Computer Science Journal of Moldova, vol.**15**, no.2(44), 2007

Affordability and Paradigms in Agent-Based Systems

Boldur E. Bărbat, Andrei Moiceanu,
Sorin Pleşca, Sorin C. Negulescu

At each level of complexity entirely new properties appear.

Psychology is not applied biology, nor is biology applied chemistry.

PHILIP W. ANDERSON

Abstract

The paper aims at substantiating that in universities with scarce resources, applied Information Technologies (IT) research is affordable, even in most advanced and dynamic sub-domains. This target is split into four specific objectives: a) to set up a framework for IT research affordability in universities representative for current East-European circumstances; b) to outline a workable approach based on synergistic leverage and to assess the paradigms prevalent in modern artificial intelligence through this "affordability filter"; c) to describe the evolution and the current stages of two undertakings exploiting paradigms founded on *emergence* (the sub-domains are *stigmergic coordination* and *agent self-awareness*); d) to summarise for both sub domains the mechanisms and the architectonics (the focus is on computer science aspects; implementation details will be given in future papers). The results in both directions appear as promising and reveal significant potential for transdisciplinarity. From this perspective, the paper is a call to improved cooperation.

Key-words. Open, heterogeneous, dynamic and uncertain environments (OHDUE); Emergence; Synergy; Stigmergy; Agent self-awareness; Agent-oriented software engineering (AOSE).

©2007 by B.E. Bărbat, A. Moiceanu, S. Pleşca, S.C. Negulescu

1 Introduction

As regards advanced IT, it is generally accepted that academic research in East-European countries is still limited not by scientific potential but rather by financial or logistic boundaries. Considering the case of the "Lucian Blaga" University of Sibiu, the paper tries to show that – without aiming at immediate spectacular results – such research can be afforded even in demanding fields like agent-based systems (ABS). The key seems to be updated context awareness and a short delay in exploiting it. This context is delineated by:

Environment. Present-day IT environments, except for some irrelevant applications, are *open* and *heterogeneous* (the resources involved are unalike and their availability is not warranted), *dynamic* (the pace of exogenous and endogenous changes is high) and *uncertain* (both information and its processing rules are revisable, fuzzy, uncertain and intrinsically non-deterministic – as every stimuli generator) [47]. All four features are there in any modern IT system, albeit with different weights. (Over)simplifying, it could be considered that the first two features are universal (i.e., any significant activity environment in the post-industrial age enjoys them), whereas the last two are their – more or less avoidable – consequences. However, in IT this general role – still essential – is surmounted by an even weightier one: in the Internet era of "computing as interaction" [1], dynamic and uncertain environments are inherent, since deterministic applications are practically vanishing. In short, nowadays, relevant applications are grounded in open, heterogeneous, dynamic, and uncertain environments (OHDUE).

Affordability. For the current Romanian (technological, social, and economic) context, this is the key feature focused on because applications should be not just inexpensive, but also convenient as tools for ordinary interactants in the emerging "information (or broad-band) society".

Anthropocentrism. Here it means focusing on the human being as user, beneficiary, and, ultimately, raison d'être of any application. Anthropocentrism becomes common in IT because it is crucial for user acceptance (the main macro-architectural features looked for are *flex-*

ibility and *user-friendliness*). Only powerful IT companies can impose technocentric solutions. Hence, affordability requires rather "user-pulled" than "technology-pushed" research. (Anthropocentric interface design is commented in [10].)

Technological Trends. Prigogine's idea that the most interesting scientific activities seem to occur at domain interfaces [28] is a confirmed – and, because of financial reasons, often (first of all, now, in Romanian universities) the *only affordable* – path for applied research [21] [20]. Again, only powerful IT companies can afford to exploit effectively mature technologies. Academic research is confined to find "Prigogine niches"; they can be found mainly through innovative, emerging technologies.

Considering the context described above, the target is split into four specific objectives: a) to set up a framework for IT research affordability in universities representative for current East-European circumstances (the University of Sibiu is taken as instance); b) to outline a workable approach based on synergistic leverage and to assess the paradigms prevalent in modern artificial intelligence (AI) through the "affordability filter"; c) to describe the evolution and the current stages of two undertakings in challenging and innovative areas, exploiting (related but distinct) paradigms founded on *emergence* (after filtering, the sub-domains selected are *stigmergic coordination* and *agent self-awareness*); d) to summarise for both sub-domains the mechanisms applied and the architectonics of the current experimental models (considering the journal profile, the focus is on computer science aspects regarding the sub-domains; implementation details as well as aspects related to complexity science, will be given in future papers).

Accordingly, the rest of the paper is organised as follows: the *affordable synergy*-based *approach* is outlined in Section 2 and used for the *paradigm filtering* (*and blending*) in Section 3. The next two sections describe abridged the emergence-based core undertakings stemming from: a) *simulated emergence* (in Section 4 stigmergic coordination (SC) simulates self-organization in ant colonies); b) *emulated emergence* (in Section 5 Gödelian self-reference tries to emulate self-awareness of interface agents). The paper concludes (Section 6) that results in both directions appear as promising and reveal significant potential for transdisciplinarity as well as for teamwork. From this perspective, the paper is a call to improved cooperation in developing innovative and still affordable ABS. (Why *Agent-Based* Systems? Because of two nuances – both related to affordability: a) Although many such systems are *Multi-Agent Systems* (MAS) – or, at least labelled as such – it is not necessary to exclude applications with a few agents. b) Agent-*based*, not agent-*oriented*, to cover also very simple agents in unpretentious applications.)

