To be published in a special issue of *Collection of Czechoslovak Chemical Communications*, October 2006

ARTIFICIAL CHEMISTRY AND REPLICATOR THEORY OF COEVOLUTION OF GENES AND MEMES

Vladimír KVASNIČKA[1] and Jiří POSPÍCHAL[2]

Institute of Applied Informatics, Faculty of Informatics and Information Technologies
Slovak Technical University, 84216 Bratislava, Slovakia
e-mail: [1]kvasnicka@fiit.stuba.sk, [2] pospichal@fiit.stuba.sk

Dedicated to the memory of Prof. Jaroslav Koutecký, founder of Czech and Slovak quantum chemistry. His personal and professional contacts with Slovak chemists in the sixties initiated the formation of small groups of young quantum chemists. These groups, islands of positive deviations in theoretically oriented chemistry, were latterly almost all transformed into serious scientific laboratories, very active in many-body perturbation theories.

A simple replication theory of coevolution of genes and memes is proposed. A population composed of pairs of genes and memes, the so-called m-genes, is postulated as a subject of Darwinian evolution. Three different types of operations over m-genes are introduced: Replication (an m-gene is replicated with mutations onto an offspring m-gene), interaction (a memetic transfer from a donor to an acceptor), and extinction (an m-gene is eliminated). Computer simulations of the present model allow to identify different mechanisms of gene and meme coevolution.
Key words: Chemostat; Replicator; Meme; Memetics; Coevolution; Darwinian evolution; Social learning; Evolvability.

Physicochemical theory of replicators elaborated in the seventies by Eigen and Schuster[14,15] can be used as a very simple and effective formal tool for computational (*in silico*) simulations of Darwinian evolution. This approach will be applied in the present communication to an evolutionary study of an interaction between genes and memes. Memes [2,5,10,11] rank among very controversial concepts of theory of human culture. This meme concept was initially introduced by evolutionary biologist Richard Dawkins in his seminal book [10] *The Selfish Gene* as an information unit copied from one human brain to another human brain by imitation. Moreover, he postulated that memes have properties of replicators and therefore their population may be a subject of evolution [5,11]. English dictionaries define meme as an element of behaviour or culture passed on by imitation or other non-genetic means and memetics as the study of memes and their social and cultural effects. Memetics offers conceptually simple explanations of the nature and evolution of human culture; a paradigm of memes as replicators looks very attractive for audience mostly outside social sciences. On the other hand, memetics is strongly rejected by many social scientists as a theoretical approach based on dubious postulates and superficial analogies. Hot disputes continue in the following three directions: (1) Whether culture is properly seen as composed of independently transmitted information units, (2) whether memes have the necessary qualification to serve as replicators, and (3) whether evolutionary approaches such as memetics offer the most natural framework for a theory of culture. Recent Aunger's *Darwinizing Culture*[2] is the first book to attempt a thorough critical and conservative appraisal of the potential of memetics. This text summarizes the points of agreement and/or disagreement on memetics and concludes with some suggestions for the progressive directions, particularly with respect to the means by which empirical research in this area may be undertaken.

The purpose of this communication is to suggest a model of coevolution of genes and memes [6, 8, 13, 17, 30], where it is postulated that a memetic environment may change selection of genes; in other words, gene fitness is affected by memes. This field of cultural anthropology was initiated by Cavalli-Sforza and Feldman [8]. In their seminal book *Cultural Transmission and Evolution: A Quantitative Approach* they used the approach based on differential equations from population dynamics. This approach was in the beginning of nineties extended by Zhivotovsky and Feldman[18] by probabilistic analysis. Unfortunately, both these approaches are strongly phenomenological, they had to specify ad-hoc concepts of population dynamics, which were crucial for cultural phenomena. In the present communication we substituted this classical differential equations and probabilistic approach by a modern multiagent evolutionary paradigm [9], where dynamics is governed by interactions between single agents. The rules of these interactions can be simply formulated and modeled. Agent based approach to modeling of social dynamics is best known from the work of Epstein and Axtell[16] oriented towards students of computer science**;** however, they did not specifically model gene-culture coevolution.

In our approach memes are not independent of genes, they form pairs composed of a gene and a meme (these pairs are called the m-genes). The subject of Darwinian evolution is a population composed of these pairs instead of two relatively independent populations composed of genes and memes. In the proposed coevolutionary approach fitness of m-genes is composed of two parts: fitness of the respective gene itself and interaction fitness between the respective gene and the respective meme. This second term reflects an influence of memes onto genes and may be characterized as a (direct or indirect) cultural influence on genes through a (cultural) modification of environment (niche) in which genes exist. The memes coevolve simultaneously with genes; a Darwinian selection exists only for gene-meme composites. Three different types of m-gene “transformations” are postulated:

(1) Replication – randomly selected m-gene is copied with mutations, creating another m-gene. There exist two possibilities how to create a meme of emerging offspring, the first one is a simple copy with mutations of parental meme, whereas the second possibility consists in local optimization of the new meme in nearest neighborhood of the parental meme. This second possibility can be interpreted as a kind of social learning, where an offspring adapts the parental meme to a form more appropriate for its gene. The replication causes the so-called vertical transfer of memes from parents onto offspring.

(2) Interaction – two randomly selected m-genes (that are classified as a donor and an acceptor) are transformed in such a way that an acceptor meme is substituted by a modified donor meme. Similarly as for replication, the new donor meme can be created in two possible ways. The first one consists in a simple mutation of the donor meme, whereas the second one consists in local optimization of the donor meme. In other words, the approach to creation of donor meme can be considered as a genuine social learning, where the acceptor adapts the donor meme to an optimal form with respect to its gene. The interaction process performs the so-called horizontal transfer of memes from donor to acceptor.

(3) Extinction – a randomly selected m-gene is eliminated from the population. This simple process is immediately applied after a replication process is used, which would otherwise increase the number of m-genes in population by one. This means that extinction process ensures a constant number of m-genes in population.

We have specified two types of the social learning processes that may appear when either vertical or horizontal meme transfers are applied; both learning processes are connected with local optimization of memes (for their fixed gene from the m-gene pair). There exists also another quite different possibility consisting in local optimization of gene with respect to a fixed respective meme (when genes code architecture of cognitive organs – brains). This interesting alternative approach to the learning of gene parts of m-genes is applicable to calculation of fitness, in particular when the gene-part of fitness is locally optimized. We get the so-called effective fitness that reflects evolvability of m-genes on fitness landscape (this approach to calculation of fitness is often called in the literature the Baldwin effect [3, 4, 27].

The fitness of m-genes is evaluated by making use of a generalized version of Kauffman *KN* functions[1,24], which recently became very popular for construction of "realistic" rugged fitness landscapes (see Appendix A). The same type of fitness landscape was recently used by Bull et al.[7] in their simulation of coevolution of genes and memes. They studied coevolution, where two independent populations are considered, one for genes and another one for memes. The present approach to coevolution of genes and memes is quite different from Bull's approach; we will study only one population composed of pairs of a gene and a meme (m-genes), and this population is the subject of Darwinian evolution.

A COEVOLUTION OF GENES AND MEMES

Let us consider two different sets that are composed of all possible genes and memes (see Appendix B), $\mathcal{G}=\{\boldsymbol{g},\boldsymbol{g}',...\}$ and $\mathcal{M}=\{\boldsymbol{m},\boldsymbol{m}',...\}$. A composition, called the *m-gene*, is defined by $(\boldsymbol{g},\boldsymbol{m})\in\mathcal{G}\times\mathcal{M}$, then a *chemostat* is a finite multiset composed of *A* m-genes (see Fig. 1)

$$\mathcal{P}=\{(\boldsymbol{g}_1,\boldsymbol{m}_1),(\boldsymbol{g}_2,\boldsymbol{m}_2),...,(\boldsymbol{g}_A,\boldsymbol{m}_A)\} \tag{1}$$

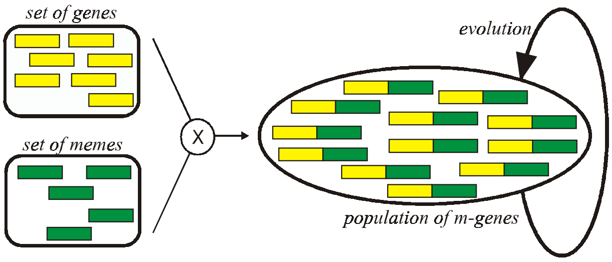


Fig.1.
A population $\mathcal{P}$ is composed of m-genes that are represented by compositions (determined by a "cross" product) of genes and memes. Darwinian evolution runs over the chemostat $\mathcal{P}$ and not separately over genes and memes. This approach is a manifestation of our postulation that memes may exist only in coexistence with genes.

