RIA

Revista Iberoamericana de Argumentación

ἐπεὶ δὲ ταύτην τὴν ἐπιστήμην ζητοῦμεν

Programming Synthetic Innovation?

Jordi Vallverdú Segura
Philosophy Department
Universitat Autònoma de Barcelona
E08193 Bellaterra (BCN), Catalonia, Spain
Jordi.vallverdu@uab.cat

RESUMEN

El presente artículo analiza cómo los sistemas de IA han estado proporcionando información sobre el mundo bajo dos métodos básicos: heurísticas y metaheurísticas. Tras un breve análisis sobre el significado de la innovación y los elementos fundamentales de un sistema cognitivo, el autor sugiere un nuevo enfoque para crear sistemas de IA capaces de innovar.

PALABRAS CLAVE: heurística, IA, innovación, metaheurística, programación, robot.

ABSTRACT

This paper analyzes how AI systems have been providing new data about the world, basically under two main approaches: heuristics and metaheuristics. After a short analysis about what innovation means and about the basic elements of any cognitive system that elaborates theories on information, the author suggest a new path to create AI systems that could be able to innovate.

KEYWORDS: AI, heuristics, innovation, metaheuristics, programming, robot.

Revista Digital de Acceso Abierto http://e-spacio.uned.es/ojs/index.php/RIA
Editada por el Departamento de Lógica, Historia y Filosofía de la Ciencia

1. INTRODUCTION

What are we talking about when we say "innovation"? This concept can be defined thus «*Innovation* is the application of new solutions that meet new requirements, inarticulate needs, or existing market needs».[1] Far from entering into a conceptual debate on the nature, details and meanings of this word, it is clear to us that an innovative practice implies the skill to create something different or new, at least for the users of this information. In the past, philosophers like Plato believed that nothing new could be created by human mind, because all real things preexisted in the ideal world, but this is a childish, outmoded, false and stupid view about the human world. Change is the law of the human universe, reinforced by a cognitive uniqueness that pushes humans towards continuous questioning and learning: neoteny and brain plasticity are cognitive intentional forces that drive human bodies (Gould, 1977; McKinney & McNamara, 1991). Cognition is the result of an evolutionary implemented morphology.

Anyhow, from my humble point of view there are two ways to obtain innovative results, and I'll use a metaphor establishing an analogy between knowledge and board games. Board games are the sum of two things: game pieces (including the board, that defines its limits and shape) and rules to operate with these pieces. According to this, any possible innovation in one movement can be the result of:

I. A recombination of pieces: following the same rules of a game, the recombination of pieces inside a game offers new results. The logic outcomes of this procedure explain why several authors reach similar results at a certain historical moment. Perhaps the combinations of the pieces guided under certain rules can be really enormous, but even in that case they are limited. The fulfillment of the Period Table of Elements followed such pattern: assuming certain properties of the atomic world, there were niches that had to be occupied by specific atoms. This is part of the predictive power of a good scientific theory. Here innovation is a new combination of concepts inside the accepted paradigm.

II. The creation of new rules and/or pieces: at some point of a research, existing or prevailing rules are not enough to solve a problem or even cannot be able to explain or predict it. Then it is necessary to introduce some new pieces that violate at a certain level the basic rules of the existing game or even to define new rules. This is what Thomas S. Kuhn called a *paradigm shift*. The

[1] From Wikipedia, http://en.wikipedia.org/wiki/Innovation, accessed on August 5th 2013.

Copernican revolution or Darwin's explanations about the evolution of life are examples of it. The point is: when is a paradigm shift necessary? How do their contemporaries know that they must abandon the general set of rules and concepts of their habitual sciences in order to trust and adhere to new ones? A simple anomaly, even a big one, is not enough to justify so drastic a change.

These are my two basic ideas about how innovations are produced, which can be understood with a more common vocabulary as *heuristics* (recombination of pieces) and *metaheuristics* (creation of new rules and pieces, and going beyond traditional optimization uses of the term in computer sciences). Henceforth I'll use these terms to talk about innovation in AI.