2 An Affordable Approach: Synergy as Leverage

To draw a workable approach, the previous assertions about the context are restated as premises and commented upon, focusing on their consequences as regards the very approach:

- *Affordability Is Sine Qua Non.* Moreover, it is vital for both research aspects: a) *Effort.* A university research undertaking is considered affordable if it proves workable as a project ending with a few experimental models validated "in vitro" (for instance, within the narrow scope of a PhD thesis [43]). b) *Applicability.* Applications embodying the research results should be intrinsic *usable* (to represent solutions first to toy problems but soon to real-world problems – albeit small-scale one – so as to be at least roughly conclusive) and easy *extendable* for further research. Thus, the undertaking must ensure affordability for future end users with quite scarce resources, i.e. the applications must be not just inexpensive, but also convenient as tools for ordinary interactants in the emerging "information (or broad-band) society" (for instance, a research project regarding inductive [33] [53] [54] non-algorithmic [51] e-Learning [47], should be tested also by individual students on usual configurations).

- *Anthropocentric Design Is Crucial For User Acceptance.* The goal of anthropocentrism entails a kind of "anisotropic" flexibility: at the outer surface (that means at the interface level), the application has to match only user expectations, i.e. to *adapt itself* to the user (the users are not forced to adapt themselves). Nevertheless, as this requirement does not hold also for the internal application layers, in order to improve efficiency – above all, under time criticality – the application and the controlled process itself can (and should) *be adapted* and *adapt* mutually. However, how can we reach such a comprehensive flexibility, expressed through both its exogenous facet (*adaptability*), and its endogenous one (*adaptivity*)? While *adaptability* (adapting system behaviour to user *explicit* requests) can be attained rather easily through generic, parameterised functions, *adaptivity* (learning to shape, refine, and update several actions so as to meet *implicit* user expectations) involves, first of all, *intentionality* and *ability to learn* [26] [6] [7]. That means genuine distributed and multi-faceted AI.

- *Technological Trends in Artificial Intelligence Means Agent Technologies.* The trend is well-established, steady, obvious, and uncontested. However, general consensus stops here. The palette of divergent opinions is impressive and involves almost all facets of agent-orientation (AO). Some of the most controversial issues are [9]: complexity (many simple agents or a few complex ones?), degree of parallelism (coarse-grain or fine-grain, genuine or simulated?), degree of autonomy (strict human control but low agent effectiveness or high autonomy but uncertain – even potentially dangerous – agent behaviour?), design mechanisms (are object-oriented tools suitable for implementing agent architectures or are needed new tools?), etc. The very nature of agent logic is disagreed upon [17]: should be reasoning based on symbolic or on sub-symbolic inferences? Thus, comparing and selecting paradigms is a hot issue in any context.

Such very restricting premises require an approach with a high "performance/resource ratio". In turn, that entails a kind of "resource amplifier". The ancient, best known, verified, versatile and affordable universal amplifier is *synergy* [29]. Thus, *in nuce*, the approach could be defined: "in search of synergy, from Aristotle to Haken", i.e., "synergy whenever and wherever it emerges" (reminding the slogan about information, the idea sounds familiar). Where does synergy come from? Considered pragmatically, in its comprising Aristotelian (or Daoist) meaning, synergy has three sources (adapted and extended after [21] [9]):

A) *Homogeneous Amassing of Many Simple Entities.* For instance, although ants behave rather as robots than as agents, the system they belong to is not a "multi-*robot*" system, but a "multi-*agent*" one. This wonder is due to the synergistic effect of their interaction: beyond the individuals (ants or ant-like entities), the team (colony, society, system) comes out. That is the preterminological meaning of "synergy" due to Aristotle: the whole is stronger than the sum of its parts. This kind of synergy (referred to as "*classical synergy*") is intrinsically linked to *multiplicity* and *parallelism.* Obviously, if it would be only "one part", the "whole" could not be stronger than itself. However, multiplicity implies *parallelism*, not just because the ,,parts" creating the swarm *coexist* but because their incessant *interaction* (imposed by the real-world dynamics).

   Since synergetics is a well established science, Haken's principles [32] [37] should be considered too (in parentheses are comments on their role in this paper): a) "Subsystems slaved by the system" (rather irrelevant when the system – for instance, an ant colony – remains unexplained). b) "Cooperating subsystems" (does not apply directly since ants do not communicate). c) The "threshold" principle (of unquestionable importance in both nature and IT: a few ants are surely unable to run a colony and, as well, a few artificial ones are unable to solve research problems, no matter how long they try). d) The "self-organization" principle (the conceptual crux of emergence).

B) *Heterogeneous Interacting of Few Complex Entities.* Dissimilar entities create "added value" rather by *complementarity* (e.g., in symbiosis) than by *crowds*. Physical entities can be substituted by areas of expertise, sub-fields, or paradigms. However, despite the simple idea of combining paradigms to reach synergy, in IT as a whole and in AI in particular, paradigms are still poorly connected. In principle, any paradigms can be mixed up but, since engineering requires *symbols*, integrating the symbolic paradigm with subsymbolic ones (mainly based on biological models) is the target at hand. This kind of synergy will be called here "*inter-paradigmatic*" (see next section).