Each m-gene (***g***,***m***) is evaluated by a *fitness function f*

$$f\left[(\boldsymbol{g},\boldsymbol{m})\right]=(1-\omega)F(\boldsymbol{g})+\omega H(\boldsymbol{g},\boldsymbol{m}) \tag{2}$$

where the function *F* evaluates a gene ***g*** by a nonnegative real number from the closed interval [0,1] specifying "an evolutionary quality" of a gene ***g***, $F:\mathcal{G}\rightarrow[0,1]$, and the function *G* evaluates an m-gene by a real number from [0,1] that corresponds to a "cultural" interaction between a gene $\boldsymbol{g}$ and a meme $\boldsymbol{m}$, $H:\mathcal{G}\times\mathcal{M}\rightarrow[0,1]$. The parameter $0\leq\omega\leq 1$ is a measure of interaction between genes and memes (it will be called the *cultural parameter*), a maximal value of fitness is specified by $f(\boldsymbol{g},\boldsymbol{m})\leq f_{\max}=1$. The fitness may be formally specified by $f:\mathcal{G}\times\mathcal{M}\rightarrow[0,1]$. Darwinian evolution[32] over the chemostat (population) $\mathcal{P}$ may be considered as an optimization algorithm that looks for a global maximum on composed $\mathcal{G}\times\mathcal{M}$ fitness landscape

$$\left(\boldsymbol{g}_{\text{opt}},\boldsymbol{m}_{\text{opt}}\right)=\arg\max_{(\boldsymbol{g},\boldsymbol{m})\in\mathcal{G}\times\mathcal{M}} f\left[(\boldsymbol{g},\boldsymbol{m})\right] \tag{3}$$

The concept of evolution as an optimization process was introduced already by Wright[32]. The discrete and very complex optimization problem from Eq. (3) belongs to a class of hard numerical NP-complete problems. This is the main reason why the optimization problems like (3) are solved by the so-called *evolutionary algorithms* [22]. We will use a very simple version of evolutionary algorithm tightly related to the idea of *chemostat* [12, 19], which is very popular in the so-called artificial chemistry [12, 25].

Let us postulate three elementary "reactions", which are fundamental for the proposed "chemostat" evolutionary algorithm over a population of m-genes:

(1) **Replication**, an m-gene (***g***,***m***) is replicated to another m-gene (***g***',***m***') (see fig. 2)

$$\underbrace{(\boldsymbol{g},\boldsymbol{m})}_{parent}\rightarrow\underbrace{(\boldsymbol{g},\boldsymbol{m})}_{parent}+\underbrace{(\boldsymbol{g}',\boldsymbol{m}')}_{offspring} \tag{4}$$

Components of the offspring m-gene (***g'***,***m'***) from the right-hand side of (4) are closely related to the respective parent components (usually with a small "distance") and are specified as follows:

$$\boldsymbol{g}'=O_{\text{mut}}^{(\text{gene})}(\boldsymbol{g}) \tag{5a}$$

$$\boldsymbol{m}'=\arg\max_{\tilde{\boldsymbol{m}}\in U(\boldsymbol{m})} H(\boldsymbol{g}',\tilde{\boldsymbol{m}}) \tag{5b}$$

where $O_{\text{mut}}^{(\text{gene})}$ is a gene mutation operator[22] (specified by a probability $P_{\text{mut}}^{(\text{gene})}$). The second formula means that an offspring meme $\boldsymbol{m}'$ is created as a solution of local maximization problem in the neighborhood $U(\boldsymbol{m})=\left\{\tilde{\boldsymbol{m}}=O_{\text{mut}}^{(\text{meme})}(\boldsymbol{m})\right\}$ with fixed cardinality U_{card} (considered as a parameter of the method), and $O_{\text{mut}}^{(\text{meme})}$ is a meme mutation operator (specified by a probability $P_{\text{mut}}^{(\text{meme})}$). This means that an offspring meme $\boldsymbol{m}'$ is not a simple mutation of the parent meme $\boldsymbol{m}$, but it results from local hill-climbing "memetic" learning process with respect to a fixed offspring gene $\boldsymbol{g}'$. In other words, we can say, that the offspring does not automatically accept a parent meme, but it is trying to locally optimize (a kind of social learning process) the parent meme $\boldsymbol{m}$ with respect to its gene $\boldsymbol{g}'$ transferred genetically from its parent. If we put U_{card}=1, then Eq. (5b) is reduced to simple formula $\boldsymbol{m}'=O_{\text{mut}}^{(\text{meme})}(\boldsymbol{m})$.

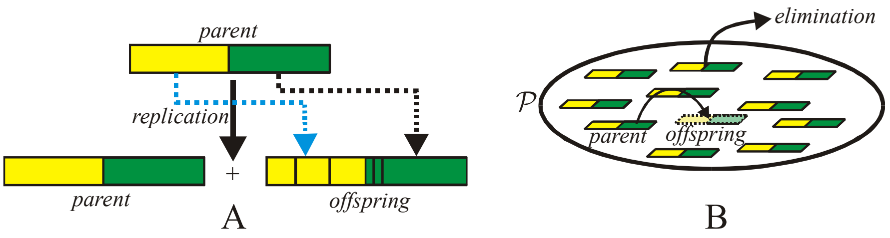


Fig. 2.
Diagrammatic interpretation of the replication process (4). (A) Randomly selected parent m-gene is replicated, both components of an offspring m-gene are specified as slightly mutated versions of parent counterparts (manifested by the appearance of thin vertical columns in the offspring). (B) Schematic outline of the replication process in a population, randomly selected parent m-gene is replicated (with mutations) to an offspring m-gene, and moreover, randomly selected m-gene is eliminated from the population.

Formally, the replication process may be expressed by a stochastic operator

$$\left(\boldsymbol{g}',\boldsymbol{m}'\right)=O_{\text{repli}}\left(\boldsymbol{g},\boldsymbol{m}\right) \tag{6}$$

In a limiting case, a resulting m-gene (***g***’,***m***’) may be simply equal to the argument (***g***,***m***). We have to emphasize that for replication processes we distinguish two types of mutation operators, the first one is a gene mutation and the second one is a meme mutation, where both are characterized by different mutation probabilities $P_{\text{mut}}^{(\text{gene})}$ and $P_{\text{mut}}^{(\text{meme})}$. This distinguishing allows to separate processes of gene and meme mutations; it implies that we may introduce different “evolutionary” rates of genes and memes. The probability of an application of the replication operator (an analog to the kinetic rate constant k) is specified by

$$\text{prob}_{(\boldsymbol{g},\boldsymbol{m})(\boldsymbol{g}',\boldsymbol{m}')}=\exp\left[\alpha\left(f\left(\boldsymbol{g},\boldsymbol{m}\right)-f_{\max}\right)\right] \tag{7}$$

where $f_{\max}=1+\omega$ is an estimated maximal value the fitness and $\alpha>0$ is the so-called *slope parameter* (its greater values, $\alpha\rightarrow\infty$, cause a negligible value of probability, $\text{prob}_{(\boldsymbol{g},\boldsymbol{m}),(\boldsymbol{g}',\boldsymbol{m}')}\rightarrow 0$, for $f\left(\boldsymbol{g},\boldsymbol{m}\right)<f_{\max}$). It means that replication is more probable for m-genes with fitness close to its maximal value $f_{\max}$. The preferable selection of m-genes with higher fitness for replication is of great importance for achieving a global optimum (or its closely related approximation) of the optimization problem (3) by the proposed evolutionary algorithm based on the metaphor of “chemostat” [12, 19].

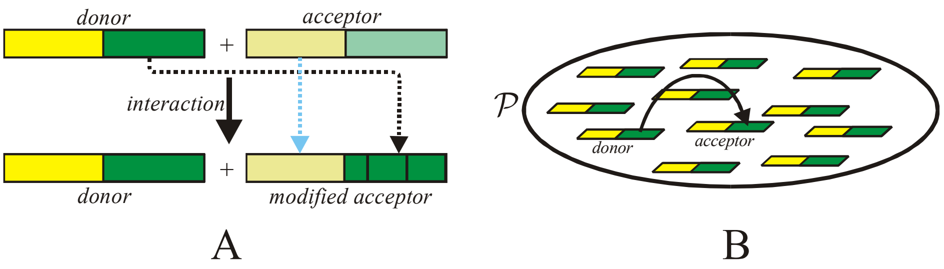


Fig. 3.
Diagrammatic interpretation of the interaction process (8). (A) Randomly selected donor transfers its meme ***m*** to a randomly selected acceptor and then this meme is locally optimized with respect to the original acceptor gene ***g***’. (B) Schematic outline of the interaction process in a population, two randomly selected m-genes “interact” such that a meme from the donor is transferred (represented by an oriented arrow) to the acceptor and the transferred meme is then locally optimized with respect to the acceptor gene.

(2) **Interaction**, two m-genes (***g***,***m***) (called the donor) and $\left(\boldsymbol{g}',\boldsymbol{m}'\right)$ (called the acceptor) interact in such a way that there exists an oriented “memetic” transfer of a meme from the donor to the acceptor (see Fig. 3)

$$\underbrace{\left(\boldsymbol{g},\boldsymbol{m}\right)}_{donor}+\underbrace{\left(\boldsymbol{g}',\boldsymbol{m}'\right)}_{acceptor}\rightarrow\underbrace{\left(\boldsymbol{g},\boldsymbol{m}\right)}_{donor}+\underbrace{\left(\boldsymbol{g}',\boldsymbol{m}''\right)}_{\substack{modified\\acceptor}} \tag{8}$$

The memetic part of acceptor is created from the transferred donor meme by its local optimization with respect to its fixed gene part ***g***′ (cf. eq. (5b))

$$\boldsymbol{m}'' = \arg\max_{\tilde{\boldsymbol{m}} \in U(\boldsymbol{m})} H\left(\boldsymbol{g}', \tilde{\boldsymbol{m}}\right) \tag{9}$$

In similar way as in (5b), it means that an acceptor meme $\boldsymbol{m}''$ is created as a solution of local maximization problem in the neighborhood $U(\boldsymbol{m}) = \left\{\tilde{\boldsymbol{m}} = O_{\text{mut}}^{(\text{meme})}(\boldsymbol{m})\right\}$ with fixed cardinality (considered as a parameter of the method) U_{card}. The modified acceptor is composed of a meme, which originates from the donor and is modified by local optimization with respect to original acceptor gene $\boldsymbol{g}'$. This interaction process corresponds to one of fundamental properties of memetic systems [5]; in particular, memes are spreading throughout population not only "vertically" by replication, but also "horizontally" when a donor offers its meme to other acceptors by the interaction process. The interaction process may be expressed by a stochastic operator

$$\left(\boldsymbol{g}', \boldsymbol{m}''\right) = O_{interact}\left(\left(\boldsymbol{g}, \boldsymbol{m}\right), \left(\boldsymbol{g}', \boldsymbol{m}'\right)\right) \tag{10}$$

