and those created during the process; and *Base of Knowledge (BK),* which contains the Rules used for the Representation of knowledge, into a determined Domain. There exists a two-way flow: *from the MI to IU,* and *from MI to BA,* but only one between *BK* and *MI,* and never in the reverse sense, except if we accept the system´s capacity of learning.

3. Inference in SBR

Such *Inference* consists in establishing the certainty of some statement, from the disposable information into Base of Facts and Base of Knowledge. We can use two methods, *concatenation going forward*, or *concatenation going backwards.* In the first case, we depart from Rules with verified affirmations in their antecedent, advancing through the affirmations which we find in their consequents. Whereas in the second case, we depart of Rules verified in certain consequents (all the consequents must be also verified in this sense), and we turn back to the antecedent. This converts its affirmations in new sub-objectives for the proof, searching for Rules where appear in their consequent, and so on.

The Rules shows a great advantage compared to the Classical Logic [5]. In Classical Logic, as you known, the Reasoning was Monotonic, with inferences without contradiction with the pre-existing, in RBS. Nevertheless, in the RBS, we may delete facts or affirmations from the Base of Facts, according to the new inferences. This makes the Reasoning Non-Monotonic, because we can modify the conclusion. Then, a question arises: what should we do with the conclusion of the affirmation now invalided? For this problem [3], we need to introduce the concept of *Type of Dependence of a Rule*, which can be *reversible:* if we delete the affirmations, then we automatically delete the above inferred facts, or *irreversible:* the facts, once inferred, are not deleted or changed.

And in the case of some applicable rules to time, which of them should be first executed? Such Rules constitute, for each step, the *Conflict Set* (obviously, a dynamic set). The subjacent decision problem is called *Resolution of Conflicts* or *Control of Reasoning.* There exist different strategies, to elect a Conflict Set, such as *Ordering of Rules*; *Control of Agendas*; *Criterion of Actuality*, and *Criterion of Specificity.* About the first and the second, the commentaries are unnecessary: they consist in the disposition of the Rules in the order as must be executed. The *Criterion of Actuality* consists in applying first the Rules in whose Antecedent there exists the up-to-date information. The Motor of Inference must be charged of the control of their respective moments. The *Criterion of Specificity* leads to executing the more specific Rules first, that is, those with more facts in their antecedent. So, between R_1*: if a, then b*, and R_2*: if a and d, then c*, we must select R_2, because it is more specific than R_1.

We also have need the *Mechanisms of Control in RBS.* So, by using*Mechanism of Refractority:* we are prevented from executing again a Rule already utilized, unless there is no other information such (in general, anomalous) case; *Rule Sets*: they allows the activation or neutralizing of Block´s Rules; *Meta-Rules*: they are rules which treat (or reason about) other Rules. Such Meta-Rules can collaborate in the Control of Reasoning, with the change or assignation of priorities to different Rules, according to the evolution of the circumstances.

4. Frames

They constitute the most general and most integrating method among all the Representation Procedures [3, 5]. They us to introduce some different elements. For instance, Rules in Frame Systems. Such System are usually denoted as *FS.* We must distinguish between *Facets,* as properties of the Field, and *Devils,* as procedures associated to the Frame System.

5. Scripts

They are structures of knowledge [1, 3, 4, 5] which must organize the information relative to dynamical stereotyped situations, that is, always, or almost always identical sequence of steps, or at least very similar. For instance, going to a certain cinema or a certain big store. The words and the subjacent ideas remind one of movies.

The *elements of a Script* can be *Scenes, Roles, Objects, Places, Names, Conditions, Instruments,* and *Results.* Its signification is obvious from the name: for instance, the Scenes must be events described sequentially, each scene being necessary for the realization of the subsequent one. With Results, we mention the facts we have obtained, once we have finished the sequence described in the Script.

6. Searching Methods

We will distinguish between Blind Search Procedures and Heuristic Procedures. In the first case, the oldest of them, it is possible to apply *Breadth Search* and *Depth Search.* But the problems associated to Combinatorial Explosion occur when the ramification index, or branching factor (the average cardinal of the successors of each node) increase beyond reasonable bounds.

For this reason, a more efficient search procedure is required, such as the introduction of *heuristic functions,* which give the estimation of the distance among the actual node and the final node. Thus, we chose to cal it *Heuristic Search.*

7. Introduction to Fuzziness

We define the "world" [4] as a complete and coherent description of how things are or how they could have been. Often, in problems related to the "real world", which is only one of the "possible worlds", the Monotonic Logic does not work often. But such a type of Logic is classically used in formal worlds, such as Mathematics. It is a real problem, because the "common sense" logic is non-monotonic, and this is our usual logic. We can see the more essential foundations of Fuzzy Theory in books as [2], [3] or [10].

