Evolving Towards an Evolutionary Epistemology

Luís Moniz Pereira
lmp@di.fct.unl.pt
http://centria.di.fct.unl.pt/∼lmp/
Centro de Inteligência Artificial - CENTRIA
Universidade Nova de Lisboa
Quinta da Torre
2829-516 Caparica, Portugal

Abstract: This work concerns a non-traditional approach to logic and epistemology, based on a challenging, albeit conjectural, articulation of views proceeding from Evolutionary Psychology and Biology, Non-Monotonic and Decision Logics, and Artificial Intelligence. The hinges to the latter inevitably suggest the emergence of an innovative symbiotic form of evolutionary epistemology, which will eventually have the ability to be shared, be it with robots, aliens or any other entity who must needs perform cognition to go on existing and program their future. Creating situated computers and robots means carrying out our own cognitive evolution by new means.

Keywords: Logic, Decision Logics, Epistemology, Evolutionary Biology, Evolutionary Psychology, Artificial Intelligence, Non-Monotonic Reasoning, Robots.

1 Evolution and the Brain

The first bipedal primates establish the separation between the human species and the other simians. To fathom the abilities of the human brain it is necessary to understand what exactly the problems our ancestor primates were were trying to solve that led them to develop such an extraordinarily intricate brain. We cannot look at the modern human brain, and its ability to create science, as if the millions of evolution-years which attuned it to its present configuration had never taken place. Among the eventual problems we certainly have those of status, territorialism, mating, gregariousness, altruism versus opportunism, the building of artifacts, and the mappings of the external world.

To Homo Sapiens Sapiens' brain, considered indistinguishable from current ones, we assign an estimated age of one or two hundred thousand years. The Palaeolithic started at about 60 or 30 thousand years before that, the period in which language, and much later writing, began to develop. By the Upper Palaeolithic times however, from 40,000 to 10,000 before the present, the tempo of cultural evolution quickened dramatically. According to the theory of population genetics, most of the change was far too fast to be tracked closely by genetic evolution.

As the psychiatrist must look at a patient's past in order to better understand him in the present, so must we look also at our species' past in order to understand our modern peculiarities. This stance is called Evolutionary Psychology — a fascinating field of study — born some 40 years ago. It is a consummate example of successful ongoing scientific unification, engendered by a deeply significant combination of Psychology, Anthropology, Archaeology, Evolutionary Biology, Linguistics, Neuroscience, and Artificial Intelligence (Buss 2005). Evolutionary Psychology has been studying the brain from the evolutionary perspective, thereby originating some extremely relevant contributions. In that perspective, it has been strongly supported by Anthropological Archaeology in its empirical study of the cultural evolution of mankind (Shennan 2002).

2 Evolutionary Psychology: Genes and Memes

In human life, we have two reproductive mechanisms: one is sexual reproduction, in which the replication unit is the gene; the other is mental reproduction. Some authors from Evolutionary Psychology have construed the notion of "meme", in complement and contrast to that of gene. A meme is that which substantiates a second reproductive system executed in the brain. It is the mental unit corresponding to the gene. Memes gather in assemblies, in patterns, similar to the way genes gather in chromosomes. Memes are patterned by ideologies, religions, and common sense ideas. Indeed, certain memes work well together, mutually reinforcing each other, others not, so that correcting (and correctional) mechanisms may be triggered.

We have a genetic reproduction system and, on top of it, Nature — through evolution — has created a second one, which we employ in pedagogy. We reproduce ideas: generally, good ideas propagate and replicate, being selected over the bad ones, although no one is around to guarantee it. Genes persist because they reproduce, and memes are the reproduction units associated with the brain — the brain being a reproductive organ. What we are doing, in schools and universities, is to reproduce knowledge. Educational systems consist of a means for "infecting" students with good memes, ideas having proven themselves able enough to self-reproduce and persist, while despising others that fail to survive. There are many educational systems variants, for instance madrasas.

When people interact they communicate ideas, and those which are infectiously good tend to reproduce. There are assemblies of ideas, sets of beliefs, which reproduce together. The memes in such memeplexes — like the genes in chromosomes — are in competition amongst themselves and also with the gene base. They exist because they are part of a reproductive mechanism which is necessary to achieve faster local adaptations, as genes take too long to reproduce with respect to the time scale of the individual bearing the memes. Thus the individual phenotype may be given more of a chance to reproduce its genotype. This leads directly to the meme-gene co-evolution.

Memes however could not spread but for the biologically valuable tendency of individuals to imitate, something afforded by the brain. There are plenty of good reasons why imitation should have been favoured by conventional natural selection working on genes. Individuals that are genetically predisposed to imitate enjoy a fast track to skills that may have taken others a long time to build.