In a limiting case, if this operator could not be applied to arguments, then the resulting m-gene ($\boldsymbol{g}'$,$\boldsymbol{m}''$) is simply equal to the second argument (acceptor) ($\boldsymbol{g}'$,$\boldsymbol{m}'$). In order to get complete specification of the interaction operator, we have to introduce the probability of its application to arguments

$$\text{prob}_{(\boldsymbol{g},\boldsymbol{m})(\boldsymbol{g}',\boldsymbol{m}')} = \max\left\{0, 1 - \exp\left[\alpha\left(f\left(\boldsymbol{g}', \boldsymbol{m}'\right) - f\left(\boldsymbol{g}, \boldsymbol{m}\right)\right)\right]\right\} \tag{11}$$

where $\alpha > 0$ is the so-called *slope parameter*. Its greater values cause a dichotomic behavior (either zero or unit) of the probability. The probability is positive if a fitness of donor is greater than a fitness of acceptor, $f\left(\boldsymbol{g}, \boldsymbol{m}\right) > f\left(\boldsymbol{g}', \boldsymbol{m}'\right)$, and it is proportional to a difference $f\left(\boldsymbol{g}, \boldsymbol{m}\right) - f\left(\boldsymbol{g}', \boldsymbol{m}'\right)$. It means that there exists a "one-way" memetic transfer from a donor, which must have a greater fitness than the acceptor.

(3) **Extinction**, an m-gene ($\boldsymbol{g}$,$\boldsymbol{m}$) is removed from the population $\mathcal{P}$

$$\underbrace{\left(\boldsymbol{g}, \boldsymbol{m}\right)}_{parent} \rightarrow \varnothing \tag{12}$$

Each replication process is automatically accompanied by extinction. Since the replication increases the number of m-genes in population by one (cf. Eq. (4)), application of extinction (a randomly selected m-gene is eliminated from the population) ensures a constant number of m-genes in chemostat (i.e. the size of chemostat – population is invariant in the course of evolution).

At the end of this Section we add a few remarks on the possibility to reflect the evolvability of a particular m-gene on the fitness landscape (Baldwin effect [3, 4, 27]). We may introduce the effective fitness, which is assigned to each m-gene ($\boldsymbol{g}$,$\boldsymbol{m}$) as a local maximum of the standard fitness (2) within the neighborhood $U(\boldsymbol{g}) = \left\{\boldsymbol{g}' = O_{mut}(\boldsymbol{g})\right\}$ with respect to a fixed meme $\boldsymbol{m}$

$$f_{\text{eff}}\left[\left(\boldsymbol{g}, \boldsymbol{m}\right)\right] = \max_{\boldsymbol{g}' \in U(\boldsymbol{g})} f\left[\left(\boldsymbol{g}', \boldsymbol{m}\right)\right] \tag{13}$$

where the cardinality $U_{\text{card}}^{(\text{eff})}$ of the neighborhood $U(\boldsymbol{g})$ is kept fixed in all effective fitness evaluations. It means that the concept of effective fitness reflects local properties of fitness landscape; if for an m-gene there exists in its closest neighborhood a fitter possibility, then the respective m-gene is evaluated by an effective fitness greater than its original counterpart. In an alternative way, we may say that the effective fitness is closely related to a possible "*evolvability*" of the respective m-gene. If its effective fitness is greater than its standard value (i.e. $f_{\mathrm{eff}}(\boldsymbol{g},\boldsymbol{m})>f(\boldsymbol{g},\boldsymbol{m})$), then the m-gene has a chance to be evolved in its closest neighborhood such that its standard fitness will increase; on the other hand, if a particular m-gene is situated on a local maximum of the fitness landscape (i.e. $f(\boldsymbol{g},\boldsymbol{m})=f_{\mathrm{eff}}(\boldsymbol{g},\boldsymbol{m})$), then the respective m-gene is not evolvable in its closest neighborhood.

In the proposed evolutionary algorithm based on a metaphor of chemostat, selection pressure in population of m-genes is created by replication and interaction processes based on fitness. M-genes with a greater fitness have a greater chance to take part in a replication or interaction process (a measure of quality of m-genes); on the other hand, m-genes with a small effective fitness are rarely used in the replication process or as a donor in the interaction. This simple manifestation of Darwin's natural selection ensures gradual evolution of the whole population. In the present approach the mentioned principle of fitness selection of m-genes is preserved, but it is now combined with an additional selection pressure due to constancy of the number of m-genes in the chemostat. An m-gene outgoing from the replication reaction eliminates a randomly selected m-gene. Moreover, we have to distinguish between the performance of replication and interaction processes; the replication process should be applied with substantially higher frequency than interaction process, that is simple replications of parents into an offspring are more frequent than transfers of a meme from donors to acceptors (see Algorithm 1).

Algorithm 1.

```
1  chemostat $\mathcal{P}$ is randomly generated;
2  for t:=1 to $t_{max}$ do
3  begin $(\mathbf{g},\mathbf{m}):=O_{select}(\mathcal{P})$;
4        if random<prob$(\mathbf{g},\mathbf{m})$ then
5        begin $(\mathbf{g}',\mathbf{m}'):=O_{repli}(\mathbf{g},\mathbf{m})$;
6              $(\mathbf{g}'',\mathbf{m}''):=O_{select}(\mathcal{P})$;
7              $(\mathbf{g}'',\mathbf{m}'')\leftarrow(\mathbf{g}',\mathbf{m}')$;
8        end;
9        while random<$P_{interact}$ then
10       begin $(\mathbf{g},\mathbf{m}):=O_{select}(\mathcal{P})$; $(\mathbf{g}',\mathbf{m}'):=O_{select}(\mathcal{P})$;
12             if random<prob$((\mathbf{g},\mathbf{m}),(\mathbf{g}',\mathbf{m}'))$ then
13             begin $(\mathbf{g}',\mathbf{m}''):=O_{interact}((\mathbf{g},\mathbf{m}),(\mathbf{g}',\mathbf{m}'))$;
14                   $(\mathbf{g}',\mathbf{m}')\leftarrow(\mathbf{g}',\mathbf{m}'')$;
15             end;
16       end;
17 end;
```

The algorithm is initialized by a population composed of randomly generated m-genes that are all evaluated by the fitness (see row 1). The algorithm is composed of two blocks that are activated with different probabilities, the first one (with probability `prog(g,m)`, which reflects fitness of the corresponding m-gene in such a way that the resulting probability is proportional to the fitness) for a replication (rows 5 – 8) and the second one (with probability $P_{interact}$) for an interaction (rows 9-16). The replication block is initiated by random selection (realized by an operator O_{select}) of an m-gene $(\mathbf{g},\mathbf{m})$; this m-gene is further replicated (with probability specified by `prob(g,m)`) by an operator O_{repli}. The resulting product $(\mathbf{g}',\mathbf{m}')$ is evaluated by a fitness and then it is returned to the population such that it eliminates a randomly selected m-gene $(\mathtt{g''},\mathtt{m''})$. The interaction block is repeated with probability $\mathtt{P_{interact}}$, two m-genes are randomly selected and then both undergo an interaction.

TWO FURTHER GENERALIZATIONS OF THE PRESENT THEORY

One of the most serious restrictions of the present theory is a postulate that a particular gene interacts only with one meme. In many cases of theoretical interest (in theoretical memetics) this severe restriction seriously limits the applicability of the present theory; therefore, we suggest its generalized version, which to some extent overcomes the mentioned restriction. Moreover, this simple generalization nicely demonstrates flexibility and effectivity of the present theoretical approach, which allows not only simple computational simulations, but also its effortless modifications and augmentations.

The first purpose of this Section is to outline a simple generalization of the theory such that a gene may be simultaneously coupled with P different memes. Let us introduce P meme sets

$$\mathcal{M}^{(1)}=\{\boldsymbol{m}_1,\boldsymbol{m}_1',...\},\ \mathcal{M}^{(2)}=\{\boldsymbol{m}_2,\boldsymbol{m}_2',...\},...,\ \mathcal{M}^{(P)}=\{\boldsymbol{m}_P,\boldsymbol{m}_P',...\} \tag{14}$$

A generalized form of the concept of m-gene looks as follows:

$$(\boldsymbol{g};\boldsymbol{m}_1,\boldsymbol{m}_2,...,\boldsymbol{m}_P)\in\mathcal{G}\times\mathcal{M}^{(1)}\times\mathcal{M}^{(2)}\times...\times\mathcal{M}^{(P)} \tag{15}$$

Each generalized m-gene is evaluated by fitness

$$f\left[(\boldsymbol{g};\boldsymbol{m}_1,\boldsymbol{m}_2,...,\boldsymbol{m}_P)\right]=(1-\omega)F(\boldsymbol{g})+\frac{\omega}{P}\sum_{i=1}^{P}H_i(\boldsymbol{g},\boldsymbol{m}_i) \tag{16}$$

This means that a fitness is composed of a fitness part assigned to the gene $\boldsymbol{g}$ and P interaction parts H_i, which correspond to a cultural interaction of the respective gene $\boldsymbol{g}$ and the ith meme $\boldsymbol{m}_i$. Each interaction term on the right-hand side of (16) is specified by a distinct function that maps memes onto closed interval [0,1], $H_i:\mathcal{G}\times\mathcal{M}^{(i)}\to[0,1]$ (for i=1,2,…,P). The above definition (16) of fitness may be simply reformulated in form, where an m-gene is composed of a variable number of memes.