C) *Trans-Disciplinarity.* The third source matches Prigogine's idea mentioned in Section 1, about domain interfaces. It is a confirmed path for research (because of affordability, it becomes the only one). The prefix *trans* (instead of the usual *inter* or *multi*) highlights the trend towards osmotic-like confluences.

Since not all connotations of those three kinds of synergy are suggested by the term "*emergent synthesis*" (more frequently used in modern IT contexts), here the term "*synergy*" is preferred.

More common approach facets (e.g., micro-continuity, rapid and/or successive prototyping, generic architecture) have been described in [6] [7] [8] [11] [20] [13] [16] and do not need further comments.

3 Filtering and Blending Paradigms

(Instead of) Methodology. First *synergy* itself was deeper investigated, in its relationship to both *complexity* and symbols [9] [19] [21]. The results highlighted the criteria for selecting and the directions for combining paradigms able to boost SC (details in Section 4). Here they are updated, extended to all emergence-based paradigms as well as to transdisciplinarity [29] [9] [17] [13] [16] [15], and adapted to the paper target:

- a) As regards *complexity*, older results were reinforced: Synergy implies *multiplicity*; multiplicity entails *parallel interaction*; this implies two kinds of complexity: *structural* (regarding the system) and *cognitive* (regarding the way the system is perceived by its users). From another outlook, complexity may lie at two levels: *entity* or *system*. In this respect classical synergy is exclusive: entities should be simple, and complexity should emerge at system level, through the huge number of interacting components. Thus, the threshold principle becomes crucial also from an engineering standpoint: "where from start a huge number?" Indeed, albeit almost any biologically inspired model (even when considerably modified) has proved to be helpful to AI, the number of entities should stay affordable. However, artificial neural networks are founded on massive, fine-grain parallelism (as a premise for connectionism); likewise, evolutionary algorithms yield relevant results only with numerous populations; but massive parallelism is hardly affordable with scarce resources, even when intense simulation is involved. (Fortunately, there are fewer ants in a colony than neurons in a neural network or chromosomes in an evolutionary algorithm.)

- b) As regards *symbols*, it seems that not even nature can afford to deal with very many complex entities: the strength of synergy seems to be proportional not only to the scale of parallelism itself (number of entities involved) but also to the extent of sub-symbolic depiction; in other words, self-organization emerges easier in sub-symbolic contexts. According to Heylighen [34], there are at least seven characteristics of self-organising systems; here are relevant four of them: global order from local interactions; distributed control; robustness and resilience; non-linearity and feedback. Therefore, sub-symbolic paradigms, with their vast intrinsic synergistic potential, are a good choice for exploiting the first synergy source. Combining this result with those mentioned above, the best choice is the biologically inspired paradigm involving the smallest number of entities, i.e., *stigmergic coordination*. Though, the symbolic approach (Newell-Simon hypothesis) is unavoidable for at least four reasons: b1). Primarily when complexity (of all kinds) is high, software effectiveness becomes uncertain and direct human intervention is rather welcomed. Since modern IT systems are anthropocentric, the users (no matter whether system engineer, application developer, manager, and so on) should be allowed to monitor the system and to communicate with, using familiar semantics – i.e., symbol-based languages, humans are accustomed to. b2). The same reason holds for any human-agent interaction (vital for interface agents). b3). To prevent unintended – and not reasonably predictable – conduct in OHDUE, agent actions having potential ethical implications should be more rigorously controlled. Consequently, without obstructing agent autonomy, their "ethical architectonics" should be based on symbolic processing [14] [44]. b4). From a software engineering perspective, symbolic processing is not a (best) *choice*, but a *must* because any engineering undertaking involves symbols: design means to *project*, and any relation to the future implies symbols.

- c) To get synergy also from the second source, components based on different paradigms could be mixed in the same model, provided that both extreme irreconcilable dogmata (the Newell-Simon hypothesis versus the physical-grounding (ethological) paradigm applied in Brooks' automata challenging the need for symbol-based explicit representation) are deprived of their *necessity* conditions, i.e., are considered only as *sufficiency* conditions. Moreover, as corollary of the assertions above, some paradigm blend involving the symbolic paradigm is not just possible but highly desirable: humans – seen as the apex of symbolic reasoning – act as counterparts of sub-symbolic entities, in getting inter-paradigmatic synergy. Hence, any blend should include both symbols AND sub-symbols (indeed, here the "AND" should be read almost as the similar Boolean operator). After all, the way humans make inferences proves that nature created in our brains the amazing blend of (a kind of) "von Neumann"-like algorithmic procedures (in the left hemisphere) with non-algorithmic (creative, heuristic, emerging) procedures (in the right hemisphere). As a rule, in agent strategic decision making, the layer of mental (symbolic) context should prevail over the layer of situational (sub-symbolic) context. (For the time being, it is safer to set up ABS founded also on *reason* not just on *instincts*! Nevertheless, in the long run, it will depend on the future "Zeitgeist-stance".)

- d) The third synergy source, being grounded on transdisciplinarity, seemed to be, until recently, outside the scope of software engineering but the promising *memetic* approach [25] could help, since most paradigms in modern AI have an obvious memetic character ([15] [16] and future papers).