2. VARIABLES OF COGNITIVE INNOVATION

In any decision-making or problem-solving process (henceforth, DM/PS) several variables can be found that determine the kind of solution and consequences we can obtain. These may be summarized as:

a) Coherence: the strength of internal coherence between objects and rules in an innovative process will be stronger in heuristic approaches than in metaheuristic ones. The more coherent, the more optimal. Coherence is reinforced thanks to the global assured interconnections among elements of the process.

  a. Global coherence: when all the objects and rules under analysis fit perfectly at the same time.

  b. Local coherence: a local coherence is achieved but coherence bonds among local and global objects are not established, supposing that the whole system will be compatible despite of the advances obtained at local level.

b) Stability/reliability: a corollary of coherence is stability. Any ordered system will be stable and, consequently, less prone to changes that introduce uncertainty or chaos into it. There is a natural bias towards heuristics and against metaheuristics. Even in the case of metaheuristic approaches, a continuous change is not a good choice: instability is a hard price to pay.

c) Minimalism: as a rule of any meta/heuristics, it explains the economic necessity of using as few as possible resources to explain or make anything. As Johannes Kepler cleverly summarized: *natura simplicitatem amat* (nature loves simplicity). If minimalism is not followed, the whole system may create very complex tools that even in the case of obtaining good results, increase too much the use of resources and difficult the real scalability of the system. This was the problem of the Ptolemaic motto “saving the appearances”.

d) Certainty: for any system involved into DM/PS actions, there is the belief that the world will react following the known rules. Any too strict belief in this fact will make impossible a fast reaction to an unknown event/outcome. Expectation is part of this process.

e) Framing: is necessary that a DM/PS system may be able to react to the most important inputs (therefore, it needs to identify them among noisy signals), but at the same time an intelligent system can operate coherently and survive without an understanding of the environmental variables. This is the classic bottom-up approach in AI defended by Rodney Brooks (1990, 1991) and followed in our days by morphological computing studies (Pfeifer, Bongard & Grand, 2007).

f) Flexibility: to modify the necessary rules and objects to improve its performance.

g) Time reaction (time constraint): this is one of the most important variables in any DM/PS process.

h) Long-term activities: the system can allocate resources (working with a multilayer or subsumption architecture) to work on possible future outcomes.

i) Daily situations: for example, domestic robots will need to handle with home environments, which are highly unstable and under changes (people and objects moving, hundreds of actions being performed,…). These changes happen fast and require quick answers.

Once we have clarified the possible variables involved into a process that must take decisions and/or obtain new knowledge, it’s time to analyze the several approaches to learning and innovative strategies followed in AI during the last decades.

3. AI AND INNOVATION

There is a classic approach in AI that follows the ideas sketched in the definition of heuristics and that has created what was called "expert system". An expert system (henceforth ES) is a computer system that emulates the decision-making ability of a human expert and several ES have been successfully created since the beginnings of AI. They are based in the classic principles of a symbolic approach to intelligence, defined by the pioneer Herb Simon (1995). Let's see this approach, close to my definition of heuristics and historically the first approach to innovation. Subsequently, we will analyze the metaheuristic approach.

3.1. HEURISTICS AND AI

Between 1955 and 1956, Alan Newell and Herbert Simon wrote a program they called "Logic Theorist". It's main purpose was to prove automatically some of the theorems present at Russell & Whitehead's *Principia Mathematica.* Russell and Whitehead published between 1910 and 1913 their *Principia Mathematica*, in which they re-established the foundations of pure mathematics in logical terms (Flach, 2005), something not so useful for practical purposes if we consider the fact that both authors required 379 pages to justify the truth of '1+1=2' (in Volume I, §54.43 and completed in Volume II, §110.643). Logic Theorist proved 38 of the first 52 theorems (Ch.2), and the proof of theorem 2.85 was even more elegant than the original one by Russell & Whitehead. These results were presented at the Darmouth Conference, in the summer of 1956: this conference, leaded by John McCarthy, gave birth to the field of AI (in fact, McCarthy coined here the name of "Artificial Intelligence"). It was at the Darmouth Conference that the crucial results of the common research between Herbert Simon (and economist) and Allen Newell (a mathematician) were presented. They had created an heuristic theorem-proving program, using the computer JOHNNIAC, at RAND Corporation. Simon tells that after informing Russell by mail about these results, they received an ironic answer: «if we'd told him this earlier, he and Whitehead could have saved ten years of their lives. He seemed amused and, I think, pleased». (Stewart 1994). Logic Theorist can be considered the first Expert System (ES).