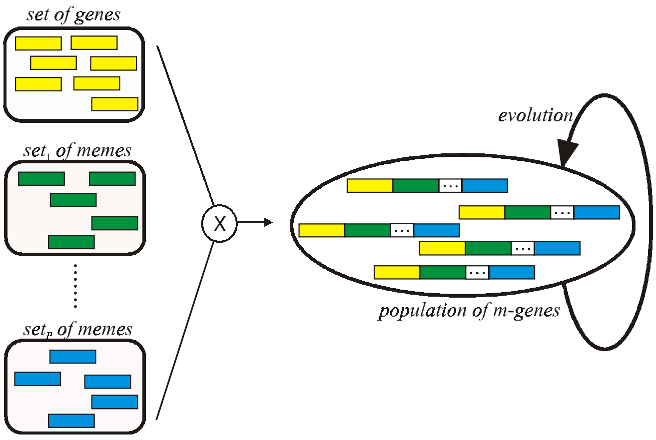


Fig. 4.

In the generalized approach a population $\mathcal{P}$ is composed of m-genes which are represented by compositions (determined by a “cross” product) of genes and P different kinds of memes. Each m-gene is composed of a gene and P memes which correspond to different memetic kinds.

As a simple application of this generalized coevolutionary theory of genes with two or more different memes we present its specification when an environmental niche is explicitly considered. In particular, many biologists or social scientists [2, 30] doubt whether memes may directly affect fitness of genes. They suggest overcoming this serious problem of the whole memetics by introducing an environmental niche. Then they postulate that the memes may affect only this niche and gene fitness is specified by its structure plus an interaction between gene and environmental niche (see diagram A in Fig. 5). This means that within the model a vicarious interaction does not exist between genes and memes, but it is mediated by an environmental niche. Subjects of Darwinian evolutions are composites called mn-genes that are created by a gene, environmental niche, and meme (see diagram B in Fig. 5). Their fitness is calculated as follows

$$f\left[\left(\boldsymbol{g},\boldsymbol{n},\boldsymbol{m}\right)\right]=\left(1-\omega\right)F\left(\boldsymbol{g}\right)+\frac{\omega}{2}\left(H_{GN}\left(\boldsymbol{g},\boldsymbol{n}\right)+H_{NM}\left(\boldsymbol{n},\boldsymbol{m}\right)\right) \qquad (17)$$

where H_{GN} (H_{NM}) represents an interaction term between gene and environmental niche (environmental niche and meme). We see that a memetic part of mn-gene does not interact directly with a particular gene, but it interacts vicariously through an environmental niche.

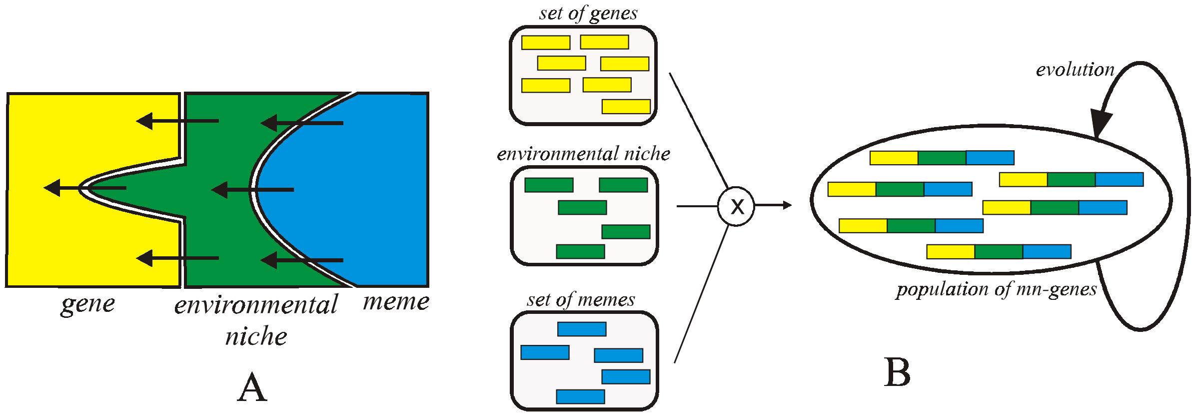


Fig. 5.
Diagram A shows a type of used interaction in an cooperative model with environmental niche, which mediates an interaction between memes and genes. Diagram B is a special kind of Fig. 4, where composite genes are created from three different parts: gene, environmental, and meme.

The second purpose of the present Section is to construct, after Eigen’s replicator equations, a system of kinetic differential equations that describe a dynamics of the present system of m-genes. A concentration of an m-gene $(\boldsymbol{g},\boldsymbol{m})$ in a time t in the population $\mathcal{P}$ is specified by a variable $c_{(\boldsymbol{g},\boldsymbol{m})}(t)$, and then, applying the mass-action law of chemical kinetics to all processes (4), (8), and (12), we get the following system of differential equations

$$\begin{aligned}\dot{c}_{(\boldsymbol{g},\boldsymbol{m})} &= c_{(\boldsymbol{g},\boldsymbol{m})}\left(k_{(\boldsymbol{g},\boldsymbol{m}),(\boldsymbol{g},\boldsymbol{m})}-\phi\right)+\sum_{\substack{(\boldsymbol{g}',\boldsymbol{m}')\\ \neq(\boldsymbol{g},\boldsymbol{m})}} k_{(\boldsymbol{g}',\boldsymbol{m}'),(\boldsymbol{g},\boldsymbol{m})}\, c_{(\boldsymbol{g}',\boldsymbol{m}')} \\ &\quad+\sum_{(\boldsymbol{g}',\boldsymbol{m}')} c_{(\boldsymbol{g}',\boldsymbol{m}')}\sum_{(\boldsymbol{m}'')}\left(k_{(\boldsymbol{g}',\boldsymbol{m}'),(\boldsymbol{g},\boldsymbol{m}''),(\boldsymbol{m})}c_{(\boldsymbol{g},\boldsymbol{m}'')}-k_{(\boldsymbol{g}',\boldsymbol{m}'),(\boldsymbol{g},\boldsymbol{m}),(\boldsymbol{m}'')}c_{(\boldsymbol{g},\boldsymbol{m})}\right)\end{aligned} \qquad (17)$$

The upper row of right-hand side terms corresponds to replication and extinction processes, whereas the bottom row is assigned to interaction processes. We emphasize that the upper row is formally identical to Eigen's replicator equations [14, 15, 21, 23] which phenomenologically describe a Darwinian evolution on molecular level. Rate constants $k_{(g,m),(g',m')}$ and $k_{(g',m'),(g,m''),(m)}$ correspond to replication process (2) and interaction process (8), respectively; loosely speaking, these rate constants may be to some extent considered as probabilities (7) and (11).

Let us postulate the following two conditions for rate constants:

(1) Diagonal replication rate constants are much greater than nondiagonal ones

$$0 < k_{(g,m),(g',m')} \ll k_{(g,m),(g,m)} \qquad \forall \left(g,m\right) \neq \left(g',m'\right) \tag{18a}$$

(2) Interaction rate constants are much smaller than diagonal replication constants

$$0 < k_{(g,m),(g',m'),(m'')} \ll k_{(g,m),(g,m)} \qquad \forall \left(g,m\right),\left(g',m'\right),\left(m''\right) \tag{18b}$$

Both these conditions considerably simplify discussion of properties of differential equations (17); in particular, the interaction terms and replications with mutations may be considered as small perturbations. The first condition (18a) reflects an intrinsic assumption of any evolutionary calculation that mutation events are very rare. The second condition (18b) is realized in Algorithm 1 by introduction of the probability tuned such that $0 < P_{interact} \ll 1$. This means that almost all elementary transformations are performed in a "replication mode", whereas only very small fraction is performed in an "interaction mode".

A dilution flux ϕ from (17) is determined by the condition that a sum of all concentrations of m-genes is unitary, $\sum c_{(g,m)} = 1 \Rightarrow \sum \dot{c}_{(g,m)} = 0$. We get

$$\begin{aligned} \phi = & \underbrace{\sum_{(g,m)} k_{(g,m),(g,m)} c_{(g,m)}}_{dominant\ part} + \sum_{(g',m') \neq (g,m)} k_{(g',m'),(g,m)}\, c_{(g',m')} \\ & + \sum_{(g',m')} c_{(g',m')} \sum_{(g,m),(m'')} \left(k_{(g',m'),(g,m''),(m)} c_{(g,m'')} - k_{(g',m'),(g,m),(m'')} c_{(g,m)} \right) \end{aligned} \tag{19}$$

According to the above two conditions (18a-b) the dilution flux is mainly determined by the dominant part specified entirely by diagonal replication constants, other two remaining terms are negligible with respect to the dominant term.

Applying both conditions (18a-b) we get that the differential equations are dominantly determined by the first two terms in (17), i.e. the dynamics of m-gene concentrations is very similar to the dynamics of Eigen's replicators [14, 15]. Moreover, according to the first condition (18a), the mutations in the course of replications are very rare events, then during whole evolution the population is composed predominantly of one sort of a particular m-gene (of course this statement is not true for transient stages of evolution when a dominant m-gene is substituted by another more fitter m-gene. Summarizing, both the above conditions (18a-b) considerably simplify discussion of properties of differential equations. The interaction terms and replications with mutations may be considered as small perturbations and therefore we may expect that an evolutionary dynamics of m-genes is very similar to the dynamic of Eigen's replicators [14, 15, 23].

CHEMOSTAT SIMULATION OF COEVOLUTION OF GENES AND MEMES

The chemostat approach outlined in the previous Section of this communication (see Algorithm 1) will be used as an algorithmic framework for a simulation of Darwinian coevolution between genes and memes. An initial composition of chemostat is set to identical binary strings. Numerical values of single parameters are specified in Table 1.

