The Fifth International Conference on Neural Networks and Artificial Intelligence
May 27-30, 2008, Minsk, Belarus

Towards Modeling Cognitive Evolution

Vladimir G. Red'ko
Scientific Research Institute for System Analysis, Russian Academy of Sciences,
Vavilova Street, 44/2, Moscow, 119333, Russia,
vgredko@gmail.com, http://www.niisi.ru/iont/staff/rvg/index_eng.php

Abstract *- Approaches to investigation of evolutionary roots of human intelligence are discussed. It is argued that the natural way to this study is to analyze cognitive evolution, evolution of animal cognitive abilities by means of mathematical/computer modeling. The area of investigations "Simulation of Adaptive Behavior" that can be considered as a first step towards modeling cognitive evolution is briefly characterized. The project "Animat Brain" (a framework for simulation of adaptive behavior) is described. The sketch program for future modeling of the cognitive evolution is proposed.*

Keywords *- Cognitive evolution, adaptive behavior, problem of intelligence origin, computer modeling*

I. INTRODUCTION

A. What Is a Path to Human Level Intelligence?

Several interesting attempts to analyze ways to understand and to model human level intelligence are initiated now. The special session "Towards Human-Level Intelligence" was organized by W. Duch and N. Kasabov at World Congress on Computational Intelligence in 2006 (at WCCI 2006). The similar session "Cognitive Architectures: Towards Human-Level Intelligence" will be at WCCI 2008 (organized by W. Duch, B. Goertzel, B.-T. Zhang) [1]. Several scientists, who work in fields of evolutionary modeling, neural networks, autonomous adaptive systems, cognitive science and psychology, proposed the Decade of Mind Initiative [2], intending to organize interdisciplinary investigations of mind and natural intelligence. Discussions of research steps directed to investigate the mind of scientists take place in Russian Neural Network Society; these discussions are organized by W.L. Dunin-Barkowski. Interesting proposals for modeling neurocognitive intelligent processes are made in the collective monograph "Challenges for Computational Intelligence" [3]. The current work proposes an approach towards investigations of evolutionary roots of natural intelligence by means of modeling cognitive evolution.

B. Actuality of Modeling Cognitive Evolution

Studies of cognitive evolution are related with a very profound epistemological problem: why is *human* mind applicable to cognition of *nature*? To emphasize the problem, let us consider physics. The power of physics is due to effective use of mathematics. However, why are mathematical deductions applicable to studies of real physical phenomena? Indeed, a mathematician makes logical inferences, proves theorems, basing on his mind, independently from physical world. Why are his results applicable to real nature? More generally, the problem can be stated as follows: why is human mind applicable to cognition of nature?

In order to investigate the stated problem seriously, it is reasonable to analyze it by means of mathematical and computer models. Modeling cognitive evolution, we can analyze, why and how did animal and human cognitive features emerge, how did applicability of human mind to cognition of nature origin.

Fortunately, there is a direction of research "Simulation of Adaptive Behavior", some models of which can be considered as a first step of modeling cognitive evolution. This direction is outlined in Section II. The project "Animat Brain" that proposes a framework for modeling of different forms of adaptive behavior is described in Section III. A particular model of simple agents that can be used as elements of the Animat Brain architecture is presented in Section IV. The sketch program for future modeling of the cognitive evolution is proposed in Section V.

II. SIMULATION OF ADAPTIVE BEHAVIOR

In the early 1990s, the animat approach to artificial intelligence researches was proposed [4]. The term "animat" originates from two words: animal + robot = animat. The main goal of this field of research is "designing animats, i.e., simulated animals or real robots whose rules of behavior are inspired by those of animals. The proximate goal of this approach is to discover architectures or working principles that allow an animal or a robot to exhibit an adaptive behavior and, thus, to survive or fulfill its mission even in a changing environment. The ultimate goal of this approach is to embed human intelligence within an evolutionary perspective and to seek how the highest cognitive abilities of man can be related to the simplest adaptive behaviors of animals" [5].

Note that the ultimate goal of the animat approach is similar to the goals of modeling cognitive evolution.

The animat researches are highly interdisciplinary;

they are at the interface of neurosciences, cognitive science, ethology and ecology, on the one hand, of computer science and robotics, on the other hand.