Next Summer, in 1957, the second event that changed the history of computational sciences, the Cornell Summer School in Logic (1957), took place: there were plenty of researchers attending this course (Martin Davis, Hilary Putnam, Paul Gilmore, Herbert Gelernter,....) Gelernter, from IBM and a heuristic enthusiast, was provoked by Abraham Robinson and gave a lecture on these methods applied to proof seeking field. His lecture influenced deeply Gilmore, David and Putnam, who wrote their Herbrand-based proof procedure programs. This technique led to the so-called "Property B Method".

Without the aim of developing an exhaustive history about AI debates,[2] I consider it necessary, at least, to offer here some brief historical notes on its basic schools. Following the historical approach of Robinson (2000), there were two basic AI approaches:

a) The MIT View: they considered AI as a heuristic, procedural, associative way of producing artificial-generated knowledge. Marvin Minsky or Seymour Papert were members of this approach. For these authors, formal logic was inadequate for the representation of knowledge required by any general approach to AI. They considered it a too much static and rigid view, preferring a procedural one.

b) The Edinburgh-Stanford View: on the other hand, we could find the logical view, leaded by John McCarthy, who considered that AI knowledge could be mechanized because it could be axiomatized declaratively using First Order Logic. They considered computational logic as the only way to achieve an Artificial Intelligence.

To be honest, both approaches were highly symbolic and had more in common than differences we could find among them. In the middle of a AI's civil war, they were also called *neats* (logicists) and *scruffies* (proceduralists). It was later that the two AI really confronted approaches appeared, which can be summarized as *top down* and *bottom up* approaches (Vallverdú, 2006):

i. Top Down: symbol system hypothesis (Douglas Lenat, Herbert Simon). The *top down* approach constitutes the classical model. It works with symbol systems, which represent entities in the world. A reasoning engine operates in a domain independently on the symbols. SHRDLU (Winograd), Cyc (Douglas Lenat) or the several examples of successful expert systems are examples of it.

ii. Bottom Up: physical grounding hypothesis (situated activity, situated embodiment, connexionism). On the other side, the *bottom up* approach (leaded by Rodney Brooks), is based on the physical grounding hypothesis. Here, the system is

[2] It has been discussed with more detail at: Vallverdú, J. (2006) "Choosing between different AI approaches? The scientific benefits of the confrontation, and the new collaborative era between humans and machines", *TripleC*, 4(2): 209-216.

connected to the world via a set of sensors and the engine extracts all its knowledge from these physical sensors. Brooks talks about "intelligence without representation": complex intelligent systems will emerge as a result or complex interactive and independent machines.

In this sense, the MIT View and the Edinburgh-Stanford View both belonged to the *top down* approach. This classic rule-follower method operating with symbols has provided incredible machines (DENDRAL, MYCIN, PROSPECTOR,…) able to perform accurately and with great precision not only tedious tasks (mapping human genome or finding meaningful chemical QSAR/QSPR structures, see Vallverdú, 2011), but also great intellectual results (playing chess better that great masters – like Deep Blue-, or solving automatically really complex mathematical proofs – like EQP).[3] But there is a different approach to computer or artificial heuristics that I'll analyze in the next section.

3.2. METAHEURISTICS AND AI.

Although I've employed a different meaning of the term "metaheuristics", in AI and computer sciences research this word already existed defining[4]

> a procedure designed to find a good solution to a difficult optimization problem. Metaheuristics make few assumptions about the optimization problem being solved, and so they are usable for a variety of problems. Compared to simpler heuristics, metaheuristics are more abstract procedures that use low-level heuristics or search algorithms; thus, metaheuristics use concrete heuristics (or algorithms).