Table 1. Values of parameters used in numerical simulations

No.	Parameter	Value
1	N, length of genes and memes	40
2	$P_{\mathrm{mut}}^{(\mathrm{gene})}$, probability of one-point gene mutation $P_{\mathrm{mut}}^{(\mathrm{meme})}$, probability of one-point meme mutation	0.0001 0.001
3	P_{interact}, probability of an interaction event (see Algorithm 1)	0.5
4	α, slope parameter for calculation of probabilities (7) and (11)	4
5	ω, "cultural" parameter in (2)	0.1
6	A, size of population	500
7	$t_{\max}$, maximal number of elementary evolutionary epochs	2×10^6
8	U_{card}, size of neighborhood in (5b) and (9)	1 (10)
9	$U_{\mathrm{card}}^{(\mathrm{eff})}$, size of neighborhood in (13), Baldwin effect	1 (10)

We specify functions F and H that are needed for fitness evaluation of m-genes (see Eq. (2)) by Kauffman's rugged functions [1, 24]. Their general properties are summarized by Altenberg [1]; for completeness, see Appendix A. In particular the function F is specified as a standard NK function, where $N = 40$ and $K = 3$ (number of pleotropisms). Slightly more complicated is a specification of the function H that expresses an interaction between genes and memes. We used the so-called generalized function *NKCS* [24] (we put C=S=1), which Kauffman introduced as a proper model to allow the systematic study of various aspects of natural evolution between interacting species. A plot of typical results is displayed in Fig. 6 (for parameters specified in Table 1). This figure nicely demonstrates the basic property of all our simulations, in particular that the plot of mean value of function F (it corresponds to a fitness of an isolated gene) is nondecreasing and is of a "staircase" form, whereas the plot of mean value of H (it corresponds, loosely speaking, to a fitness of meme with respect to its "partner" gene), is not monotonic and mostly random. This very important conclusion may be formulated as the first observation concluded from our simulations.

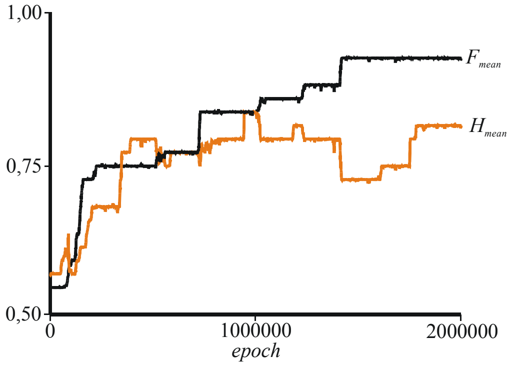


Fig. 6.
Typical plots of mean values of functions F and H obtained for parameters specified in Table 1. We see that a plot for F (corresponding to a fitness of an isolated gene) forms a typical nondecreasing "staircase" graphs, whereas a plot of H (corresponding to an interaction between genes and memes) is nonmonotonic and mostly random.

***1st observation**. A gene part F of fitness forms **always** a "staircase" nondecreasing function, whereas a memetic part H of fitness is not monotonic; it may contain both decreasing as well as increasing stages.*

This first observation has a very interesting, almost "philosophical" interpretation. Biological evolution is always of a cumulative and nondecreasing character, but its memetic counterpart though partially of an increasing character contains also a considerable random part (it is not of a cumulative and nondecreasing character). This interesting feature of our simulations is caused by the fact that memes form only an "environment" (we may say a cultural niche) for an evolution of genes, while memes are "evolution-pulled" only indirectly through their interaction with genes.

The fitness (2) contains a positive parameter ω with the help of which an interaction (cultural) term H may be gradually incorporated into fitness. We have done a series of simulations with parameters specified in Table 1, when the parameter ω was gradually changed from the initial value 0.1 to its final value 2.0. The obtained results may be summarized by the following observation.

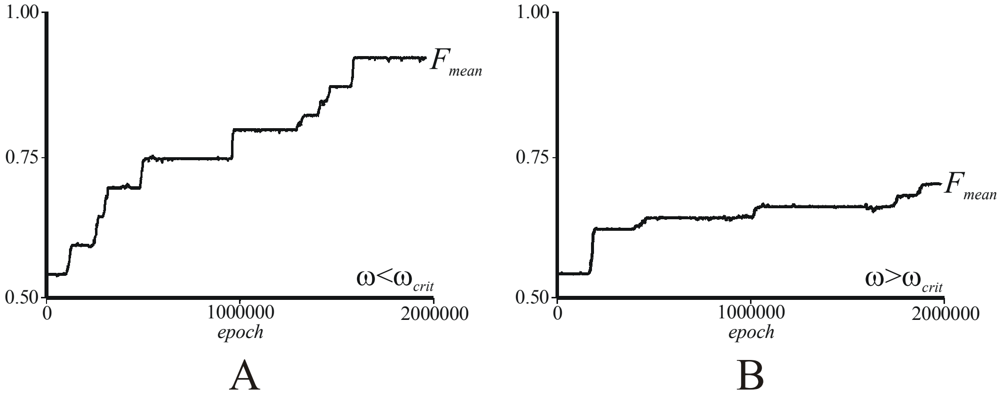


Fig. 7.
Two different plots of mean values of function F, which corresponds to a gene part of the fitness. Diagram A contains a plot for those values of parameter ω that are smaller than a critical value ω_{crit}. This plot shows that gene fitness gradually increases almost to its maximal unit value. Diagram B contains a similar plot, but for such values of ω that are greater than the critical value ω_{crit}. In this case a gradual increase in function F is substantially decelerated.

***2nd observation**. There exists a critical value ω_{crit} of cultural parameter (for a fixed value of interaction probability P_{interact}); for $\omega<\omega_{\text{crit}}$ a gene part F of fitness is not much affected, but for $\omega>\omega_{crit}$ part F is substantially decelerated* (see Fig. 7).

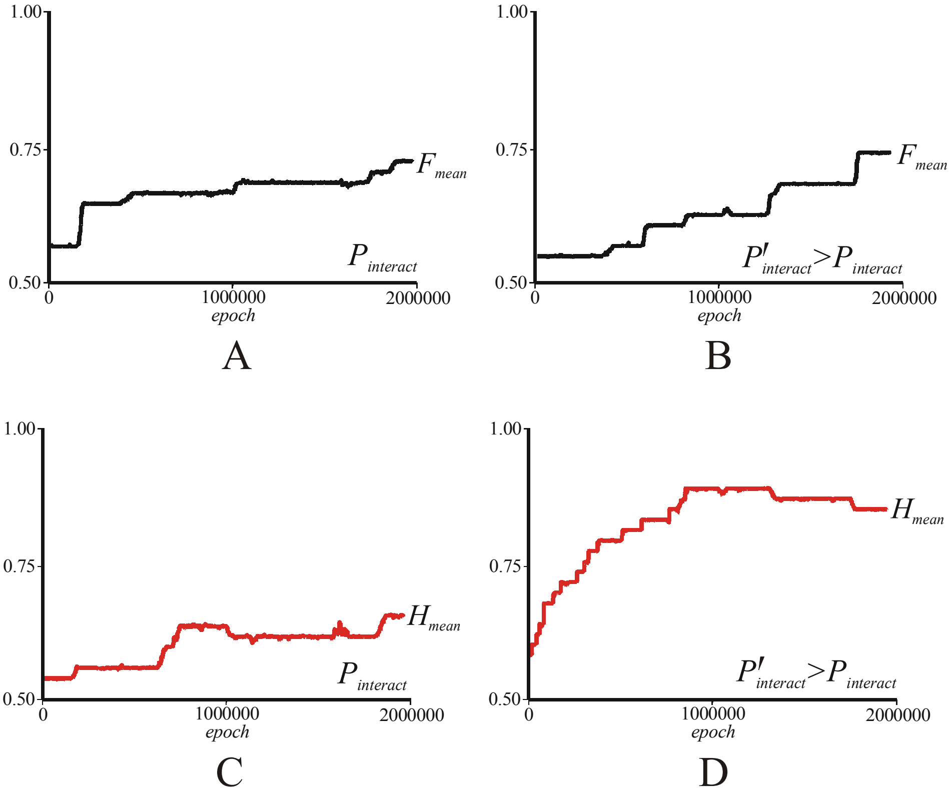


Fig. 8.
Plots of mean values of functional parts *F* and *H* of the fitness for two different values of probabilities $P_{interact}$ and $P'_{interact} > P_{interact}$. Diagrams A and C correspond to $P_{interact} = 0.5$, whereas diagrams B and D correspond to $P'_{interact} = 0.95$. We see that an increase in the probability (diagrams A and B) does not affect the plots of the gene fitness part *F*; on the other hand, an increased value of probability substantially changes (diagrams C and D) plots of interaction fitness part *H*, which is considerably accelerated to higher functional values.

A similar phenomenon was also observed by Bull et al. [7]. They interpreted the observation as a result of the existence of a “phase transition” depending on the parameter ω. There exist two different types of plots such as displayed in Fig. 7. If parameter ω exceeds its critical value, then one type of plot is dramatically changed to the other type.

The chemostat algorithm (see Algorithm 1) contains an important parameter corresponding to probability P_{interact} which controls stochastically an inclusion of interaction between two randomly selected m-genes. We have done two independent simulations, the first one for smaller values of probability P_{interact} and the second one of its increased value P'_{interact} ; the obtained results are summarized by the following observation.

***3rd observation**. An interaction probability P_{interact} does not affect substantially gene part F of fitness while a cultural part H is accelerated by increasing the probability (see Fig. 8).*

As was already mentioned, the probability P_{interact} controls a horizontal transfer of memes, its intensification by increasing the probability caused an increase in selection pressure on memes, which manifests itself by acceleration of memetic evolution. In other words, this means that the frequency of horizontal memetic transfer ranks among important parameters of the present coevolutionary theory; it may partially control relative evolutionary rate of memes with respect to genes.