Consequently, the brain and its accompanying mind are the result of a deep symbiosis, a genetic product influenced by the mechanism of memetic reproduction. In this faster system of adaptation we have reached the point of being able to predict our own memetic (and genetic) mutations, as necessary changes to prepare for the future by anticipating it. That is why we imagine the future — we create hypothetical scenarios, predict the possible futures, and choose to pursue some of them while avoiding others.

However, beyond simple reproductive success there are also pressing concerns in regard to social interaction. As communal beings, we need to develop some sort of status in order to be respected, copied, or obeyed. We must worry about territorial expansion and its defence, if we are to have descendants and, moreover, descendants with descendants. We need to sign contractual agreements with those who share our social and cultural ecology. There is also the important requisite of personal expression opportunity. If we do not express ourselves, no one will copy even our dearest memes, let alone our scientific theory memeplexes.

In this view, scientific thought emerges from distributed personal interaction, albeit it at a special and temporal distance, and never in an isolated way. It must be erected from several confluences, or in teams, as is the case in science. In truth, knowledge is not constructed in an autonomous way; rather it is engendered by networks of people. In science it is important to work as a team. The stereotype of the isolated and enlightened aristocratic scientist has been defeated for quite some time: science is institutionalized, organized and has proper methodologies, conferences. It is processed in appropriate environments, one being the educational one, in which we carry out memetic proliferation.

3 Specific Modules versus General Intelligence

Theoretical and field archaeologists, like Steven Mithen in (Mithen 1996), are bringing in historical and pre-historical evidence that our ancestors began with a generic intelligence, such as we find in apes. There has been a broad discussion — in fact reproduced within the Artificial Intelligence (AI) community — about whether intelligence is a general functionality or else best envisaged as divided into specific ability modules or components. When it first appeared, Evolutionary Psychology developed a trend, which Chomsky had begun in insisting on innate specialized areas for language processing in the brain, and it was generally accepted that a plethora of specific modules for a diversity of certain brain functions do exist. Indeed, in the beginnings of Evolutionary Psychology, the likes of Steven Pinker, Leda Cosmides, John Tooby, and David Buss, in consonance with AI's own vision of specific modules, believed all brain function to be founded on such modules — for language, for mating, religion, etc.

Meanwhile, archaeologists have demonstrated, via historical record, that human species went from a first phase of general intelligence to a second phase of three major specialized modules: one for natural history and naive physics (knowledge of Nature); one for knowledge and manufacture of instruments; and one for cultural artifacts, i.e. the rules of living in society and the very politics of coexistence. These three specialized intelligences were separated, and it is only at a newer stage — corresponding to Homo Sapiens, with the appearance of spoken language — that it becomes necessary to have a cupola module, articulating the other ones. How else do the major different specialized modules connect, and how can people communicate among themselves? That need gave birth to the generic cupola module, a more sophisticated form of general intelligence, the cognitive glue bringing the specialized modules to communicate and cooperate.

4 The Evolution of Reason: Logic

The formal systems of logic have ordinarily been regarded as independent of biology, but recent developments in evolutionary theory suggest that biology and logic may be intimately interrelated. (Cooper 2001) outlines a theory of rationality in which logical law emerges as an intrinsic aspect of evolutionary biology. This biological perspective on logic, though at present unorthodox, could change traditional ideas about the reasoning process (Hanna 2006).

Cooper examines the connections between logic and evolutionary biology and illustrates how logical rules are derived directly from evolutionary principles, and therefore have no independent status of their own. Laws of decision theory, utility theory, induction, and deduction are reinterpreted as natural consequences of evolutionary processes. Cooper's connection of logical law to evolutionary theory ultimately results in a unified foundation for an evolutionary science of reason. According to Cooper, today, in the general drift of scientific thought, logic is treated as though it were a central stillness. For the most part, the laws of logic are taken as fixed and absolute. Contemporary theories of scientific methodology are logico-centric. Logic is seen commonly as an immutable, universal, meta-scientific framework for the sciences, as well as for personal knowledge. Biological evolution is acknowledged, but it is accorded only an ancillary role, as a sort of biospheric police force, whose duty is to enforce the logical law among the recalcitrant. Logical obedience is rewarded and disobedience punished by natural selection, it is thought. All organisms with cognitive ability had better comply with the universal laws of logic on pain of being selected against!