From the point of view of my initial definition, this classic notion of "metaheuristics" is still under the notion of strict rules applied to a limited number of objects: a finite and determined working memory, with a rigid inference engine and a static agenda prioritizing certain rules from the limited universe of the database.

A good heuristic tool makes also the system blind to a necessary update, like it usually happens in human domains with overspecialized experts: they know a lot about their research field but cannot understand important issues about different fields that could contribute to their advancement. The revolution of biological sciences during the second half of 20th century cannot be understood, for example without taking into account the new ideas introduced by experts in Physics who followed the ideas of

[3] A good list of Automated Theorem Provers can be found at: http://en.wikipedia.org/wiki/Automated_theorem_proving. Accessed on August 8th, 2013.
[4] From: http://en.wikipedia.org/wiki/Metaheuristic, accessed on August 9th 2013.

Schrödinger present in *What is Life?* (1944) and applied them both instrumentally and theoretically to Biology. Here, there is a gap, a jump from one field to another one, which was totally impossible for classic rule/concepts-followers of both disciplines. This is what I've tried to point out when talking about paradigm shifts. But there is a second question: the solution to unexpected problems (looking for autonomous machines inside dynamic environments), epitomized by the frame problem (Hayes, 1971).

3.2.1. EVOLUTIONARY INNOVATION

One solution to go beyond artificial systems that were enslaved by old-fashioned or inefficient rules and patterns[5] was to design systems able to evolve or even to learn by imitation (or behavioral cloning). Implementing stochastic optimizations, some metaheuristic systems obtained what was not a globally optimal solution but a large set of feasible solutions, using less computational resources than algorithms, iterative methods or simple heuristics. A large number of them are biologically inspired: genetic algorithms, genetic programming, evolutionary programming, differential evolution, evolution strategies, swarm intelligence, ant colony optimization algorithms,… All these different approaches try to use successful strategies that natural evolution has employed and which they have tried to apply or adapt to an artificial entity. Evolutionary algorithms, for example (Singh, 2006):[6]

- work with a population of candidate solutions and not a single point,
- work with coding of parameters instead of parameters themselves,
- do not require any domain knowledge (gradient information etc.) and just use the payoff information,
- are stochastic methods, i.e., use probabilistic transition rules and not deterministic ones,
- apply to a variety of problems and do not work in just a restricted domain.

These changes, from heuristics to metaheuristics, are incredibly useful and efficient but are still operationally-oriented, that is, *they are not self-referring.*

This last mentioned possibility, to make the very tools and ideas evolve, is still

[5] Even in the case of bootstrap larning, it is not enough to reach an optimal human-like cognitive flexibility. See Kohonen, T. (1997) *Self-Organizing Maps*, USA: Springer.
[6] From: http://www.southasianuniversity.org/~vivek/RTU-talk1.pdf, accessed on August 8th, 2013.

not applied in AI. Anyhow, what does innovation in this sense imply for learning purposes? The next section tries to explain this.

3.2.2. INNOVATION AS "LEARNING TO LEARN"

After the creation of the Perceptron, by Frank Rosenblatt in 1957, the first artificial neural network, several strategies emerged as ways to improve the learning capacities of these Neural Networks (henceforth NN). Supervised, unsupervised or semi-supervised methods were created to create better NN. These three approaches are also applied by humans to educate their children as well as to train people in any specific skill (sometimes combining some of these methods).

New ideas can emerge from combinations of existing pieces of information and well-known rules, but their number is limited to the possible outcomes that this game makes possible. Nothing more, nothing less. The role of AI systems dealing with extraordinary amounts of data with great store and calculus power has been extremely successful in the area of expert systems. Their incredible achievements in some fields are even beyond human cognitive capacities, but at the same time it is only a scalability question: to perform informational activities inside a domain of specific rules about what is a piece of information and how it can be related to other pieces.