For vertical as well as horizontal transfers of memes, the so-called social learning was introduced, which consists in local optimization of received memes, in the course of application of either a replication process or an interaction process. Loosely speaking, we may say that a respective gene does not automatically accept a new meme (from its parent and donor, respectively), but it tries to modify (locally optimize) this meme into a form more suitable for the gene than its original received form. The process of this search for more appropriate form of a received gene is called the learning, and since this learning is performed within the act of transfer of a new meme from one gene to another, this type of learning is called the social learning. The obtained results are summarized as the next observation.

4th observation. *An introduction of meme learning in replication process, in particular for smaller values of cultural parameter* ω, *causes an acceleration of meme evolution by exerting stronger selection pressure on them* (see Fig. 9).

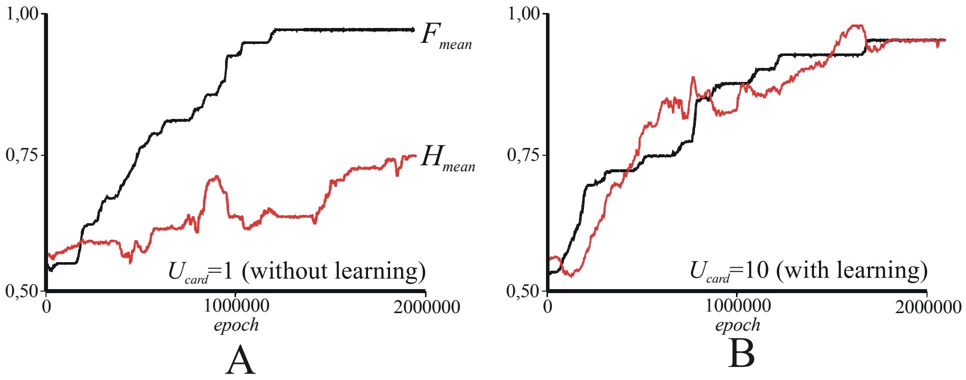


Fig. 9.
An illustrative plots of the influence of social learning on coevolution of genes and memes. We see that the evolution of genes is not affected by an introduction of social learning, it affects considerably only memes.

In the first part of this paper (following Eq. (5b)), we have discussed the so-called gene learning, which may be applied when m-genes are evaluated by fitness. The approach of gene learning was formulated via the concept of an effective fitness, which is defined as a local maximum of the standard fitness (2) in the nearest neighborhood of the respective gene. The effective fitness manifests a potential evolvability[4] of the given m-gene, i.e. m-genes with a greater effective fitness than their original (noneffective) counterparts are more effective in the forthcoming stage of evolution. Theoretical approaches based on application of evolvability to genes, which code cognitive organs with a necessary plasticity, are in literature called the Baldwin effect [3, 4, 27]. In our simulations the Baldwin effect has been included by replacement of a standard fitness by an effective fitness (see Fig. 10).

5th observation. *An introduction of gene learning within evaluation of m-genes by fitness (the so-called Baldwin effect) causes a substantial acceleration of gene evolution and it pulls vicariously also a meme evolution (see Fig. 10).*

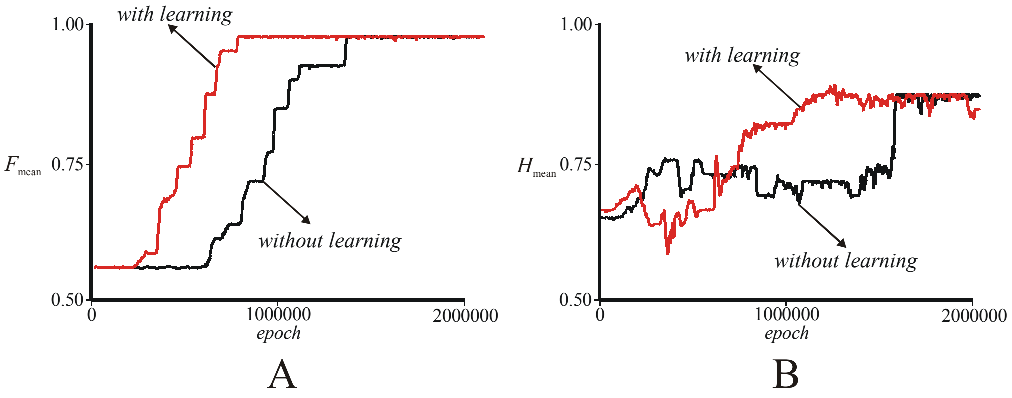


Fig. 10.
Illustrative plots for a demonstration of the Baldwin effect. Diagram A shows plots of mean values of gene fitness F for situations when the gene learning is included or not included. We see than an inclusion of gene learning substantially accelerates the gene evolution. Diagram B shows plots for mean values of interaction parts of fitness H, the forms of these plots being not very different. Loosely speaking, meme evolution is pulled by genes to higher values.

CONCLUSIONS

Although scientific value of memetics still remains a matter of opinion in social sciences [2], in computer sciences [9, 26, 27, 28, 29] (particularly in artificial societies and social simulations) it ranks among very popular approaches how to overcome information limits of Darwinian evolution. Since there is impossible to realize genetically a transfer of acquired information from parents onto offspring, but some of this information is vitally important for successful surviving of offspring, there must exist other than genetic transfer of information. It seems that the most natural solution to these problems is a meme approach [2, 5, 10, 11], based on the postulation that genes and memes form composites called in this paper the m-genes. The present coevolutionary theory is fully based on the postulation that memes are not appearing independently and freely with respect to genes, but they are obligatorily bound to genes in composites, which are subjects of Darwinian evolution. This represents a formal attempt to model a coevolution of genes and memes by introducing a population of m-memes with specified types of elementary interactions. Loosely speaking, this approach may be alternatively understood as a performance of multiagent system, where each agent is specified not only by a gene (e.g., it specifies architecture of its cognitive organ), but also by a meme (a knowledge database that facilitates surviving of agent in its environment).

Acknowledgement. The work was supported by the Scientific Grant Agency of Slovak Republic under grant 1/4053/07 and by the Slovak Research and Development Agency under grant 20-002504.

REFERENCES

1. Altenberg L. in: *Handbook of Evolutionary Computation* (T. Baeck, D. F. Fogel, Z. Michalewicz, Eds), p. B2.7.2. Oxford University Press, New York 1998.
2. Aunger R. (Ed.): *Darwinizing Culture: The Status of Memetics as a Science*. Oxford University Press, New York 2001.

3. Baldwin J. M.: *Am. Nat.* **1896,** *30,* 441.
4. Belew R. K., Mitchell M. (Eds): *Adaptive Individuals in Evolving Populations: Models and Algorithms*. Addison-Wesley, Reading (MA) 1996.
5. Blackmore S.: *The Meme Machine*. Oxford University Press, New York 1999.
6. Boyd R., Richerson P.J.: *Culture and the Evolutionary Process.* University of Chicago Press, Chicago 1985.
7. Bull L., Holland O., Blackmore S.: *Artif. Life* **2000,** *6***,** 227.
8. Cavalli-Sforza L. L., Feldman M. W.: *Cultural Transmission and Evolution: A Quantitative Approach***.** Princeton University Press, Princenton 1981.
9. Conte R., Hegselmann R., Terna P. (Eds): *Simulating Social Phenomena.* Springer, Berlin 1997.
10. Dawkins R.: *The Selfish Gene*. Oxford University Press, Oxford 1976.
11. Dennett D.: *Darwin's Dangerous Idea*. Penguin Press, Hammondsworth 1995.
12. Dittrich P, Ziegler J., Banzhaf, W.: *Artif. Life* **2001**, *7*, 225.
13. Durham W.H.: *Coevolution: Genes, Culture and Human Diversity.* Stanford University Press, Stanford 1991.
14. Eigen M., Schuster P.: *Naturwissenschaften* **1977**, *64*, 541; **1977**, *65*, 7; **1977**, *65***,** 341.
15. Eigen M.: *Naturwissenschaften* **1971,** *58*, 465.
16. Epstein J.M., Axtell R. L.: *Growing Artificial Societies: Social Science From the Bottom Up*. MIT Press, Cambridge, MA 1996.
17. Feldman M. W., Laland K. N.: *Trends in Ecology and Evolution* **1996,** *11*, 453.
18. Feldman M. W., Zhivotovsky L. A.: *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11935.
19. Gillespie D. T.: *J. Phys. Chem.*, **1977**, *81*, 2340.
20. Hillis W. D. in: *Artificial Life II* (C. G. Langton, C. Taylor, J. D. Farmer, S. Rasmussen, Eds), p. 313. Addison-Wesley, Redwood City (CA) 1992.
21. Hofbauer J., Sigmund K.: *The Theory of Evolution and Dynamical Systems*. Cambridge University Press, Cambridge 1988.
22. Holland J. H.: *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor 1975.
23. Jones B. L., Enns R. H., Rangnekar S. S.: *Bull. Math. Biol.* **1976**, *38*, 15.
24. Kauffman S.: *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford University Press, Oxford 1993.
25. Kvasnička V., Pospíchal J. in: *Advances in Soft Computing* (R. Roy, T. Furuhashi, and P. K. Chawdhry, Eds), p 481. Springer, London 1999.
26. Kvasnička V., Pospíchal J.: *Adv. Complex Syst.* **1999**, *2*, 395.
27. Kvasnička V., Pospíchal J.: *Artif. Life* **2000**, *5*, 319.
28. Kvasnička V., Pospíchal J.: *Collect. Czech. Chem. Commun.* **2003**, *68*, 139.
29. Kvasnička V., Pospíchal J.: *Evol. Cognition* **1999**, *5*, 75.
30. Laland K. N., Odling-Smee J., Feldman M. W.: *Behav. Brain Sci.* **1999**, *23*, 131.
31. Paredis J. in: *The Handbook of Evolutionary Computation* (T. Baeck, D. F. Fogel, Z. Michalewicz, Eds), 1st Suppl. Oxford University Press, New York 1998.
32. Wright S.: *Genetics* **1931**, *16*, 97.