4 Simulated Emergence

The problem is shaped by affordability restrictions and by the question (unanswered in Section 3) "where from start a huge number?" Unfortunately, all biologically inspired paradigms model massively parallel societies/systems. Thus, all are affordable only through simulation. (Even the less demanding SC involves usually at least tens of entities.)

On the other hand, albeit AO is already a well-established course in AI – and even in IT as a whole – at the engineering stage its effectiveness is rather unsuitable for affordable ABS, no matter what paradigms are applied, because – despite the increasing number of biologically inspired models – the newer paradigms they are founded on are in a yet syncretic stage, embodying a promising (but too little exploited) niche in itself – for both applied research and effective implementation. Thus, paradoxically, despite the increasing shift from predominant algorithmic reasoning towards subsymbolic reasoning paradigms (seen rather as complement, than as alternative) and although the fidelity towards the biological model is sometimes quite low, letting place for components sticking to older paradigms – mainly the symbolic one – inter-paradigmatic synergy is rather not manifest enough.

To exploit the niche, as graft paradigm there was chosen the symbolic one (for obvious conceptual and engineering reasons; moreover, albeit not always asserted explicitly, symbolic processing is already present in some successful algorithms based on SC – e.g., Dorigo's *Elitist Ant System*). On this groundwork there was set the engineering construction [45] [21] [20] [46] [23] [48] [31]: a) designing specific mechanisms to graft symbolic components onto the sub-symbolic foundation (the filtered biological model); b) tailoring the mechanisms to sub-fields (primarily, manufacturing control); c) building an experimental model as a test-bench for this sub-field but, considering also future extensions. Even if in SC emergence is impressive, the trouble to understand what is in fact going on at system level, is less upsetting than in the case of more familiar sub-symbolic paradigms (as artificial neural networks or evolutionary algorithms) since ant behaviour is easier to follow due to its simplicity: the ant travels from the ant hill to the food source and back guided *only* by pheromones.

Specifically, some (less quantifiable) synergy was achieved deviating from the biological model applied in the Elitist Ant Systems by adding symbolic processing components (firstly adapting the environment and secondly instituting limited central coordination). Focusing on affordability and keeping a definite engineering perspective, the immediate purpose of [20] was to save computer resources in applying stigmergic control to industrial problems by exploring the relationship between the number of digital ants and problem complexity. The long-range target is to follow the analogy to superconductivity: moving the threshold in order to improve performance and/or save computing resources. Since the ideal threshold is expressed by Heaviside step-functions, as asymptotes to sigmoid functions, every discrete function expressing a solution instance will be represented by a fitting sigmoid:

$$Sq = Obj/(1 + exp(Td - s)) \quad (1)$$

where: s("swarm") is the number of ants; Td ("threshold") is the value of $s$where self-organization [5] becomes manifest (the number of ants closest to the inflexion point of the sigmoid); Obj ("objective") is the optimal solution (the minimal number of iterations for reaching the optimal route given by the benchmark program); Sq ("solution quality") is the instance result, reflecting the degree of self-organization expressed through the ratio between the number of iterations for the best simulation run and the numbers of iterations for the current run (both for reaching 95% from the optimal solution). The term "swarm" was chosen for its ambiguity: it has the connotation of both "crowd" (discrete variable) and "multitude" (indefinite). Thus, it allows simplifying the language without affecting mathematical rigor (requiring a continuous domain for a sigmoid); of course, during simulations, only integer values where assigned to s, considered as cardinal of the swarm set. Td is obtained from (1):

$$Td = ln(Obj/Sq - 1) + s$$

The corresponding sigmoid for $Td = 4$ is represented in Figure 1.

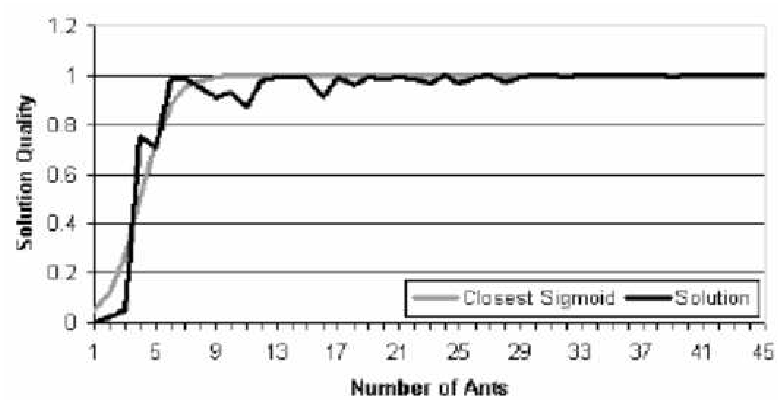

Figure 1. A very low threshold of self-organization in stigmergic coordination (taken from [20]).

As regards early validation attempts in industrial environment, synergy achieved through SC (called in [20] "Stigsynergy") will be tested as component integrated in a more comprehensive software toolkit for virtual enterprises [27]. Tools for catalysing emergence in such projects will be described in more implementation-oriented papers.

5 Emulated Emergence

Albeit both governed by emergence, the two kinds of undertaking are quite different: whereas when *simulating* common ant behaviour, emergence is expectable also from artificial ants and – despite obvious nonlinearity – their performances can be improved "incrementally" (as shown in Section 4), when *emulating* emergence the basic argument about "holism versus reductionism" transcends theory and becomes crucial for applications. Indeed, here are hidden not only the internal processes of self-organization (unanswered "Why?"), but also any "stigmas" (unanswered "How?").