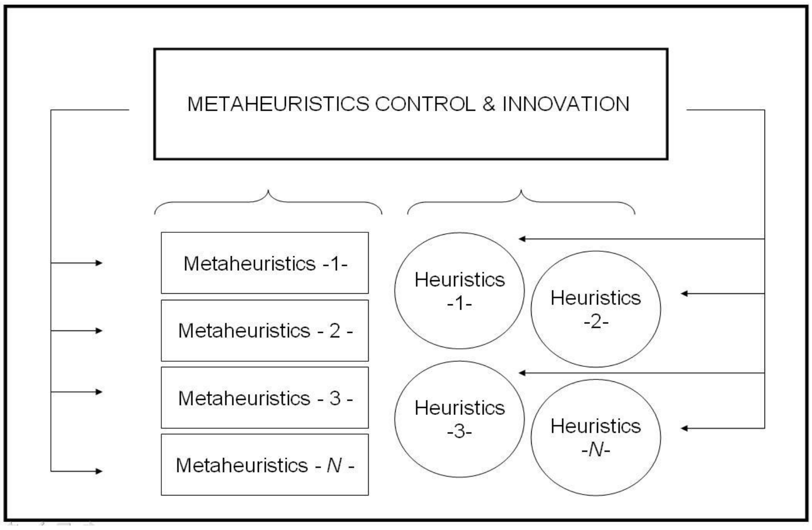
Until now we have had not a Copernicus, Galileo, Newton, Nietzsche, Wittgenstein or Einstein-like machine, able to discover new ways of thinking. This process of paradigm shifting is scarce among humans, and implies a great intellectual (and social) cost. Will machines ever be able to do it? I'm not referring myself to the very often apocalyptic notions of singularity, but I'm talking about how and under which mechanisms (some) humans are able to create innovative points of view from a conceptual perspective: new concepts, new instruments, new realities to be understood. This is the true meaning of innovation in AI: to run skills to redefine what is an observation or a significant input and how to use rules to infer better information. Will AI systems be able to create new statistical methods? Or will they be able to design new instruments to illuminate different sides of reality?

4. FROM KUHN TO IPHONES AND END REMARKS

There are two basic ideas that connect innovation and AI: *self-reference* and *paradigm-shift*. Classic heuristic approaches have been extremely successfully applied to several knowledge fields and metaheuristics have introduced evolutionary, multiple and stochastic approaches to problem-solving activities. But a true innovative AI system will be able to make something that some brilliant humans have been able to do: discover new ways to create and validate knowledge. Can creativity and innovation be programmed? Yes, I think so. Only one thing is needed, to design specific machines that are able to be at the frontier of contemporary "common sense" (scientific, epistemological). These machines will need to integrate cognitive subsumption architectures while, at the same time, designing new ways to understand reality coherently. I'm not talking about artificial systems that design possible *deus ex machina* models, conceptual toys for lazy scientists, but about systems that design new meanings for existing reality or even better, that help us to discover new aspects of our surrounding reality.

There is a second point: we need to improve the semantic habilities of computer systems in order to create a *metaexpert system*, able to integrate the specific knowledge of most human knowledge domains so that it would be possible for it to discover new relationships among different fields. This is a Big Data and a Big Cognition problem, impossible for humans due to their cognitive capacities, not only as individuals but also working together. At a certain level, this would be only a metaview of existing knowledge, a classic heuristics problem, but at the same time, a global perspective on the existing ideas could open the door to new ways to understand reality. How to evaluate different evidences that come from different disciplines and methods will be really interesting, and could make possible a *holistic science*.

How can we integrate all these ideas? My answer is that it is necessary to create modular systems that combine classic heuristic and metaheuristic approaches under a mixed architecture, with a third external metaheuristic layer that processes globally all the information and that can be able to reprogram strategies/work for the several modules at the same time and can also reprogram them automatically. A visual example of the suggested architecture could be thus:



Again, the necessary equilibrium between stability and innovation requires a specific analysis of computational nature that is not solved here, but only theoretically sketched. Finally, the way to drive the system from basic activities with well-known solutions to hypothetical research on information and rules should be emotion-like oriented by moods. To surpass efficiently *cul-de-sacs* or dead-end crazy approaches and make an optimized use of computational/database/energy resources, these machines should be ruled by an artificial emotional system that could help to add existential meaning to the information (for example: something as interesting, dangerous, helpful,…), something which is considered vital for human knowledge.