Several universities and laboratories work actively in this field. These laboratories include:

• AnimatLab (Paris, leaded by J.-A. Meyer) that is conducting several projects on autonomous adaptive systems (http://animatlab.lip6.fr/index.en.html). AnimatLab approach implies that animat control system can formed by means of 1) leaning, 2) ontogenetic development, and 3) evolutionary search.

• Artificial Intelligence Laboratory, University of Zurich (http://www.ifi.unizh.ch/groups/ailab/, leaded by R. Pfeifer). The approach of this laboratory includes 1) modeling biological systems, 2) investigation of principles of intelligence that are common for animals and human, 3) using these principles at designing robots and other animats.

• Laboratory of Artificial Life and Robotics, Institute of Cognitive Sciences and Technologies (Rome, leaded by Stefano Nolfi, http://gral.istc.cnr.it/). This laboratory works in the field of evolutionary robotics.

• The Neurosciences Institute, founded by G. Edelman (USA, San Diego, http://www.nsi.edu/). The researchers of this institute develop generations of brain models and study adaptive behavior of Brain-Based Devices.

Investigations of adaptive behavior are based on nontrivial computational intelligence methods:

• Neural networks,

• Genetic algorithm and other methods of evolutionary computations,

• Classifier Systems [6],

• Reinforcement Learning [7].

It should be underlined that the *phenomenological approach* is often used at adaptive behavior simulations. This approach implies that there are formal rules for adaptive behavior and these rules are not necessarily connected with concrete microscopic neural or molecular structures of living organisms. It is natural to expect, that the phenomenological approach for adaptive behavior studies should be effective at modeling cognitive evolution (at least at the initial stages of investigations), because it is very difficult to understand cognitive features on the basis of the analysis of all complex variety of functioning of neurons, synapses, and molecules.

Analysis of adaptive behavior investigations shows that instead of large research work, the developed models are rather fragmentary. So, a general framework for modeling animat behavior would be useful. In order to develop such framework, the project "Animat Brain" was proposed. The architecture of animat control systems that are analyzed in this project is outlined in the next section.

III. PROJECT "ANIMAT BRAIN"

We have developed several versions of Animat Brain[1] design [8-10]. The Animat Brain architecture supposes that the animat control system consists of neural network (NN) blocks. Each block is a formal functional system (FS) [11] that forms the action $\mathbf{A}(t)$ in accordance with given state $\mathbf{S}(t)$, and makes prognosis of the next state for given $\mathbf{S}(t)$ and $\mathbf{A}(t)$. The state $\mathbf{S}(t)$ characterizes both external and internal environments of the animat at given moment of time t (t = 1,2,…). At any moment, only one FS is active, in which the current action is formed. There are connections between FSs; the active FS can transmit activation to every FS through these connections.

Each FS consists of two NNs: the actor and the predictor. Operation of the active FS can be described as follows. The state vector $\mathbf{S}(t)$ is fed to the FS input. The actor forms the action $\mathbf{A}(t)$ in accordance with given state $\mathbf{S}(t)$, i.e. the actor forms the mapping $\mathbf{S}(t)$ -> $\mathbf{A}(t)$. The predictor makes prognosis of the next state for given vectors $\mathbf{S}(t)$ and $\mathbf{A}(t)$, i.e. the predictor forms the mapping $\{\mathbf{S}(t), \mathbf{A}(t)\}$ -> $\mathbf{S}^{\mathbf{pr}}(t+1)$. Activation is transmitted from one FS to others in accordance with connectivity matrix $\|C_{ij}\|$, the value C_{ij} characterizes the probability that the j-th FS is activated by the i-th FS.

The animat receives reinforcements (rewards and punishments) which are related to animat needs.

It is supposed that there are primary and secondary repertoires of behaviors. The primary repertoire is formed by evolution: there is a population of animats and a set of FSs, synaptic weights of NNs and connectivity matrix $\|C_{ij}\|$ are adjusted during evolutionary processes. The secondary repertoire is formed by NN learning.

A particular version of the Animat Brain is based on adaptive critic design (ACD) [12]. In this case each FS is a simple ACD. In order to investigate features of such FS, we consider a simple model of evolving population of autonomous adaptive agents and study the ACD operation by means of computer simulation [13]. The results of these simulations are outlined below.