An element of the Universe, U, can belong more or less to an arbitrary set C. It can belong to C in different degrees. From 0, when it does not belong at all to C, to 1, when it belongs totally to C. Or in any intermediate degree, like: 0.5, 0.3, 0.1..., but always between 0 and 1, both values included in their range. Such "membership degree" value can be assigned by an adequate "membership function", whose range is the closed unit interval, [0, 1]. So, the application can be expressed as in [7], by $f: C \rightarrow [0, 1]$. But information is given about the "membership degree", of such element, x, of the universe U, to the set C. In a Classical Set, therefore, the range of f should be reduced to the set {0, 1}. Given n universes of the discourse, we define a fuzzy relation, R, as a membership function that associates each n-tuple, a value of the unit closed interval, [0, 1]. The fuzzy relation, R, can be defined through such "membership function". In this way, we have 0R, ..., (1/3) R, ..., (1/2)R, ..., 1R.

The Cartesian product of two fuzzy sets, F and G, will be a fuzzy binary relation, through the minimum between the membership degrees. Sometimes, it is very useful to symbolize each fuzzy relation as one matrix, where the entries can be any real number between 0 (not related) and 1 (totally related, or simply, related). There exists a clear analogy between the composition of fuzzy relations and the product of matrices. To show this connection, it is sufficient to establish the correspondence: one +, one max; and one • one min. For this reason, the composition of fuzzy relations can also be called "max-min matrix product". As a particular case of the previous definition for the composition between fuzzy relations, we can introduce the composition between a fuzzy set and a fuzzy relation.

This can be very useful in the "Fuzzy Inference" [6], where we attempt to obtain new knowledge from the only already available. Obviously, in such a case, the fuzzy set can be represented by one row matrix, or a column matrix, depending on the order in the product.

The usual properties of the classical relations can be translated into fuzzy relations, but the transitive will be modified. R is *Reflexive:* R(x, x) = 1, for each x in the set C, into the universe. According to this, each element would be totally related with itself, when R is reflexive. R is *Symmetric:* If R(x, y) = R(y, x), for each pair (x, y). Therefore, the principal diagonal acts as a mirror, in the associated matrix. R is *Transitive:* Not in the usual way for relations or associated matrices, but now it is that $R(x, z) \geq \max(\min\{R(x, y), R(y, z)\})$ occurs.

All these mathematical methods can be very useful in Fuzzy Logic and in many branches of Artificial Intelligence. We can introduce new generalized versions of Classical Logic. So, *Modus Ponens Generalized, Modus Tollens Generalized* or *Hypothetic Syllogism.*

To each Fuzzy Predicate, we will associate a Fuzzy Set, the defined by such property, that is, composed by the elements of the Universe such that totally or partially verify such condition. For example, we can prove that the class of fuzzy sets with the operations ∪, ∩ th membership degree that belongs to the open unit interval.

8. Roughness

The concept of *Rough Set* was introduced by the Polish mathematician Zdzislaw Pawlak in 1982. Some theoretical advances with the corresponding applications have been emerging since then [2]. It is possible to apply Rough concepts to astonishing purposes, as will be the prediction of financial risk, but also in voice recognition, image processing, medical data analysis and so on.

Taking object, attribute or decision values, we will create rules for them: upper and lower approximations and boundary approximation. Each object is classified in one of these regions. For each rough set, A $\subseteq$ U, we dispose of *Lower Approximation of A*, as the collection of objects which can be classified with full certainty as members of A, and *Upper Approximation of A*, as the family of objects which may possibly be classified as members of A. Obviously, this class is wider than the aforementioned, containing between both the Rough set. *Rough Set Theory is a model of Approximate Reasoning.* According to this, we will interpret knowledge as a way to classify objects. We dispose of U, the universe of discourse, made up of objects, and an equivalence relation on U, denoted R. The procedure is to search for a collection of subsets in U (categories), so that all the elements of the same class possess the same attributes with the same values. So, we obtain a covering of U by a set of categories.

The elementary knowledge is encoded in a pair *(U, R),* made up of *"elementary granules of knowledge".* They constitute a partition in equivalence classes, into the quotient set, U/R. Given two elements, it is possible to determine when they are mutually indiscernible. In this case, we call this the *Indiscernibility Relation.* Therefore, it is possible to introduce the application which assigns to each object its corresponding class. Then, such indistinguishability allows us to introduce the *Fibre of* a_R , defined by the aforementioned relation R. So, the collection of such fibres, in the finite case, produces a union: this union of fibres is called a *granule of knowledge*. The pair (U, R) will be a *Knowledge Base*.