Comfortable as that mind set may be, Cooper believes he is not alone in suspecting it has things backward. There is a different, more biocentric, perspective to be considered. In the alternative scheme of things, logic is not the central stillness. The principles of reasoning are neither fixed, absolute, independent, nor elemental. If anything, it is the evolutionary dynamic itself that is elemental. Evolution is not the law enforcer but the law giver — not so much a police force but a legislature. The laws of logic are not independent of biology but implicit in the very evolutionary processes that enforce them. The processes determine the laws.

If the latter understanding is correct, logical rules have no separate status of their own but are theoretical constructs of evolutionary biology. Logical theory ought then in some sense to be deducible entirely from biological considerations. To paraphrase, the hypothesis is that the commonly accepted systems of logic are branches of evolutionary biology. Indeed, evolution has provided humans with symbolic thought, and symbolic language communication abilities. Objective common knowledge requires thought to follow abstract, content independent rules of reasoning and argumentation, which must not be entirely subjective, on pain of making precise communication and collective rational endeavour impossible. Such rules have become ingrained in human thought, and hold an enormous joint survival value. In human cognitive evolution, both mimetic knowledge (such as that inherent in reality-simulating maps and models), and imitation knowledge (such as that present in ritual observation, or in artefact reproduction), were essential first steps towards socially situated, joint rule following behaviour, required by, say, hunting plans.

Decision theory is the branch of logic that comes into most immediate contact with the concerns of evolutionary biology. They are bound together by virtue of their mutual involvement in behaviour. The logic of decision is concerned with choices regarding the most reasonable courses of action, or behavioural patterns. Behaviour is observable, it is amenable to scientific prediction and explanation, and there is the possibility of explaining it in evolutionary terms. This makes behaviour an interdisciplinary bridge approachable from both the biological and the logical sides. Ultimately, behaviour is the fulcrum over which evolutionary forces extend their leverage into the realm of logic. Viewed through the lenses of biology, favoured behaviour is evolutionary fit. Through the lens of logic it is rational decision behaviour, according to rules for reasoning and rules for action (Cooper, 2001).

5 Reasoning Cupola

On the heels of rational group behaviour, throughout human cultures there emerged abstract rule following social games. Game rules encapsulate concrete situation defining patterns, and concrete situation-action-situation causal sequencing, which mirrors causality-obeying physical reality. From games, further abstraction ensued, and there finally emerged the notions of situation-defining concepts, of general rules of thought and their chaining, and of legitimate argument and counter-argument moves. Together they compose a cognitive meta-game (Holland 1998).

The pervasiveness of informal logic for capturing knowledge and for reasoning, a veritable *lingua franca* across human languages and cultures rests on its ability to actually foster rational understanding and common objectivity. Crucially, objective knowledge evolution dynamics, whether individual or plural, follows ratiocination patterns and laws. Furthermore, and more recently, the very same rules of reasoning can and are employed to reason about reasoning. Moreover, new reasoning methods can and have been invented and perfected throughout human history. Examples of these are transfinite induction, *reductio ad absurdum* (proof by contradiction), recursion, abduction, and contradiction removal, to name a few.

Though some reasoning methods are well known, some are still unconscious but, like the rules of grammar, can be discovered through research. Indeed, humans can use language without learning grammar. However, if we are to understand linguistics, knowing the logic of grammar, syntax and semantics is vital. Humans do use grammar without any explicit knowledge of it, but that doesn't mean it cannot be described. Similarly, when talking about the movement of electrons we surely do not mean that a particular electron knows the laws it follows, but we are certainly using symbolic language to describe the process, and it is even possible to use that description to implement a model and simulation which exhibits precisely the same behaviour.

New purported reasoning methods may be disputed, just like any specific train of reasoning can. But reasoning can only be disputed by further reasoning, if any consensus is to be found! (Nagel 1997). Some argue that scientific and philosophical discussion is necessarily a tantamount to a culture sensitive, and culturally relative, persuasive informal ad hoc argumentation, allied to anything goes rhetoric (criticized by (Gross & Levitt 1994)). They ignore that argumentation is just another form of reasoning which has itself been made the subject of logical formalization, and are oblivious to the fact that rhetoric may be fine for preachers, but is not conducive to the two-sided communication required to reach common agreement in the all rigorous scientific praxis that lead to cumulative knowledge.

Logic, we sustain, provides the overall conceptual cupola that, as a generic module, fluidly articulates together the specific modules identified by evolutionary psychology. In that respect, it is mirrored by the computational universality of computing machines, which can execute any program, compute any computable function.

The relationship of this argument to logic is ensured by the philosophical perspective of functionalism: logic itself can be implemented on top of a symbol processing system, independently of the particular physical substrate supporting it. Once a process is described in logic, we can use the description to synthesize an artefact with those very same properties. As long as it is a computational model, any attempt to escape logic will not prove itself to be inherently more powerful.