IV. BEHAVIOR OF SIMPLE ADAPTIVE AGENTS

A. Agent Control System

We consider the model of simple adaptive agents that make actions and predict results of actions. For concreteness we analyze agent-brokers that predict future changes of the stock price and try to increase they wealth by buying and selling stocks. The agent has its resource distributed into cash and stocks. The sum of cash and stocks is the net capital of the agent $C(t)$; t is time moment; t = 1,2,… The agent decision is to change the fraction of the agent's capital $u(t)$ that is currently invested in stocks. The environment is determined by the time series $X(t)$, where $X(t)$ is the stock price at the

[1] The term "Animat Brain" was proposed by Konstantin V. Anokhin.

moment t. The goal of the agent is to increase its capital $C(t)$ by changing the value $u(t)$. The capital dynamics is described by the equation:

$$C(t+1) = C(t)\ \{1 + u(t+1)\ \Delta X(t+1) / X(t)\}\ , \qquad (1)$$

where $\Delta X(t+1) = X(t+1) - X(t)$ is the current change of the stock price. It is convenient to use the logarithmic scale for the agent resource, i.e., $R(t) = \log C(t)$. The current agent reward is $r(t) = R(t+1) - R(t)$:

$$r(t) = \log\ \{1 + u(t+1)\ \Delta X(t+1) / X(t)\}\ . \qquad (2)$$

For simplicity we assume that the variable $u(t)$ takes only two values, $u(t) = 0$ or $u(t) = 1$.

The agent control system is the ACD that consists of two NNs: a model and a critic (Fig. 1). This architecture is slightly modified as compared with that of described in previous section. However, the modification is not large; it is due to the task that the considered agents solve.

The method of reinforcement learning [7] is used in this model. The method implies that agent learning is unsupervised (without a teacher); it is based on agent rewards $r(t)$.

The goal of the ACD is to maximize the utility function $U(t)$ [7]:

$$U(t) = \sum_{j=0}^{\infty} \gamma^j r(t+j), \qquad t = 1,2,..., \qquad (3)$$

where $r(t)$ is determined by (2), γ is the discount factor ($0 < \gamma < 1$). Making the realistic assumption $|\Delta X(t+1)| << X(t)$, we specify that the state $\mathbf{S}(t)$ is: $\mathbf{S}(t) = \{\Delta X(t), u(t)\}$.

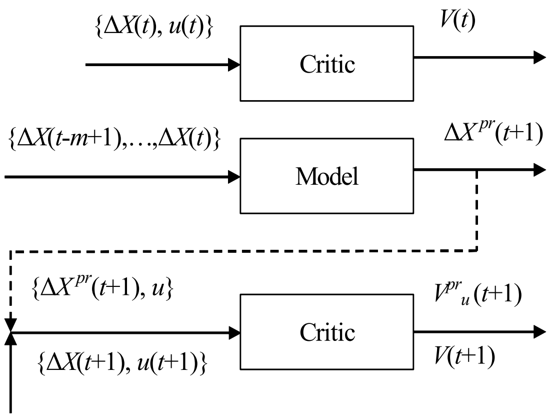

Fig. 1. The NN control system of the agent. The model predicts changes of the time series. The critic (the same NN is shown in two consecutive moments) forms the state value function for the current state $\mathbf{S}(t)$, the next state $\mathbf{S}(t+1)$, and its predictions $\mathbf{S}^{\mathbf{pr}}_{\mathrm{u}}(t+1)$ for two possible actions, $u = 0$ or $u = 1$.

The role of the model is to predict changes of the stock time series. The model is implemented as a NN with one hidden layer of tanh nodes and linear output.

The critic is intended to estimate the state value function $V(\mathbf{S})$ (estimate of U in (3)) for the current state $\mathbf{S}(t) = \{\Delta X(t), u(t)\}$, the next state $\mathbf{S}(t+1) = \{\Delta X(t+1), u(t+1)\}$, and its predictions $\mathbf{S}^{\mathbf{pr}}_{\mathrm{u}}(t+1) = \{\Delta X^{pr}(t+1), u\}$ for two possible actions, $u = 0$ or $u = 1$. The critic is also a NN of the same structure as the model.

The agent behavior is optimized by both individual learning and evolution of agent population.