We say that an object, or category, is R-rough, if it is not R-exact. For each R-rough set, Y $\subseteq$ U, we define two associate R-exact sets, the *R-lower approximation of Y,* and the *R-upper approximation of Y.* So, we can represent the Rough set, Y, through the pair ($\underline{R}$Y, $\overline{R}$Y). Observe that $\underline{R}$ Y $\subseteq$ Y $\subseteq$ $\overline{R}$ Y, and furthermore, Y is R-exact $\Leftrightarrow$ $\underline{R}$ Y= $\overline{R}$Y.

Given a Knowledge Base, K ≡ llection of classes E_K = *{R - exact sets on U}*, which is closed with respect to the usual set operations $\cup$, $\cap$ It verifies the known properties of a Boolean Algebra. More concretely, we can say a *Field of Sets.* But it is not the case when we deal with R-rough sets. Because, for instance, the union of two R-rough sets can be a R-exact set. The coincidence of this Rough Set Theory with the Classical Theory of Sets occurs when we only work with R-exact sets.

An interesting generalization of Rough Set will be the *Generalized Approximation Space,* denoted *GAS.* It consists of a triple: (U, I, ν), where *U* will be the *Universe; I,* the *uncertainty function, I: U → P(U)*, and ν the *Rough Inclusion Function.* An example of this type of Rough Inclusion Function will be the *Generalized Rough Membership Function.* So, given any subset, we have both GAS - approximations, *lower-approximation* and *upper-approximation.*

9. Comparison between fuzziness and roughness

These names us mislead into believing that they are referring to the same concept. But they are very different approaches to uncertainty in the set of data. It depends of the nature of vagueness in the problem, or the convenience in applications. Both resources cover distinct aspects of the Approximate Reasoning. For this reason, both paradigms address to solve the Boundary Problem in Non-Crisp cases. Dubois and Prade [2] establish the mutual relationship between *Rough Fuzzy Set* and *Fuzzy Rough Set.* In the first case we will pass from fuzzy sets, through filtering, by the classical equivalence relations to quotient spaces, which are fuzzy sets. Whereas, in the second case we imitate the rough set approximation, but now with fuzzy similarity (instead of equivalence) relations.

We work into the collection of fuzzy sets on U, endowed with the operations: max and min. So, {Fuz (U, [0, 1]), max, min}. This produces a Zadeh Lattice. And it provides the path to complementary operator, {Fuz (U, [0, 1]), max, min, c}. It will be a Brouwer-Zadeh Lattice.

This lets us introduce the Rough Approximation to Fuzzy Sets. Our actual purpose is double: given A$\in$Z (U), we can induce a fuzzy set in U / R, by A, and reach the approximation of A, relative to R, according to the Rough Set Theory.

The notion of Fuzzy Rough Set is dual to the above concept. We consider newly the family of fuzzy sets in the universe U, with values in the closed unit interval, Fuz (U, [0, 1]). We need to analyze the fuzzy notion of equivalence relation and then, the fuzzy partition induced. Regarding the equivalence relation, the closest concept is the T-Fuzzy Similarity Relation.

In the past, the relationship between Fuzzy and Rough concepts were studied by some mathematicians and computer scientists, as Pawlak, Nakamura, Pedrycz [6], Dubois and Prade, Pal and Skowron, and many others.

10. Networks

As far as Nets are concerned, the more actual studies to deal with Bayesian Nets, also called Belief Networks [3, 5]. Before their apparition, the purpose was to obtain useful systems for the medical diagnosis, by classical statistical techniques, such as the Bayes´s Rule or Theorem.

A ***Bayesian Net*** is a pair (G, D), with G a directed, acyclic and connected graph, and D a distribution of probability (associated with the participant variables). Such distribution, D, must verify the *Property of Directional Separation,* according which the probability of a variable does not depend upon their not descendant nodes.

The *Inference in BNs* consists in establishing on the Net, for the known variables, their values and or the unknown variables, their respective probabilities. The objective of a Bayesian Network, in Medicine, is to find the probability of success with which we can give an exact diagnosis, based on known symptoms. We need to work with the following *Hypothesis: Exclusivity, Exhaustivity*, *and Conditional Independence.* According to the *Hypothesis of Exclusivity,* two different diagnoses cannot be right at the same time. With the *Hypothesis of Exhaustivity,* we suppose at our disposition all the possible diagnosis. And by the *Conditional Independence,* the thing found must be mutually independent to a certain diagnosis.

The initial problem with such hypothesis was the usual: their inadequacy to the real world. For this, we need to introduce the Bayesian Networks (BNs). In certain cases, as in the vascular problem of the predisposition to heart attack, there already exist already reasonable Systems of Prediction and Diagnosis, such as the DIAVAL Net. From these procedures springs a new and useful sub-discipline called *Medical artificial intelligence* (MAI, in acronym).