Research regarding self-awareness in ABS is founded on Hofstadter's ideas [35] and was presented in: [13] (to illustrate the broadband technology potential from an anthropocentric and transdisciplinary perspective); [15] (in a larger interdisciplinary framework); [16] (focusing on computer science aspects, and keeping a definite engineering perspective); [47] (offering an alternative approach to e-Learning, where "*Learning*" is action-oriented and highly personalised, while "*e–*" is carried out through a software entity acting as self-referencing coach and interacting with the user as interface agent). The outline below follows [16]. Since the target is a generic architecture – based on Gödelian self-reference (GSR) – for applications meant for present-day environments (i.e., OHDUE), the key question is one of viability: is it suitable to consider self-awareness as relevant agent feature when many other strong agency characteristics are missing, even in current large-scale ABS? Yes, because: system complexity makes it desirable [24] [2] [4] [22] [40] [41] [42], agent technology makes it possible [1] [30] [36], and approaching it by GSR (the agent clones itself – usually spawning better architecture) makes it affordable. Albeit the "Self-*" memeplex invaded modern IT and despite memetic likeness, the "duplicate me" [38] instructions in the genotype and the "I clone myself" message in Figure 2 are fundamentally different. Agent self-cloning means spawning an agent identical to its parent. An example at implementation level (for systems with Windows-like application programming interface): the parent-agent main thread calls a "*CreateThread*" system function [12] passing itself as parameter. It is expectable that such self-reference could be a matrix for "strange loops", which in turn could lead to a stepwise emergence of (a primitive kind of) self-awareness, hoping that "Isomorphisms Induce Meaning" [35]. If this expectation should be too great, at least GSR should provide a workable mechanism for improving agent architecture (as "Plan B" for real-world applications, as used in [47]). Self-cloning is conservative: the agent clones itself, preserving self-representation (its "I"), but not necessarily its old world model too. Instead, the model is regenerated through "phenotypical expansion", in short through elementary learning: as the agent learns, it fills out the ontology (the black part of its rucksack), and, when assessing a significant improvement, it transfers the latest assimilated knowledge into the executable program representing statically the agent (i.e. into its "genotype") by cloning itself. Here lies the weakest link of the generic architecture, since filling out ontologies for real world problems is resource demanding (because it takes much time, it is the only key application component still in the stage of suiting only toy problems). This drawback is somewhat balanced by attractive features, as assigning semantic value to the iconic space (e.g., in the context of computer-aided semiosis in trans-cultural interfaces [18]).

From the pragmatic perspective of an application, the process is seen rather as spawning "smarter progeny" (as shown for e-Learning in [47]). Evolution is assessed through a performance metrics suited to action-oriented "Simon-type machine learning" [52] [3] (i.e., the diminishing duration of task completion).

Hence, in emulating emergence there are two key problems: a) *Feasibility*. Is the hypothesis valid? For instance, in nature (humans included) is self-awareness emerging from strange loops? b) *Effectiveness*. How can be the expected feature modelled? For instance, can agent self-awareness stem from GSR? Corollary: a "Plan B" is mandatory to save the undertaking as applied research when the basic target is too far. As a result, the e-Learning application [47] is designed to be useful

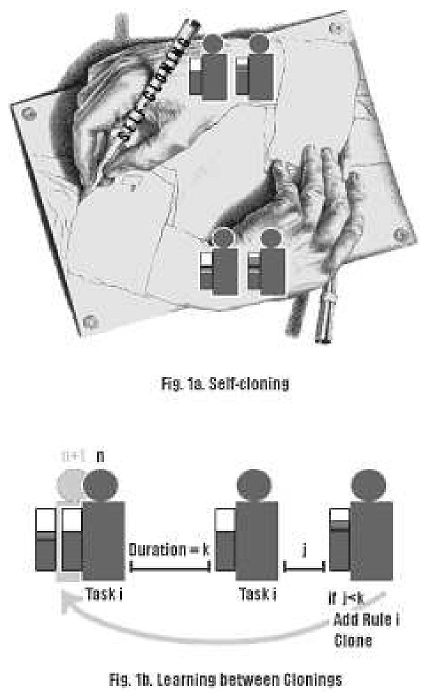

Figure 2. The Self-Cloning Loop: Doing, Learning, Knowing, Cloning (taken from [16]).

even if agent self-awareness is not yet achieved: learning is considered – in both humans and agents – as a process where most effectiveness is reached through a blend of symbolic ("left-hemisphere"-like) and subsymbolic ("right-hemisphere"-like) modi operandi. Hence, neither "apprenticeship learning", nor "by rote learning". However, the two extremes, albeit equally dangerous, are not similarly hard to fight: at least in Romania, nowadays, the average approach to learning is much closer to "by rote". Thus, the balance is redressed, favouring right hemisphere tactic, i.e., "non-algorithmic" course of action. Fortunately, that does *not* imply necessarily "sub-symbolic", because: a) Symbolic processing is unavoidable in any learning process (fact rejected only by radical cognitive theories denying knowledge decomposition and decontextualization [3]). b) Anthropocentric interfaces require symbolic human-computer communication. c) As asserted in Section 3, massive parallelism is hardly affordable.