On the other hand, there is an obvious human capacity for understanding logical reasoning, a capacity developed during the course of brain evolution. Its most powerful expression today is science itself, and the knowledge amassed from numerous disciplines, each of which with their own logic nuances dedicated to reasoning within their domain. All that has been learned empirically about evolution in general, and mental processes in particular, suggests that the brain is a machine assembled not to understand itself, but to survive. Understanding the mind at work, then, needs to be brought about by the methods of science.

6 Dynamics of Logic

It is not too difficult to imagine how a combined process of rule generation, of systematic diagnosis, and of rule revision by updating, can be used to achieve automated theory learning, in an integrated way, within the uniform setting of logic programming. (For details cf. (Pereira 2002).)

To initiate the learning, one starts with some fixed, already acquired, background knowledge in rule form, i.e. a theory, plus a rule generator to add to it new purported knowledge, in order to explain abductively known observations, whether positive or negative, in the form of facts and of explicitly negated facts.

The goal is to generate rules that define a positive concept as well as its negated counterpart concept, so that they cover all known observation instances. This automatic generation of new rules is subjected to a pre-defined bias, i.e. only some rule forms, and predicates comprising them, are allowed in the generation process. Newly generated rules may contradict one another, on some of the observation instances, and so they must be subjected to a diagnosis, to identify alternative possible minimal revisions.

To decide which revision to adopt next, desirable new possible observations are conceived of, with respect to the ongoing available theory, whose results, if known, would allow us to decide among the competing revisions. The outcomes of these so-called crucial observations are then obtained, and adopted by the program soliciting them, or are feeded into the belief revision process to carry out further actions leading to additional observation results.

Once the desired rule revisions are selected on the basis of such results, and also of programmed preference criteria, the revisions are enacted by an update procedure. Mark that revised rules can themselves be subjected to later revisions if needed.

Indeed, the whole process will be iterated on the basis of new incoming knowledge, or by knowledge confrontation of among differently evolved automated theories, with distinct backgrounds, biases, rule generators, diagnosers, revisors, preferences, planners, observations, and updating procedures, comprising a rational agent.

Agent epistemic confrontation relies on argumentation and mutual debugging. Besides the legislative and legal domain, the province of scientific discussion too relies on such procedures, and can benefit from their automation.

Tackling argumentation involves a set of tools similar to those of diagnosis and debugging. Arguments can attack another argument's assumptions either directly, by proving the negation of an assumption, or indirectly, by contradicting a conclusion of other argument which rests on its assumptions. However, such attacks by an argument upon another can, in turn, be counter-attacked in the same way, and may in response counter-counter-attack, etc. logic programming has shown how this process can be studied and conclusions drawn about competing mutually contradictory arguments, and how they each can be revised to reach agreement. Cooperative argumentation is also an available option.

7 Epistemic Tools

The canonical definition of objective scientific knowledge avidly sought by the logical positivists is not a philosophical problem nor can it be attained, as they hoped, simply by logical and semantical analysis. It is an empirical question too, that can be answered only by a continuing probe of the possible functionality of the thought process itself and its physical basis. In some cases, the cognitive tools and instruments of rationality will be found hardware independent. Even then, the appropriateness of their use in specific real circumstances and goals will need to be empirically determined. There is no universal one-size-fits-all epistemological recipe, but agreement can be had on the relative success of any given tool kit.

In any case, partial understanding may also be sought by building intelligent machines, functionalism coming to the rescue when positing that the material substrate is often not of the essence, that it suffices to realize equivalent functionality albeit over different hardware. Moreover, distinct functioning roads to the same behaviour may be had, thereby accruing to our understanding of what general intelligence means, toward their symbiotic entwining, the most recent step in evolutionary epistemology. Functionalism can only make that more adroit.

The most fruitful procedures will almost certainly include the use of Artificial Intelligence, theory and technique, aided in due course by the still embryonic field of artificial emotion, to simulate complex mental operations. This modelling system will be joined to an already swiftly maturing neurobiology of the brain, including the high-resolution scanning of computational networks active in various forms of thought.

With this background in mind, and namely the discussion about specialized modules and general intelligence, I would like introduce at this point the informal notion of "cognome", by analogy with genome, standing for an individual's particular structural combination of cognitive memes.

When considering scientific knowledge, if the computer processing of the human genome is what leads us to Bioinformatics then, by analogy, we may state that the cognome will be the basis of a future "Cognotechnology", applicable in any science. This way, the future of AI is connected to the characteristic of it being an epistemological instrument, not only for an autonomous agent, but a symbiotic one which will help humans in performing science itself.