Concluding, in order to obtain innovation we need to implement combined uses of heuristics and metaheuristics approaches, guided by a superior general metaheuristic control system able to change the goals of several inferior levels and also to introduce changes in the way by which data are acquired and processed, under a emotion-like modulation system that adds meaning to the data while guiding the whole system towards an optimization according to dynamic circumstances. As the Latin motto says: *primum vivere, deinde philosophare*. In fact, the system would be oriented toward the continuous analysis of several strategies and the selection of more efficient ways (although only local) to achieve results, becoming an evolutionary self-learning, self-programming system.

REFERENCES

Brooks, Rodney A. (1990). "Elephants Don't Play Chess". *Robotics and Autonomous Systems* 6: 3-15

Brooks, Rodney A. (1991). "Intelligence without representation". *Artificial Intelligence* 47: 139-159.

Flach, P.A. (2005). "Modern Logic and Its Role in the Study of Knowledge", in Dale Jacquette (Ed.) , *A Companion to Philosophical Logic*, chapter 42: 680-693.

Gould, S.J. (1977). *Ontogeny and Phylogeny*. Cambridge: Belknap Press.

Hayes, P.J. (1971). "The Frame Problem and Related Problems on Artificial Intelligence". Stanford AI Project Memo *(AIM-153)*, Stanford, CA: Stanford University. Accesible at: ftp://reports.stanford.edu/pub/cstr/reports/cs/tr/71/242/CS-TR-71-242.pdf

Kuhn, Th. S. (1962). *The Structure of Scientific Revolutions*. Chicago: UCP.

McKinney, M. L. & K. J. McNamara (1991). *Heterochrony: the Evolution of Ontogeny*. NY: Plenum Press

Pfeifer, R. Bongard, R. & Grand, S. (2007). *How the body shapes the way we think: anew view of intelligence*. Cambridge, MA: MIT Press.

Robinson, J.A. (2000). "Computational Logic: Memories of the Past and Challenges for the Future", in J. Lloyd et al. (Eds.), *CL 2000, LNAI 1861*. Berlin: Springer Verlag, 1-24.

Simon, H. A. (1995). "Artificial Intelligence: an empirical science". *Artificial Intelligence* 77: 95-127.

Stewart, D (1994). "Herbert A. Simon: Thinking Machines". *OMNI Magazine,* Interview June 1994. At: http://www.astralgia.com/webportfolio/omnimoment/archives/interviews/simon.html.

Vallverdú, J. (2006). "Choosing between different AI approaches? The scientific benefits of the confrontation, and the new collaborative era between humans and machines".*TripleC* 4(2): 209-216.

Vallverdú, J. (2011). "An Epistemological analysis of QSPR/QSAR models". In Castro,E.A. & Haghi, A.K. (eds). *Advanced Methods and Applications in Chemoinformatics: Research Progress and New Applications*. USA: IGI, 318-332.

ACKNOWLEDGEMENTS: Financial support for this research was received from the Spanish Government's DGICYT research project: FFI2011-23238, "Innovation in scientific practice: cognitive approaches and their philosophical consequences".

J. VALLVERDÚ: B. Phil, B. Mus, S. Sci, Ph.D., is Lecturer Professor at UAB (Catalonia, Spain), where he teaches Philosophy and History of Science and Computing. His research is dedicated to the epistemological and cognitive aspects of Philosophy of Computing and Science. He is Editor-in-chief of the *International Journal of Synthetic Emotions* (IJSE), and as researcher is member of the EuCogIII and the Convergent Science Network of Biomimetic and Biohybrid Systems Net. He has written several books on computer epistemology and artificial emotions (one of them is the *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*). Very recently, he won a Japanese JSPS fellowship for his studies on HRI at Nishidalab (Kyoto University).