6 Conclusions and Future Work

The results in both directions appear as promising and reveal significant potential for transdisciplinarity as well as for teamwork. Since it is an ongoing research, the conclusions are grouped in three time ranges: A) *short* (relevant results obtained with current experimental models), B) *medium* (directions for future work likely to be successful), and C) *long* (engineering and scientific openings):

- A1. As regards *simulated emergence*, the experimental model attested that the threshold *exists* and depends on problem *type* and *complexity*; the same solution quality can be obtained with fewer ants than used in common benchmarks. Moreover, combining *stigmergic control* with *symbolic processing* components has significant synergistic potential (the most useful mechanism proved to be "*User-Driven Heuristics*").

- A2. As regards *emulated emergence*, the current agent endorses the model, and, mainly, the usefulness of self-cloning. Its generic architecture proved to be affordable for toy problems (inductive non-algorithmic e-Learning) and the model was not difficult to implement (the only exception: filling out dynamic ontologies, even for primitive toy problems is hard work and risky outside an authentic transdisciplinary effort). In short, "Plan B" is viable.

B1. For *stigsynergy*, despite insufficient statistical relevance, it seems that: a) The sigmoid pattern seems credible because it is similar to – or, at least, consistent with – trustworthy results regarding exponential convergence [49]. This similarity is significant because both the target (*performance* vs. *affordability*) and the approach (systems achieving *global ends based on local behaviour* vs. the *principles of synergetics*) were different. b) From an engineering perspective, in operational research, comparable solution quality could be obtained with a significantly less number of ants than used in common benchmarks, saving thus at least one order of magnitude of processing time. c) The results remind the von Neumann theorem for the complexity threshold (e.g., a map with only 14 towns proved to be too simple to allow self-organization).

B2. For emulating *agent self-awareness*, the main hindrance imposed by affordability restrictions is the purely software, bodiless, agent nature: the agent will lack the awareness of its own body, crucial for the somatoception-based self-representation achievable by robots. Hence, the expected emergence of a primitive "I" should be catalysed through a powerful temporal dimension and an emphasised non-algorithmic behaviour: the main feature to be added to usual interface agent architecture and preserved through self-cloning is its primal sense of time (besides its intrinsic architectonic value, it could be helpful in future "pseudosomatoception" as surrogate for the lacking sense of space and haptic proprioception).

C1. a) From an engineering perspective it is worthwhile to try to exploit the analogy between self-organization in *stigmergic coordination* and in *electromagnetism*, based on the correspondences:

problem solving vs. superconductivity; number of digital ants vs. absolute temperature; solution quality vs. conductivity. The target would be to lower the threshold of self-organization for operational research problems modifying the macro-parameter M (M stays for *Model* in SC and for *Material* in electromagnetism). b) From a scientific perspective it seems attractive to find out whether in the real world, natural and digital beings live in, there are relevant problem classes (or, more general, processes) where "many starts from four" (as the results for 45 towns in Figure 1 may suggest).

C2. From a computer science perspective it is much too soon to claim that agents could achieve self-awareness through Gödelian self-reference *per se*. (Nevertheless, first indices mentioned in A2 and B2 are rather encouraging.) improve agent architecture, first of all its dynamic ontology, sense of time and reactivity (it should be much more event-driven).

C3. Complexity as macro-parameter must be better investigated: a) For *stigmergic coordination* besides the number of ants and the number of towns, probably other factors (e.g., map topology, anisotropic graph edges) can play a major role. b) For self-referencing agents a much improved architecture – first of all their dynamic and visual ontology, sense of time and reactivity (they should be much more event-driven) – could help emulating the autocatalytic process of self-awareness.

C4. Since synergy is essential fundamental in both proposed approaches, and transdisciplinarity is a primary source, a second-degree synergy (a kind of "synergy of synergies" expressed as second derivative) may appear, creating relationships closer to the ideas of Aristotle and Lao Zi, than to those of Haken.

References

[1] AgentLink III. *Agent based computing. AgentLink Roadmap: Overview and Consultation Report.* University of Southampton. http://www.agentlink.org/roadmap/al3rm.pdf, 2005.

[2] Amir, E., M.L. Anderson, V.K. Chaudhri. *Report on DARPA Workshop on Self-Aware Computer Systems.* Artificial Intelligence Center SRI International, 2004.

[3] Anderson, J.R., L.M. Reder, H.A.Simon. *Applications and Misapplications of Cognitive Psychology to Mathematics Education.* Texas Educational Review, 2000.

[4] Anderson, M.L., D.R Perlis. The roots of self-awareness. *Phenomenology and the Cognitive Sciences*, 4, pp. 297–333, Springer, 2005.

[5] Banzhaf, W. Self-organizing Systems. *Encyclopedia of Physical Science and Technology*, Academic Press, New York, **14**, pp. 589–598, 2002.

[6] Bărbat, B.E. Holons, Agents, and Threads in Anthropocentric Systems. *Studies in Informatics and Control Journal*, **9**, 3, pp. 253–268, 2000.

[7] Bărbat, B.E. Agent-Oriented Captology for Anthropocentric Systems. *Large Scale Systems: Theory and Applications 2001* (F.G. Filip, I. Dumitrache, S.S. Iliescu, Eds.), Elsevier, IFAC Publications, pp. 214–219, 2001.

[8] Bărbat, B.E. Emotions and Time in Captological Agents. *Third International -NAISO Symposium on ENGINEERING OF INTELLIGENT SYSTEMS*, ICSC-NAISO Academic Press Canada/The Netherlands, 99 (Abstract; full paper on CD-ROM), 2002.