And I'm not just talking about data mining, pattern recognition, ontology building, although in these fields we can approach more structured aspects of epistemology. I'm thinking about that which every scientist does, which is to abduce, invent and prophesy theories, put them to the test, create experiments, draw conclusions to support additional observations, discuss these observations and his conjectures with other scientists.

Veritably, the capacity for cognition is what allows us to anticipate the future, to preadapt and imagine scenarios of possible evolutions — of the world and of ourselves as cognitive agents — to make choices, to use preferences about some hypothetical worlds and their futures, and meta-preferences — preferences on which preferences to employ and how to make them evolve. The activity of prospecting the future is vital and characteristic of our species and its capacity to understand the real world and ourselves, living in society, where distributed cognition is the normal and regular way to do science. Prospective consciousness allows us to pre adapt to what will happen. For that, a capacity to simulate, to imagine "what would happen if", i.e. hypothetical thinking, becomes necessary. Such thinking is indispensable in science; for it gives us the rules to predict and explain what will or can happen, without which technology would not be possible.

Lately, I've been working towards automating this capacity, by implementing programs which can imagine their futures, making informed choices about them, and then modify themselves to enact those choices — the inklings free will. We call it prospective computing ((Lopes & Pereira 2007)), and have applied it to moral reasoning case studies ((Pereira & Saptawijaya 2007)). Indeed, continuous developments in logic programming (LP) language semantics which can account for evolving programs with updates (Alferes, et al. 2002, Alferes, et al. 2000) have opened the door to new perspectives and problems amidst the Logic Programming and agents community. Since is now possible for a program to talk about its own evolution, changing and adapting itself through non-monotonic self-updates, one of the new looming challenges is how to use such semantics to specify and model logic based agents which are capable of anticipating their own possible future states and of preferring among them in order to further their goals, prospectively maintaining truth and consistency in so doing. Such predictions need to account not only for changes in the perceived external environment, but also need to incorporate available actions originating from the agent itself, and perhaps even consider possible actions and hypothetical goals emerging in the activity of other agents.

While being immersed in a world (virtual or real), every proactive agent should be capable, to some degree, of conjuring up hypothetical *what-if* scenarios while attending to a given set of integrity constraints, goals, and partial observations of the environment. These scenarios can be about hypothetical observations (what-if this observation were true?), about hypothetical actions (what-if this action were performed?) or hypothetical goals (what-if this goal was pursued?).

As we are dealing with non-monotonic logics, where knowledge about the world is incomplete and revisable, a way to represent predictions about the future is to consider possible scenarios as tentative evolving hypotheses which *may* become true, pending subsequent confirmation or disconfirmation on further observations, the latter based on the expected consequences of assuming each of the scenarios.

How does natural selection anticipate our future needs? Well, by creating a cognitive machine called brain that can create models of the world, and even of itself, and process hypotheticals much like a Universal Turing Machine can mimic any other Turing machine, just like any given computer can run any program. This plasticity provides for its universal versatility (cf. (Davis 2000)).

8 Mens ex-machina

It is worth mentioning that, in contrast to Turing, Gödel was not interested in the development of computers. However, their mechanics is so connected with the operations and concepts of logic that, nowadays, it is quite commonplace for logicians to be involved, in some way or other, in the study of computers and computer science. However, Gödel's famous incompleteness theorem demonstrates and establishes the rigidness of mathematics and the limitations of formal systems and, according to some, of computer programs. It relates to the persistent issue of whether mind surpasses machine. Thus, the growing interest given to computers and Artificial Intelligence (AI) has lead to a general increase in interest about Gödel's own work. But, so Gödel himself recognizes, his theorem does not settle the issue of knowing if the mind surpasses the machine. Actually, Gödel's work in this direction (Casti & DePauli 2000, Wang 1973, Wang 1987, Pereira 2007) seems to favour (instead of countering) the mechanist position (and even finitism) as an approach to the automation of formal systems.

Gödel contrasts insight with proof. A proof can be explicit and conclusive, for it has the support of axioms and of rules of inference. In contrast, insights can be communicated only via "pointing" at things. Any philosophy expressed by an exact theory can be seen a special case of the application of Gödel's conceptual realism. According to him, its objective should be to give us a clear perspective of all the basic metaphysical concepts. More precisely, Gödel claims that this task consists in determining, through intuition, the primitive metaphysical concepts C and in making correspond to them a set of axioms A (so that only C satisfies A, and the elements in A are implied by the original intuition of C). He further admits that, from time to time, it would be possible to add new axioms.