B. Learning Algorithm

At any moment t, the following operations are performed:

1) The model predicts the next change of the time series $\Delta X(t+1)$.

2) The critic estimates the state value function for the current state $V(t) = V(\mathbf{S}(t))$ and the predicted states for both possible actions $V^{pr}_u(t+1) = V(\mathbf{S}^{\mathbf{pr}}_{\mathbf{u}}(t+1))$, where $\mathbf{S}^{\mathbf{pr}}_{\mathbf{u}}(t+1) = \{\Delta X^{pr}(t+1), u\}$, and $u = 0$ or $u = 1$.

3) The ε-greedy rule [7] is applied: the action corresponding to the maximum value $V^{pr}_u(t+1)$ is selected with probability $1-\varepsilon$, and an alternative action is selected with probability ε ($0 < \varepsilon << 1$).

4) The selected action is carried out. The transition to the next time moment $t+1$ occurs. The current reward $r(t)$ is calculated in accordance with (2) and received by ACD. The value $\Delta X(t+1)$ is observed and compared with its prediction $\Delta X^{pr}(t+1)$. The NN weights of the model are adjusted to minimize the prediction error using the error backpropagation.

5) The critic computes $V(t+1)$. The temporal-difference error is calculated [7]:

$$\delta(t) = r(t) + \gamma\ V(t+1) - V(t)\ . \qquad (4)$$

6) The weights of the NN of the critic are adjusted to minimize the temporal-difference error (4).

C. Evolutionary Algorithm

In addition to learning, the evolution of agent population is considered. An evolving population consists of n agents. Each agent has a resource $R(t)$ that changes in accordance with values of agent rewards $r(t)$.

Evolution passes through a number of generations, $n_g = 1,2,\ldots$ The duration of each generation n_g is T time steps (agent lifetime). At the beginning of any generation, initial resource of each agent is zero.

The initial synaptic weights (received at agent birth) of both NNs form the agent genome $\mathbf{G}$. The genome $\mathbf{G}$ does not change during the agent life. However, temporary synaptic weights of the NNs are changed during agent life via learning.

At the end of each generation, the agent having the maximum resource $R_{max}\ (n_g)$ is determined (the best agent of the generation n_g). This best agent gives birth to n children that constitute a new (n_g+1)-th generation. The children genomes **G** differ from their parent genome by small mutations.

D. Results of Simulations

In our computer simulations, we deal with two examples of the time series $X(t)$: sinusoidal (with the period of 20 time moments) and stochastic time series.

The simulation parameters are as follows: discount factor $\gamma = 0.9$, number of inputs of the model NN $m = 10$, number of hidden neurons of the model and critic $N_{hM} = N_{hC} = 10$, learning rate of the model and critic $\alpha_M = \alpha_C = 0.01$, parameter of the ε-greedy rule $\varepsilon = 0.05$, mutation intensity $P_{mut} = 0.1$, population size $n = 10$.

We analyze the following cases:

- Case L (pure learning); in this case we consider a single agent that learns by means of temporal difference method, see Eqs. (2)-(4);
- Case E (pure evolution), i.e., evolving population without learning;
- Case LE, i.e., learning combined with evolution, as described above.

The results are illustrated by Fig. 2, where the agent resource values attained during 200 time steps for these three cases of adaptation are shown. For the cases E and LE, we record the maximal value of agent resource in a population $R_{max}(n_g)$ at the end of each generation n_g . For the case L, we have just one agent whose resource is reset $R(T(n_g-1)+1) = 0$ after the passing of every $T = 200$ time steps; the index n_g is also incremented by one after every T time steps. In order to exclude the decrease of the value $R_{max}(n_g)$ due to the random choice of actions when applying the ε-greedy rule for the cases LE and L, we set $\varepsilon = 0$ after $n_g = 100$ for the case LE and after $n_g = 2000$ for the case L.

Analysis of agent actions demonstrates that both cases E and LE ensure the finding of optimal policy, i.e. the agent buys/sells stocks at prediction of stock price rises/falls (see the curves E and LE in Fig. 2). The pure learning is able to find only the satisfactory policy (see the curve L in Fig. 2), namely, the agent buys stock when stock price rises (or falls by a small amount) and sells stocks when stock price falls significantly.