[9] Bărbat, B.E. *Agent-Oriented Intelligent Systems.* Romanian Academy Publishing House, Bucharest, 2002 (in Romanian, "Grigore Moisil" Prize of the Romanian Academy).

[10] Bărbat, B.E. *Communicating in the world of humans and ICTs.* Chapter 8 in L. Fortunati (Ed.) *COST Action 269. e-Citizens in the Arena of Social and Political Communication*, pp. 113–142, EUR21803, Office for Official Publications of the European Communities, Luxembourg, 2005.

[11] Bărbat, B.E., R. Creulescu. Affordable Affective Avatars. Persuasion, Emotions and Language(s). *Proc. of the 1st Balkan Conf. in Informatics (BCI'2003)*, (Y. Manolopoulos, P. Spirakis, Eds.), Publ. Centre Technological Educational Institute of Thessaloniki, pp. 25–35, 2003.

[12] Bărbat, B.E., F.G. Filip. *Industrial Informatics - Real-time Software Engineering.* Technical Publishers, Bucharest, 1997 (in Romanian).

[13] Bărbat, B.E., A. Moiceanu. I, Agent. *The good, the bad and the unexpected: The user and the future of information and communication technologies, Institute of the Information Society,* Moscow (forthcoming, May 2007).

[14] Bărbat, B.E., A. Moiceanu, H.G.B. Anghelescu. *Enabling Humans to Control the Ethical Behaviour of Their Virtual Peers.* Chapter in Enid Mante-Meijer, Leslie Haddon and Eugène Loos (Eds.) *The Social Dynamics of Information and Communication Technology.* (To be published by Ashgate, Aldershot, UK, 2007.)

[15] Bărbat, B.E., A. Moiceanu, I. Pah. Gödel and the "Self"ish Meme. *Gödel – Heritage and Challenge.* Interdisciplinary Symposium, Romanian Academy, Bucharest, 2007.

[16] Bărbat, B.E., A. Moiceanu, I. Pah. Gödelian Self-Reference in Agent-Oriented Software. (Accepted for the 11th WSEAS Int. Conf. on Computers. Agios Nikolaos, Crete, July 2007.)

[17] Bărbat, B.E., S.C. Negulescu. From Algorithms to (Sub-)Symbolic Inferences in Multi-Agent Systems. *International Journal of Computers, Communications & Control*, **1**, 3, pp. 5–12, 2006. (Paper selected from the *Proc. of ICCCC* 2006.)

[18] Bărbat, B.E., S.C. Negulescu, A.E. Lascu, E.M. Popa. Computer-Aided Semiosis. Threads, Trends, Threats. (Accepted for the 11th

WSEAS Int. Conf. on Computers. Agios Nikolaos, Crete, July 2007.)

[19] Bărbat, B.E., S.C. Negulescu, C.B. Zamfirescu. Boosting Synergy in Agent-Based Optimisation Tools. *Scientific Bulletin of "POLITEHNICA" University of Timioara*, Transactions on AUTOMATIC CONTROL and COMPUTER SCIENCE, **49** (63), 4, pp. 119–124, 2004.

[20] Bărbat, B.E., S.C. Negulescu, C.B. Zamfirescu. Human-Driven Stigmergic Control. Moving the Threshold. *Proc. of the 17th IMACS World Congress (Scientific Computation, Applied Mathematics and Simulation),* (N. Simonov, Ed.), e-book, ISBN 2-915913-02-01, *Paris*, 2005.

[21] Bărbat, B.E., C.B. Zamfirescu, S.C. Negulescu. The best from ants and humans: synergy in agent-based systems. *Studies in Informatics and Control Journal*, **13**, 1, pp. 47–59, http://www.ici.ro/ici/revista/sic2004_1/art5.pdf, 2004.

[22] Bermúdez, J. L. *The Paradox of Self-Consciousness.* Cambridge, MA, MIT Press, 1998.

[23] Brueckner, S.A. *Engineering Self-organising Systems: Methodologies and Applications.* Springer Publishing 2005.

[24] DARPA. *Workshop on Self-Aware Computer Systems 2004. Statements of Position.* http://www.ihmc.us/users/ phayes/DWSAS-statements.html#top

[25] Dawkins, R. *The Selfish Gene* (30th Anniversary edition). Oxford University Press, 2006.

[26] Dennett, D. *The Intentional Stance.* The MIT Press, Cambridge, MA, 1987.

[27] Fabian, R., S. Pleca. Stigmergy-Based Software Toolkit For Virtual Enterprises. (Accepted for the 11th WSEAS Int. Conf. on Computers. Agios Nikolaos, Crete, July 2007.)

[28] Filip, F.G. Personal account, 1995.

[29] Filip, F.G., B.E. Bărbat, *Industrial Informatics - New Paradigms and Applications.* Technical Publishers, Bucharest, 1999 (in Romanian).

[30] FIPA TC Agent Management. *FIPA Agent Management Specification.* Standard SC00023K (2004/18/03). http://www.fipa.org/specs/fipa00023/SC00023K.pdf, 2004.

[31] Hadeli, P. Valckenaers, C.B. Zamfirescu, H. Van Brussel, B.S. Germain. Self-organising in Multi-agent Coordination and Control using Stigmergy. *Self-Organising Applications*, Springer Verlag, Lecture Notes in Artificial Intelligence 2977, 2004.