Gödel also advocates an 'optimistic rationalism'. His justification appeals to (1) "The fact that those parts of mathematics which have been systematically and completely developed ... show an astonishing degree of beauty and perfection." As such, (2) It is not the case "that human reason is irredeemably irrational by asking itself questions to which it cannot answer, and at the same time emphatically asserting that only reason can answer them." It follows that (3) There are no "undecidable questions of number theory for the human mind." So (4) "The human mind surpasses all machines."

However, the inference from (1) to (2) seems to be obtained from accidental successes in very limited fields to justify an anticipation of universal success. Besides, both (2) and (3) concern only a specific and delimited part of mind and reason which refer just to mathematical issues.

Gödel understands that his incompleteness theorem by itself does not imply that the human mind surpasses all machines. An additional premise is necessary. Gödel presents three suggestions to that end: (a) It is sufficient to accept his 'optimistic realism' (b) By appealing "to the fact that the mind, and the way it's used, is not static, but finds itself in constant development," Gödel suggests that "there is no reason to claim that" the number of mental states "cannot converge to infinity in the course of its development." (c) He believes that there is a mind separate from matter, and that such will be demonstrated "scientifically (maybe by the fact that there are insufficient nerve cells to account for all the observable mental operations)."

There is a known ambiguity between the notion of mechanism confined to the mechanizable (in the precise sense of computable or recursive) and the notion of materialist mechanism. Gödel enounces two propositions: (i) The brain operates basically like a digital computer. (ii) The laws of physics, in their observable consequences, have a finite limit of precision. He is of the opinion that (i) is very likely, and that (ii) is practically certain. Perhaps the interpretation Gödel assigns to (ii) is what makes it compatible with the existence of non-mechanical physical laws, and in the same breath he links it to (i) in the sense that, as much as we can observe of the brain's behaviour, it functions like a digital computer.

8.1 Is mathematical insight algorithmic?

Roger Penrose (Penrose 1994) claims that it is not, and supports much of his argument, as J. R. Lucas before him (revisited in (Lucas 1996)), on Gödel's incompleteness theorem: It is insight that allows us to see that a Gödel assertion, undecidable in a given formal system, is accordingly true. How could this intuition be the result of an algorithm? Penrose insists that his argument would have been "certainly considered by Gödel himself in the 1930s and was never properly refuted since then ... "

However, in his Gibbs lecture delivered to the American Mathematical Society in 1951, Gödel openly contradicts Penrose:

"On the other hand, on the basis of what has been proven so far, it remains possible that a theorem proving machine, indeed equivalent to mathematical insight, can exist (and even be empirically discovered), although that cannot be proven, nor even proven that it only obtains correct theorems of the finitary number theory."

In reality, during the 1930s, Gödel was especially careful in avoiding controversial statements, limiting himself to what could be proven. However, his Gibbs lecture was a veritable surprise. Gödel insistently argued that his theorem had important philosophical implications. In spite of that, and as the above citation makes it clear, he never stated that mathematical insight could be shown to be non-algorithmic.

It is likely that Gödel would agree with Penrose's judgment that mathematical insight could not be the product of an algorithm. In fact, Gödel apparently believed that the human mind could not even be the product of natural evolution. However, Gödel never claimed that such conclusions were consequence of his famous theorem.

Due to the current prevailing trend to restrict discussion about the limits of rationality, in contraposition to insight, to the well-defined and surprising advances in computer science technology and programming, the perspective of considering reason in terms of mechanical capabilities has received much attention in the last decades. Such is recognised as being core to the study of AI, which is clearly relevant to Gödel's wish of separating mind and machine.

In this stance, AI would be primarily interested in what is feasible from the viewpoint of computability, whose formal concern involves only a very limited part of mathematics and logic. However, the study of the limitations of AI cannot be reduced to this restriction of its scope. In this regard, it is essential to distinguish between algorithms for problem-solving, and algorithms simpliciter, as sets of rules to follow in a systematic and automatic way, which are eventually self-modifiable, and without necessarily having a specific and well-defined problem to solve.

9 Remarks on Future Work

How can the general reasoning issues explored here be approached, within a uniform framework, except in the context of logic? I submit indeed that logic is the only conceivable venue for the core of such an enterprise. And must we not do it? Must we not mechanize reasoning procedures so that robots and computers can perform, for us or with us, the sometimes tedious sometimes overly complex work involved in them? Surely, only so persisting will permit the grappling of the evermore demanding reasoning problems that face us.