Thus, pure learning is imperfect in our simulation, nevertheless, learning helps evolution to attain larger values of R_{max} faster (see curves E and LE in Fig. 2). We also confirm that learning helps evolution to find a good policy by some other simulations. For example, Fig. 3 demonstrates that during initial stages of evolution (n_g = 1-2) a satisfactory policy is found via learning (only after 200-300 time steps of the agent life). But at 5-th generation, the agents exhibit a satisfactory policy from the beginning of the generation. This phenomenon is known as the Baldwin effect [14, 15], i.e., initially acquired features become inhered.

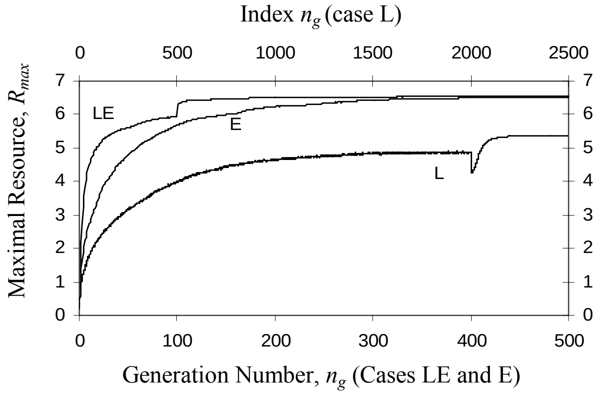

Fig. 2. The plots of $R_{max}(n_g)$ for the sinusoidal time series. The curves LE, E and L correspond to the cases of learning combined with evolution, pure evolution and pure learning, respectively. Results are averaged over 1000 simulations.

E. Discussion of simulation results

Thus, the behavior of simulated agents is interesting; in particular, nontrivial genetic assimilation of acquired features in Darwinian evolution takes place (Fig. 3). However, the agent cognitive abilities are small; learning processes are slow; so new versions of Animat Brain architectures should be analyzed. We continue the development of these architectures. In particular, we designed the control system architectures of animats that live in cellular world and began computer simulations of these animats.

V. SKETCH PROGRAM FOR FURTHER RESEARCH

Returning to modeling cognitive evolution, let us propose a sketch program for further researches.

1) *Modeling of adaptive behavior on the base of the project "Animat Brain".*

Such modeling can include simulation of simple cognitive features of animats that have natural needs: food, reproduction, safety.

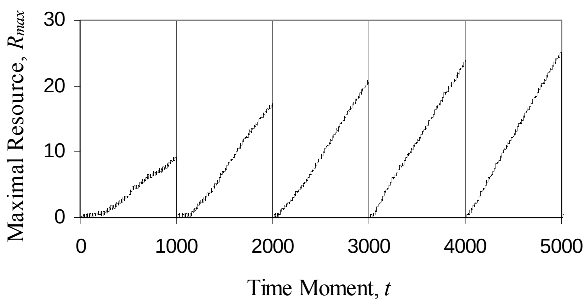

Fig. 3. The plots of $R_{max}(t)$ for the sinusoidal time series. The case LE, T = 1000. The ends of the generations are shown by vertical lines. At first two generations a satisfactory policy is found via learning. At 5-th generation the newborn agent "knows" a good police from the moment of its birth.

2) *Investigation of the transition from the physical level of information processing in nervous system of animals to the level of the generalized "notions".*

Such transition can be considered as emergence of the property "notion" in animal minds. The generalized "notions" are mental analogues of our words, which are not said by animals, but really used by them. Using notions, animals are able to have short descriptions of interactions with external environment.

3) *Investigations of processes of generating causal relations in animal memory.*

Storing relationships between the cause and the effect and adequate use of these relationships is one of key properties of active cognition of the external world regularities ("laws of nature") by animals. For example, such relationships are generated at the conditional reflex: the animal remembers the temporal relation between the conditional stimulus (CS) and the unconditional stimulus (US). This allows it to predict events in the external world and adequately use these predictions.

Natural next step is the transition from memorizing separate causal relations to systems of logic conclusions.

4) *Investigations of logic conclusions in animal minds.*

Actually, at classical conditioning, the animals do a "logic conclusion": {CS, CS -> US} => US or "If the conditional stimulus takes place, and the conditional stimulus is followed by the unconditional one, then the occurrence of the unconditional stimulus is expected". We can even state that such conclusions are similar to logical conclusions of mathematicians proving the theorems. It is important to understand, what are systems of these conclusions, to what extent the "animal logic" is similar to our, human logic.