APPENDIX A. KAUFFMAN'S *KN* FUNCTIONS

Let ***g*** be a binary string composed of *N* entries

$$\boldsymbol{g} = \left(g_1 g_2 ... g_N\right) \in \{0,1\}^N \tag{A1}$$

Each entry index $1 \le i \le N$ is evaluated by a subset composed of *K*+1 randomly selected indices (including *i*) from {1,2,...,*N*} (this subset is called the neighborhood) (see Fig. A1),

$$\Gamma(i) = \{ j_1 < j_2 < ... < j_{K+1} \} \subseteq \{1,2,...,N\} \tag{A2}$$

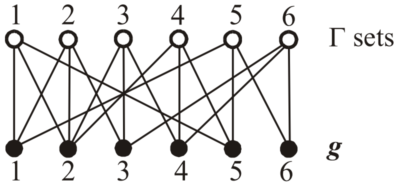


Fig. A1.
An illustrative example of Kauffman's rugged function specified for *N*=6 and *K*=2, where subsets Γs are specified as follows: $\Gamma(1)=\{1,2,5\}, \Gamma(2)=\{1,2,3\}$, $\Gamma(3)=\{2,3,4\}$ $\Gamma(4)=\{2,4,5\}, \Gamma(5)=\{1,5,6\}$, $\Gamma(6)=\{3,4,6\}$.

Kauffman's rugged function maps binary vectors of the length *N* onto positive real numbers from the interval [0,1]; this mapping is determined with respect to subsets Γ(*i*) as follows:

$$F(\boldsymbol{g}) = \frac{1}{N}\sum_{i=1}^{N} \varphi\left(\left\langle \boldsymbol{g}, \Gamma(i) \right\rangle\right) = \frac{1}{N}\sum_{i=1}^{N} \varphi\left(g_{j_1}, g_{j_2}, ..., g_{j_{K+1}}\right) \tag{A3}$$

where an auxiliary function φ is randomly specified by

$$\varphi\left(g_{j_1}, g_{j_2}, ..., g_{j_{K+1}}\right) = \text{random}\left(\underbrace{\text{int}\left(g_{j_1}, g_{j_2}, ..., g_{j_{K+1}}\right)}_{\text{RandSeed}}\right) \tag{A4}$$

where an integer $\text{int}\left(g_{j_1}, g_{j_2}, ..., g_{j_{K+1}}\right) = \sum_{l=0}^{K} g_{j_{K+1-l}} 2^l$ is used as a RandSeed of a particular random number generator with uniform distribution of positive reals from [0,1). For better understanding of the above ideas let us consider a Kauffman's function with Γ sets specified in Fig. A1. For instance, a string $\boldsymbol{g} = (001101)$ is evaluated by the generalized Kauffman's rugged function

$$\begin{aligned} F(\boldsymbol{g}) &= \frac{1}{6}\left(\varphi(000) + \varphi(001) + \varphi(011) + \varphi(010) + \varphi(001) + \varphi(111)\right) \\ &= \frac{1}{6}\left(\text{random}(0) + \text{random}(1) + \text{random}(3)\right. \\ &\quad \left. + \text{random}(2) + \text{random}(1) + \text{random}(7)\right) \end{aligned} \tag{A5}$$

where, e.g., random(3) corresponds to a uniform random number generator initiated by RandSeed=3 which produces a nonnegative real number from the semiopen interval [0,1).

Let us consider a meme ***m***, which is represented by a binary vector as follows

$$\boldsymbol{m} = \left(m_1, m_2, ..., m_N\right) \in \{0,1\}^N \tag{A6}$$

Both genes and memes form the so-called gene-meme complexes (m-gene) (***g***,***m***). Each m-gene index $i \in \{1,2,...,N\}$ is evaluated by a subset composed of K+1 gene indices and K+1 meme indices) (see Fig. A2),

$$\Pi(i) = \{j_1 < j_2 < ... < j_{K+1}, k_1 < k_2 < ... < k_{K+1}\} \tag{A7}$$

Generalized *NK* function assigned to an m-gene (***g, m***) is

$$H(\boldsymbol{g},\boldsymbol{m}) = \frac{1}{N}\sum_{i=1}^{N}\sigma\left(\left\langle \boldsymbol{g},\boldsymbol{m},\Pi(i)\right\rangle\right) \tag{A8a}$$

where

$$\left\langle \boldsymbol{g},\boldsymbol{m},\Pi(i)\right\rangle = \left(g_{j_1},...,g_{j_{K+1}} m_{k_1},...,m_{k_{K+1}}\right) \tag{A8b}$$

is composed of K+1 entries of ***g*** and K+1 entries of ***m***, and σ(...) is a random real from [0,1) assigned to a RandSeed specified by an integer assigned to a binary substring $\left(x_{j_1},...,x_{j_{K+1}} y_{k_1},...,y_{k_{K+1}}\right)$

$$\sigma\left(\underbrace{x_{j_1},...,x_{j_{K+1}} y_{k_1},...,y_{k_{K+1}}}_{\text{RandSeed}}\right) = \text{random}\left(\text{RandSeed}\right) \tag{A8c}$$

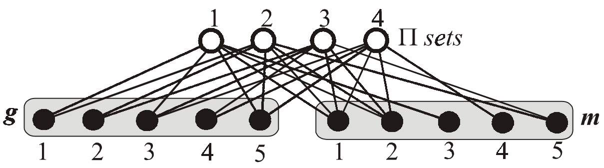


Fig. A2.
An illustrative example of Kauffman's function for an m-gene (specified for *N*=6, *K*=4), where subsets Πs are specified as follows: $\Pi(1) = \{1,3,5,1,2,3\}$, $\Pi(2) = \{1,2,5,1,2,5\}$, $\Pi(3) = \{1,2,4,1,2,5\}$, $\Gamma(4) = \{3,4,5,1,2,4\}$.

An m-gene $((\boldsymbol{g},\boldsymbol{m}) = (10110\,11000)$ is evaluated by the generalized Kauffman's function with Π sets specified by Fig. A2

$$\begin{aligned} H(\boldsymbol{g},\boldsymbol{m}) &= \frac{1}{4}\Big[\sigma\left(\left\langle \boldsymbol{g},\boldsymbol{m},\Pi(1)\right\rangle\right) + \sigma\left(\left\langle \boldsymbol{g},\boldsymbol{m},\Pi(2)\right\rangle\right) + \\ &\qquad \sigma\left(\left\langle \boldsymbol{g},\boldsymbol{m},\Pi(3)\right\rangle\right) + \sigma\left(\left\langle \boldsymbol{g},\boldsymbol{m},\Pi(4)\right\rangle\right)\Big] \\ &= \frac{1}{4}\Big[\sigma(110110) + \sigma(100110) + \sigma(011110) + \sigma(011110)\Big] \\ &= \frac{1}{4}\Big[\text{random}(54) + \text{random}(38) + \text{random}(30) + \text{random}(30)\Big] \end{aligned} \tag{A9}$$

APPENDIX B. GENERAL CLASSIFICATION OF COEVOLUTION

Hillis [18] used a *coevolutionary competitive model* as early as 1992, for construction of networks for sorting lists of sixteen numbers. Such a sorting network consists of comparisons between two numbers (and the possibility of swapping both numbers). The goal is to find a network which correctly sorts all possible lists of sixteen numbers and this with as few comparisons as possible. Hillis used two populations. The first population consisted of sorting networks. The individuals in the second population were sets of test-lists. These lists contain numbers to be sorted. In each generation a sorting network was tested on the set of test-lists at the same location. The fitness of a sorting network was defined as the percentage of correctly sorted test-lists. The fitness of the set of test-lists, on the other hand, was equal to the percentage of test-lists incorrectly sorted by the network.

The purpose of this second Appendix is to outline a theory of two different types of coevolution, in particular, we will distinguish the so-called *cooperative* and *competitive* coevolution. Let us consider a mapping $y = f(\boldsymbol{x})$ with binary argument

$$f : D \subseteq \{0,1\}^n \rightarrow R \tag{B1}$$

which evaluates binary vectors $\boldsymbol{x} \in \{0,1\}^n$ by a real number $f(\boldsymbol{x}) \in R$. Our goal is to find its global maximum

$$\boldsymbol{x}_{opt} = \arg \max_{\boldsymbol{x} \in \{0,1\}^n} f(\boldsymbol{x}) \tag{B2}$$

This discrete optimization problem is obviously NP hard, the CPU time grows exponentially with the length of binary vectors $\boldsymbol{x}$, $t_{\mathrm{CPU}} \sim 2^n$.