[32] Haken, H. *Information and Self-Organization: A Macroscopic Approach to Complex Systems.* Springer-Verlag, Germany, 1988.

[33] Hawthorne, J. Inductive Logic. In *Stanford Encyclopedia of Philosophy* (E.N. Zalta, Ed.). http://plato.stanford.edu/entries/logic-inductive/#1, 2005.

[34] Heylighen, F. The Science of Self-Organization and Adaptivity. In *The Encyclopedia of Life Support Systems*, 2002.

[35] Hofstadter, D.R. *GÖDEL, ESCHER, BACH: an Eternal Golden Braid.* (Including the Preface to the Twentieth-anniversary Edition.) Basic Books, New York, 1999.

[36] Kephart, J.O., D.M. Chess. The Vision of Autonomic Computing. *IEEE Computer*, 36, 1, pp. 41–50, 2003.

[37] Knyazeva, H., H. Haken. Synergetics of Human Creativity. In *Dynamics, Synergetics, Autonomous Agents. Nonlinear Systems Approaches to Cognitive Psychology and Cognitive Science.* World Scientific, pp. 64–79, Singapore, 1999.

[38] Luksha, P.O. Formal definition of self-reproductive systems. *Proceedings of the eighth international conference on Artificial life*, pp. 414–417, MIT Press Cambridge, MA, 2002.

[39] McCarthy, J. *Ascribing mental qualities to machines.* Technical Report, Stanford University AI Lab., Stanford, CA 94305, 1978.

[40] McCarthy, J. *Concepts of Logical AI.* http://www-formal.stanford.edu/jmc/concepts-ai.html, 2000.

[41] McCarthy, J. *Making Robots Conscious of their Mental States.* http://www-formal.stanford.edu/jmc/ consciousness/node9.html, 2002.

[42] McCarthy, J. *Notes on Self-Awareness.* www-formal.stanford.edu/jmc/selfaware/selfaware.html, 2004.

[43] Moiceanu, A. Self-Awareness in Agent-Based Systems. (Accepted for the 2nd Annual South East European Doctoral Student Conference, South East European Research Centre, Thessaloniki, June 2007.)

[44] Moiceanu, A., B.E. Bărbat. Ethical Behaviour of Self-Aware Agents. *The good, the bad and the unexpected: The user and the future of information and communication technologies*, Institute of the Information Society, Moscow (forthcoming, May 2007).

[45] Negulescu, S.C., B.E. Bărbat. Enhancing the effectiveness of simple multi-agent systems through stigmergic coordination. *Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS (EIS 2004)*, ICSC-NAISO, Academic Press Canada, 149 (summary, complete text on CD), 2004.

[46] Negulescu, S.C., C.B. Zamfirescu, B.E. Bărbat. User-Driven Heuristics for nondeterministic problems. *Studies in Informatics and Control*, **15**, 3, 289–296. (Special issue dedicated to the 2^{nd} *Romanian-Hungarian Joint Symp. on Applied Computational Intelligence*, Timioara, 2005.), 2006.

[47] Pah, I., A. Moiceanu, I. Moisil, B.E. Bărbat. Self-Referencing Agents For Inductive Non-Algorithmic e-Learning. (Accepted for the 11th WSEAS Int. Conf. on Computers. Agios Nikolaos, Crete, July 2007.)

[48] Parunak, H.V.D., S. Brueckner. Ant-like missionaries and cannibals: synthetic pheromones for distributed motion control. *Proceedings of the 4th International Conference on Autonomous agents*, Barcelona, ACM Press, pp. 467–474, 2000.

[49] Parunak, H.V.D., S. Brueckner, J.A. Sauter, R. Matthews. Global Convergence of Local Agent Behavior". *The Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'05)*, http://www.erim.org/∼vparunak/AAMAS05Converge.pdf, 2005.

[50] Ray, P., M. Toleman, D. Lukose. Could Emotions be the Key to Real Artificial Intelligence? *Proc. of the ICSC Symposia on INTELLIGENT SYSTEMS & APPLICATIONS (ISA'2000)* (F. Naghdy *et al.*, Eds.), ICSC Academic Press Canada/Switzerland, 2000.

[51] Resnick, L.B. *Education and Learning to Think*. National Academy Press, Washington, DC, 1987.

[52] Simon, H.A. Why should machines learn? In *Machine Learning. An Artificial Intelligence Approach* (Michalski, R.S. et al., Eds.), Springer-Verlag Berlin, 1984.

[53] Schulte, O. Formal learning theory. In *Stanford Encyclopedia of Philosophy* (E.N. Zalta, Ed.). http://plato.stanford.edu/entries/learning-formal, 2002.

[54] Vickers, J. The Problem of Induction. In *Stanford Encyclopedia of Philosophy* (E.N. Zalta, Ed.). http://plato.stanford.edu/entries/induction-problem/, 2006.

Boldur E. Bărbat, Andrei Moiceanu, Sorin Pleşca, Sorin C. Negulescu, Received June 5, 2007

B.E. Bărbat, S. Pleşca, S.C. Negulescu
"Lucian Blaga" University of Sibiu,
Faculty of Sciences
E–mail: *bbarbat@gmail.com*

Andrei Moiceanu
"Politehnica" University of Timişoara, Faculty of Automation and Computers