In its struggle for that rigorous description, the field of Artificial Intelligence has made viable the proposition of turning logic into a programming language (Pereira 2002). Logic can presently be used as a specification language which is not only executable, but on top of which we can demonstrate properties and make proofs of correction that validate the very descriptions we produce with it. Facing up to the challenge, AI developed logic beyond the confines of monotonic cumulativity, far removed from the artificial paradises of the well-defined, and well into the real world purview of incomplete, contradictory, arguable, revisable, distributed and updatable, demanding, among other, the study and development of non-monotonic logics, and their computer implementation.

Over the years, enormous amount of work has been carried out on individual topics, such as logic programming language semantics, belief revision, preferences, evolving programs with updates, and many other issues that are crucial for a computational architecture of the mind. We are in the presence of a state-of-the-art from whence we can begin addressing the more general issues with the tools already at hand, unifying such efforts into powerful implementations exhibiting promising new computational properties. Computational logic has shown itself capable to evolve to meet the demands of the difficult descriptions it is trying to address.

The use of the logic paradigm also allows us to present the discussion of our system at a sufficiently high level of abstraction and generality to allow for productive interdisciplinary discussions, both about its specification and its derived properties.

AI, most especially through logic programming, will continue to accomplish a good deal in identifying, formalizing, and implementing the (descriptive) Laws of Thought. Most notably, AI has taken on the challenge of opening up logic to the dynamics of knowledge in flux. And in so doing, it has been progressively meeting our expectations and requirements. Continuing this work is essential if we are to cope, albeit with the aid of computers, with the challenges of evermore accumulated and distributed knowledge, in a changing world.

A wealth of investment in logic programming research has accumulated in the form of published results and built systems, over the more than 35 years since its inception. Some of these could not be brought to full fruition at the time because of technological costs and funding limitations, or more simply on because of the lack in availability of cheap and efficient requisite equipment, nowadays having become affordable and with better performance. Even though the theory and implementation know-how were there at the time, their full fruition was not viable then.

However, to-day, younger researchers are not well aware of such past results and of their potential, some from the early eighties, even when the implementation technology may have evolved to allow them to be put to use. And all the more so because the complexity of present systems makes such immediately available "theory technologies" all the more desirable. On the other hand, the lowering cost of hardware and the appearance on the scene of new "implementation technologies" make the prior investment in theory not only realizable now, but affordable and profitable as well.

Conversely, we should not stop or refrain from investing now for the future, for the "theory technologies" of today, in addressing real felt problems, will be the providers of tomorrow's implementation and application reapers.

For instance, were it not for the impressive development in Logic Programming semantics in the past 20 years or so, and its extensions to non-monotonic and other forms of reasoning, we would not have today the impressive, sound, and efficient system implementations of such semantics (be it the stable or well-founded varieties or their hybrid combination), which have been opening up a whole new gamut of application areas as well as a spate of sophisticated reasoning abilities over them. This was made possible only through successive and prolonged efforts at theoretical generalization and synthesis, and by way of the combined integration of theory, procedure development, and practical implementation.

The lesson, then, for further investing in this still emerging information technology, is that it must be continual: both for the fruition of past investment as for the seeding of timely future delectation.

The logic programming paradigm provides a well-defined, general, integrative, encompassing, and rigorous framework for systematically studying computation, be it syntax, semantics, procedures, or attending implementations, environments, tools, and standards. Logic programming approaches problems, and provides solutions, at a sufficient level of abstraction so that they generalize from problem domain to problem domain. This is afforded by the nature of its very foundation in logic, both in substance and method, and constitutes one of its major assets.

Indeed, computational reasoning abilities such as assuming by default, abducing, revising beliefs, removing contradictions, updating, belief revision, learning, constraint handling, etc., by dint of their generality and abstract characterization, once developed can readily be adopted by, and integrated into, distinct topical application areas.

Logic programming is, without a doubt, the privileged melting pot for the articulate integration of functionalities and techniques, pertaining to the design and mechanization of complex systems, addressing ever more demanding and sophisticated computational abilities. Forthcoming rational agents, to be realistic, will require an admixture of any number of them to carry out their tasks. No other computational paradigm affords us with the wherewithal for their coherent conceptual integration. And, all the while, the very vehicle that enables testing its specification, when not outright its very implementation. Or is there?

Moreover, the issue is not just that of conceptual integration but that of conceptual cross-fertilization too. How can one employ learning in the service of belief revision? How does one remove contradictions from updates? How can learnt preferences be combined with preference updates to guide belief revision in a rational agent? How may fuzzy preferences be revised in the light of ongoing abduced assumptions and constraints? How can probabilistic and multi-valued logic programs be combined? Etc.

10 Coda

There is an ongoing meta-argumentation about what is good reasoning, what are the conclusions we can draw from a discussion (i.e. a semantics), which is inherent to all scientific activity. I foretell the computer will be used more and more as a research aide, not just to automate but also propose experiences and hypotheses and, in the end, by making our own conceptions on epistemology application repeatable and externalized, the computer will make them more objective too.