The listed items outline steps of investigations from simplest forms of animal behavior to logical rules that are used at proofs of mathematical theorems. It should be noted that certain researches on these topics are conducted already; however, a sequence of serious canonical models is absent yet.

VI. Conclusion

Thus, approaches to modeling cognitive evolution have been discussed. This modeling is related with foundation of science and with scientific cognition. The initial steps towards modeling cognitive evolution have been already done in the research area "Simulation of Adaptive Behavior". Certain cognitive features can be modeled in the framework of the project "Animat Brain". Some researches on this project have been outlined in the current work. The sketch program for further modeling cognitive evolution has been proposed. The program describes reasonable research steps from simple animal cognitive abilities to mathematical deductions.

Acknowledgment

This work is supported by the Russian Foundation for Basic Research, Grant No 07-01-00180 and the Program of the Presidium of the Russian Academy of Science "Fundamental problems of informatics and informational technologies", Project No 16.2.45.

References

1. Web Site of World Congress on Computational Intelligence (WCCI 2008): http://www.wcci2008.org.
2. J.S. Albus, G.A. Bekey, J.H. Holland, N.G. Kanwisher, J.L. Krichmar, M. Mishkin, D.S. Modha, M.E. Raichle, G.M. Shepherd, and G. Tononi, "A Proposal for Decade of Mind Initiative," Science, vol. 317, no. 5843, p. 1321, September 2007.
3. Challenges for Computational Intelligence, W. Duch and J. Mandziuk, Eds. Berlin, Heidelberg, New York: Springer, 2007.
4. From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior, J.-A., Meyer and S.W. Wilson, Eds. Cambridge: MIT Press, 1991.
5. J.Y. Donnart and J.-A. Meyer, "Learning reactive and planning rules in a motivationally autonomous animat," IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, vol. 26, pp. 381-395. 1996.
6. J.H. Holland, K.J. Holyoak, R.E. Nisbett, and P. Thagard, Induction: Processes of Inference, Learning, and Discovery. Cambridge: MIT Press, 1986.
7. R. Sutton and A. Barto, Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998.
8. K.V. Anokhin, M.S. Burtsev, I.Yu. Zarayskaya, A.O. Lukashev, and V.G. Red'ko, "Project 'Animat Brain': Development of a Model of Adaptive Behavior on the Basis of the Functional System Theory," in Proceedings of 8-th Russian Conference on Artificial Intelligence. Moscow: Fizmatlit, 2002, vol. 2, pp. 781-789 (In Russian).
9. V.G. Red'ko, D.V. Prokhorov, and M.S. Burtsev. "Theory of functional systems, adaptive critics and neural networks," in Proceedings of International Joint Conference on Neural Networks (IJCNN 2004), Budapest, 2004, pp. 1787-1792.
10. V.G. Red'ko, K.V. Anokhin, M.S. Burtsev, A.I. Manolov, O.P. Mosalov, V.A. Nepomnyashchikh, and D.V. Prokhorov, "Project 'Animat Brain': Designing the Animat Control System on the Basis of the Functional Systems Theory," in Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior. M.V. Butz, O. Sigaud, G. Pezzulo, and G. Baldassarre, Eds. Berlin, Heidelberg: Springer Verlag, 2007. pp. 94-107.
11. P.K. Anokhin, Biology and Neurophysiology of the Conditioned Reflex and Its Role in Adaptive Behavior. Oxford: Pergamon, 1974.
12. D.V. Prokhorov and D.C. Wunsch, "Adaptive critic designs," IEEE Trans. Neural Networks, vol. 8, no.5, pp. 997-1007, 1997.
13. V.G. Red'ko, O.P. Mosalov, and D.V. Prokhorov, "A Model of Evolution and Learning," Neural Networks, vol. 18, pp. 738-745, 2005.
14. J.M. Baldwin, "A New Factor in Evolution," American Naturalist, vol. 30, pp. 441-451, 1896.
15. Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect, P. Turney, D. Whitley, R. Anderson, Eds. Special Issue of Evolutionary Computation, vol. 4, no 3, 1996.