There are many different types of clinical tasks to which Expert Systems can be applied, as *Generating alerts and reminders*; *Diagnostic assistance* ; *Therapy planning*; *Agents for information retrieval*; *Image recognition and interpretation*, and so on. In the fields of treatment and diagnosis, A I possesses very important realizations, giving us for instance the following tools: PIP (1971), at MIT; MYCIN (1976), a Rule-Based System, due to Stanford University, works on infectious diseases; CASNET (1979) is due to Rutgers University and works on ophtalmological problems; INTERNIST (1980), due to Pittsburgh, on inner medicine; AI/RHEUM (1983), at Missoury University, on Reumathology; SPE (also 1983), at Rutgers, analyses the electrophoresis of proteins; TIA (1984), at Maryland, on the therapy of ischemic attacks, and many others.

11. Turing Test and Chinese Room Argument

Finally, we proceed at analyzing both of the announced questions, *Turing Test*, and *Chinese Room Argument.*

The Turing Test (TT). An interrogator is connected to one person and one machine, via a terminal, and therefore cannot see their counterparts. Its task is to find out which of the two candidates will be the machine, and which will be the human, only by asking them questions. If the interrogator cannot make a decision within a reasonable time, then the machine is considered to be intelligent. The most important argument against the TT is that it indeed provides only a test for human intelligence.

The Chinese Room Argument. John Searle´s argument [5] is intended to show that *implementing a computational algorithm that is formally isomorphic to human thought processes cannot be sufficient to reproduce thought.* Something more is required. So, it will be considered a refutation of both, Turing Test and Functionalism. It begins with this hypothetical premise: Suppose that A I research has succeeded in constructing a computer that behaves as if it understood Chinese. It takes Chinese characters as *inputs,* and produces other different characters, which it presents as *output,* by following the instructions of a computer program. Its attempt with this argument is to refute a certain conception of the role of computation in human cognition. To understand this argument, it will be necessary to distinguish among *Strong A I,* and *Weak (or Cautious) A I.* According to the first of them, any system that implements the right computer program with the right inputs and outputs thereby has cognition in the same sense as human beings. According the second of them, the computer is nothing more than a useful tool in studying human cognition, as in studying many other scientific domains.

The contrast is that according the Strong version, *the correct simulation <u>is</u> really <u>a mind.</u>* Whereas according the weak version, *the correct simulation <u>is only a model of the mind</u>*.

Its proof contains three premises and one conjecture,

AXIOM 1: *Implemented programs are syntactical processes,* i. e., computer programs are formal (syntactic).
AXIOM 2: Minds have semantic contents, i. e., human minds have mental contents (semantics).
AXIOM 3: *Syntax by itself is neither constitutive of nor sufficient for semantics.*
CONCLUSION: The implemented programs are not constitutive of, nor sufficient for minds.

Therefore, according to Searle [5], *Strong A I is false.* But the Chinese Room Argument may be expressed by two basic principles, each of which would be stated in four words:

1st) Syntax is not Semantics. Because syntax by itself, is not constitutive of semantics, nor by itself sufficient to guarantee the presence of semantics,
and
2nd) Simulation is not duplication.

We cannot describe what the machine is doing as "thinking". And neither does the human being understand a word of Chinese. Therefore, we must infer that computer does not understand Chinese either.

Thus, Searle concludes that *Strong (but not Weak) AI is a mistake.*

References

[1] Barr, A., & Feigenbaum, E. A. (1981). *The Handbook of Artificial Intelligence,* Volumes 1-3. William Kaufmann Inc.
[2] Dubois, D., & Prade, H. (Eds.) (2000). *Fundamentals of Fuzzy Sets*. Dordrecht, Netherlands: Kluwer Publ..
[3] Fernández Galán, S., et al. (1998). *Problemas resueltos de Inteligencia Artificial Aplicada. Búsqueda y Representación.* Madrid: Addison-Wesley.
[4] Garrido, A. (2004). *Logical Foundations of Artificial Intelligence*. FotFS V, Institut fur Angewändte Mathematik, Friedrich-Wilhelm Universität, University of Bonn.
[5] Mira, J., et al. (1995). *Aspectos Básicos de la Inteligencia Artificial.* Madrid: Editorial Sanz y Torres.
[6] Pedryck, W. (1993). *Fuzzy Control and Fuzzy Systems*. New York: Wiley and Sons Co..
[7] Szczepaniak, P. S.; Lisoba, P. J., & Kacprzyk G. (2000). *Fuzzy Systems in Medicine.* Heildelberg: Physica Verlag.
[8] Turing, A. (1950). *Computing machinery and intelligence.* Mind 59, 433-60.
[9] Wiener, N. (1947). *Cybernetics.* New York: MIT Press and John Wiley.
[10] Wang, Z. (1997). *A Course in Fuzzy Systems and Control*. New Jersey: Prentice-Hall.