The language of logic is universally used both by the natural sciences and the humanities, and more generally is at the core of any source of human derived common knowledge, so that it provides us with a common ground on which to reason about our theories. Since the field of cognitive science is essentially a joint effort on the part of several distinct knowledge fields, we believe such language and vocabulary unification efforts are not just useful, but mandatory.

Epistemology will eventually have the ability to be shared, be it with robots, aliens or any other entity who must needs perform cognition to go on existing and program its future. Creating situated computers and robots means carrying out our own cognitive evolution by new means; with the virtue of engendering symbiotic, co-evolving, and self-accelerating loops.

Computerized robots will reify our scientific theories, making them objective, repeatable, and part of a commonly constructed extended reality, built upon a multi-disciplinary unified science. Artificial Intelligence and the Cognitive Sciences, by building such entities, provide a huge and stimulating step towards furthering that construction. To this end, the functionalist stance is most helpful.

Participation by philosophers and psychologists in the venture will take the form of partaking in the development of the conceptual underpinnings of the required computer implemented logic tools, and of employing such tools for attaining a computerized functional rational reconstruction of historical epistemological argument, of scientific theory evolution, and of evolutionary theories of human cognitive development (Wheeler & Pereira 2008).

REFERENCES

J. J. Alferes, et al. (2002). 'Evolving Logic Programs'. In S. F. et al. (ed.), *Procs. 8th European Conf. on Logics in Artificial Intelligence (JELIA'02)*, LNCS 2424, pp. 50–61. Springer.

J. J. Alferes, et al. (2000). 'Dynamic Updates of Non-Monotonic Knowledge Bases'. *J. Logic Programming* **45**(1-3):43–70.

D. M. Buss (ed.) (2005). *The Handbook of Evolutionary Psychology.* John Wiley & Sons Inc.

J. L. Casti & W. DePauli (2000). *Gödel - A life of Logic.* Basic Books.

W. S. Cooper (2001). *The Evolution of Reason: Logic as a Branch of Biology.* Cambridge Studies in Philosophy & Biology. Cambridge University Press.

M. Davis (2000). *The Universal Computer: The Road from Leibniz to Turing.* W.W. Norton & Co.

P. R. Gross & N. Levitt (1994). *Higher Superstition.* The Johns Hopkins University Press.

R. Hanna (2006). *Rationality and Logic.* MIT Press.

J. Holland (1998). *Emergence — From Chaos to Order.* Addison-Wesley.

G. Lopes & L. M. Pereira (2007). 'Prospective Logic Agents'. In J. M. Neves, M. F. Santos, & J. M. Machado (eds.), *Procs. 13th Portuguese Intl.Conf. on Artificial Intelligence (EPIA'07).* Springer LNAI.

J. R. Lucas (1996). 'Minds, Machines, and Gödel: A Retrospect'. In P. Millican & A. Clark (eds.), *Machines and Thought*, vol. 1, pp. 103–124. Oxford University Press.

S. Mithen (1996). *The Prehistory of Mind.* Thames and Hudson Ltd.

T. Nagel (1997). *The Last Word*. Oxford University Press.

R. Penrose (1994). *Shadows of the Mind: a search for the missing science of consciousness*. Oxford University Press.

L. M. Pereira (2002). 'Philosophical Incidence of Logical Programing'. In D. G. et al (ed.), *Handbook of the Logic of Argument and Inference*, vol. 1 of *Studies in Logic and Practical Reasoning*, pp. 425–448. Elsevier.

L. M. Pereira (2007). 'Gödel and Computability'. In *Procs. 13th Portuguese Intl.Conf. on Artificial Intelligence (EPIA'07)*, vol. 4874 of *LNAI*, pp. 63–72. Springer.

L. M. Pereira & A. Saptawijaya (2007). 'Modelling morality with prospective logic'. In J. M. Neves, M. F. Santos, & J. M. Machado (eds.), *Procs. 13th Portuguese Intl.Conf. on Artificial Intelligence (EPIA'07)*. Springer LNAI.

S. Shennan (2002). *Genes, Memes and Human History — Darwinian Archaeology and Cultural Evolution*. Thames & Hudson Ltd.

H. Wang (1973). *From Mathematics to Philosophy*. Routledge.

H. Wang (1987). *Reflections on Kurt Gödel*. The MIT Press.

G. Wheeler & L. M. Pereira (2008). 'Methodological Naturalism and Epistemic Internalism'. *Synthčse* **163**(3):315–328.