B1. Competitive Darwinian coevolution

Let us postulate that the vector $\boldsymbol{x}$ of arguments may decompose into two subvectors that represent different coevolved subsystems P and Q

$$\boldsymbol{x} = \boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}} \tag{B3}$$

$$y = f(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}) \tag{B4}$$

where $|\boldsymbol{x}_{\mathrm{P}}| = p$ a $|\boldsymbol{x}_{\mathrm{Q}}| = q$. An optimal solution of (B2) may be written as follows

$$\boldsymbol{x}_{\mathrm{opt}} = \boldsymbol{x}_{\mathrm{P,opt}} \oplus \boldsymbol{x}_{\mathrm{Q,opt}} \tag{B5}$$

$$\boldsymbol{x}_{\mathrm{P,opt}} = \arg \max_{\boldsymbol{x}_{\mathrm{P}} \in \{0,1\}^p} f(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q,opt}}) \tag{B6}$$

$$\boldsymbol{x}_{\mathrm{Q,opt}} = \arg \max_{\boldsymbol{x}_{\mathrm{Q}} \in \{0,1\}^q} f(\boldsymbol{x}_{\mathrm{P,opt}} \oplus \boldsymbol{x}_{\mathrm{Q}}) \tag{B7}$$

For each subsystem P and Q we introduce populations $\mathcal{P}$ and $\mathcal{Q}$, respectively, composed of potential solutions (replicators) of suboptimization tasks (B6-7)

$$\mathcal{P} = \{\boldsymbol{x}_{\mathrm{P},1}, \boldsymbol{x}_{\mathrm{P},2}, \ldots, \boldsymbol{x}_{\mathrm{P},A}\} \text{ and } \mathcal{Q} = \{\boldsymbol{x}_{\mathrm{Q},1}, \boldsymbol{x}_{\mathrm{Q},2}, \ldots, \boldsymbol{x}_{\mathrm{Q},B}\} \tag{B8}$$

Fitness of replicators is specified as follows

$$f_{\mathrm{P}}(\boldsymbol{x}_{\mathrm{P}}) = \max_{\boldsymbol{x}_{\mathrm{Q}} \in \mathcal{Q}} f(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}) \tag{B9}$$

$$f_{\mathrm{Q}}(\boldsymbol{x}_{\mathrm{Q}}) = \max_{\boldsymbol{x}_{\mathrm{P}} \in \mathcal{P}} f(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}) \tag{B10}$$

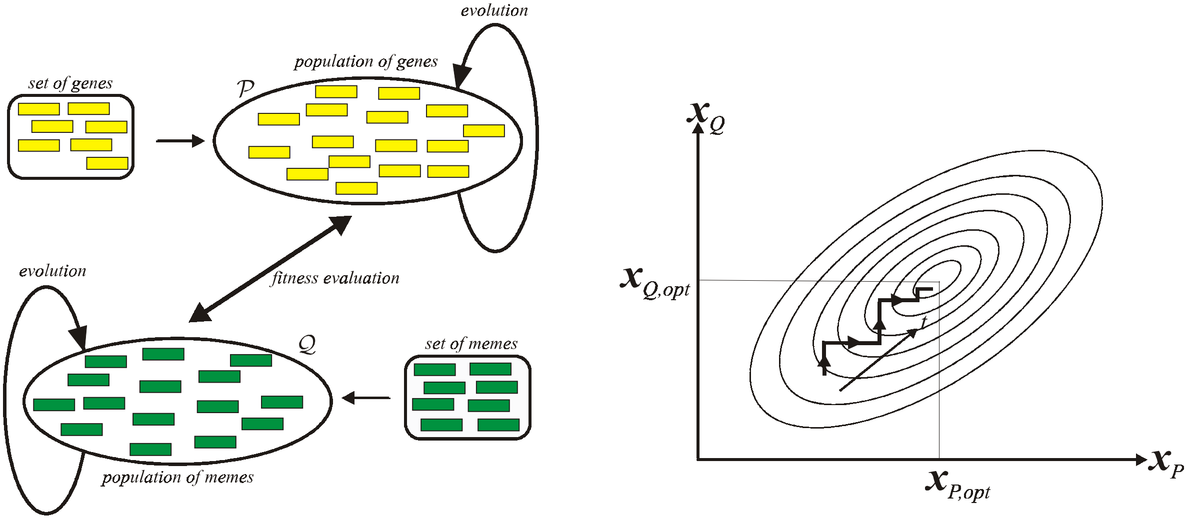
This means that in the case of competitive coevolution both populations $\mathcal{P}$ and $\mathcal{Q}$ interact through evaluation of fitness (e.g. if we calculate the fitness of replicators $\boldsymbol{x}_{\mathrm{P}} \in \mathcal{P}$, then it is specified by (B9) as a result of its interaction with all replicators from another population $\mathcal{Q}$). Loosely speaking, the fitness of an individual – replicator is calculated through competition "duels" with individuals from other chemostats. All that is required is to know which individual is better, no further quantification is needed.

An algorithm for the competitive evolution is specified as follows (see Fig. B1):

Algorithm B1.

```
chemostats U($\mathcal{P}$) and U($\mathcal{Q}$) are randomly generated
and all replicators are evaluated by fitness;
for t:=1 to $t_{max}$ do
begin $\mathbf{x}_P := O_{select}(\mathcal{P})$;
        if random<prob(fitness($\mathbf{x}_P$)) then
        begin $\mathbf{x}_P' := O_{mut}(x_P)$;
                $\mathbf{x}_P'$ is evaluated by fitness;
                $\mathbf{x}_P'' := O_{select}(\mathcal{P})$;
                $\mathcal{P} := (\mathcal{P} - \{\mathbf{x}_P''\}) + \{\mathbf{x}_P'\}$;
        end;
        $\mathbf{x}_Q := O_{select}(\mathcal{Q})$;
        if random<prob(fitness($\mathbf{x}_Q$)) then
        begin $\mathbf{x}_Q' := O_{mut}(\mathbf{x}_Q)$;
                $\mathbf{x}_Q'$ is evaluated by fitness;
                $\mathbf{x}_Q'' := O_{select}(\mathcal{Q})$;
                $\mathcal{Q} := (\mathcal{Q} - \{\mathbf{x}_Q''\}) + \{\mathbf{x}_Q'\}$;
        end;
end;
```

The algorithm is initialized by random generation of both chemostats $\mathcal{P}$ and $\mathcal{Q}$. An outer time cycle `t` is used for a specification of total evolutionary steps in both chemostats.



A B

Fig. B1.
(A) Schematic visualization of a competitive coevolution of two populations - chemostats $\mathcal{P}$ and $\mathcal{Q}$, where both chemostats evolve quasi-independently; they interact through fitness evaluation. (B) An optimization process in the competitive coevolution has stair-case form.

B2. Cooperative Darwinian coevolution

In a similar way as in the previous competitive coevolution, the function $f\left(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}\right)$ may be alternatively written as

$$f\left(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}\right)=\underbrace{f\left(\boldsymbol{x}_{\mathrm{P}} \oplus \tilde{\boldsymbol{x}}_{\mathrm{Q}}\right)}_{F\left(\boldsymbol{x}_{\mathrm{P}}\right)} + \underbrace{f\left(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}\right)-f\left(\boldsymbol{x}_{\mathrm{P}} \oplus \tilde{\boldsymbol{x}}_{\mathrm{Q}}\right)}_{H\left(\boldsymbol{x}_{\mathrm{P}}, \boldsymbol{x}_{\mathrm{Q}}\right)} \tag{B11}$$

where $\tilde{\boldsymbol{x}}_{\mathrm{Q}}$ is an arbitrary "state vector" from the second subproblem Q (usually is taken as locally/temporarily best solution for Q subsystem). This means that function value $f\left(\boldsymbol{x}_{\mathrm{P}} \oplus \boldsymbol{x}_{\mathrm{Q}}\right)$ is decomposed into two parts: (1) the first part specified by a function F that evaluates a fitness of $\boldsymbol{x}_{\mathrm{P}}$ and (2) the second part specified by a function H that evaluates an interaction fitness between $\boldsymbol{x}_{\mathrm{P}}$ and $\boldsymbol{x}_{\mathrm{Q}}$. The optimization problem (B2) looks as follows

$$\left(\boldsymbol{x}_{\mathrm{P,opt}}, \boldsymbol{x}_{\mathrm{Q,opt}}\right)=\arg \max_{\left(\boldsymbol{x}_{\mathrm{P}}, \boldsymbol{x}_{\mathrm{Q}}\right)}\left[F\left(\boldsymbol{x}_{\mathrm{P}}\right)+H\left(\boldsymbol{x}_{\mathrm{P}}, \boldsymbol{x}_{\mathrm{Q}}\right)\right] \tag{B12}$$

Summarizing, a chemostat $\mathcal{P}$ is a multiset composed of ordered pairs $\boldsymbol{x}=\left(\boldsymbol{x}_{\mathrm{P}}, \boldsymbol{x}_{\mathrm{Q}}\right)$; their fitness is specified by

$$\operatorname{fitness}(\boldsymbol{x})=F\left(\boldsymbol{x}_{\mathrm{P}}\right)+H\left(\boldsymbol{x}_{\mathrm{P}}, \boldsymbol{x}_{\mathrm{Q}}\right) \tag{B13}$$

The cooperative type of coevolution is usually typical of an interaction of different replicators that are in a common symbiotic relation. Then these different replicators may be formally considered as one super-replicator composed of quasi-isolated part corresponding to original replicators.

An algorithm for the cooperative evolution is specified by the following algorithm (see also Fig. B2).

Algorithm B2.

```
chemostat $\mathcal{P}$ is randomly generated and all its
replicators are evaluated by fitness;
for t:=1 to $t_{max}$ do
begin x:=$O_{select}(\mathcal{P})$;
      if random<prob(fitness(x)) then
      begin $x_P$':=$O_{mut}(x_P)$;
            $x_Q$':=$O_{mut}(x_Q)$;
            x':=($x_P$', $x_Q$');
            x' is evaluated by fitness;
            x'':=$O_{select}(\mathcal{P})$;
            $\mathcal{P}$:=($\mathcal{P}$-{x''})+{x'};
      end;
end;
```

This pseudocode may serve as an algorithmic background for the construction of Algorithm 1, which corresponds to a cooperative coevolution of gene-meme pairs.

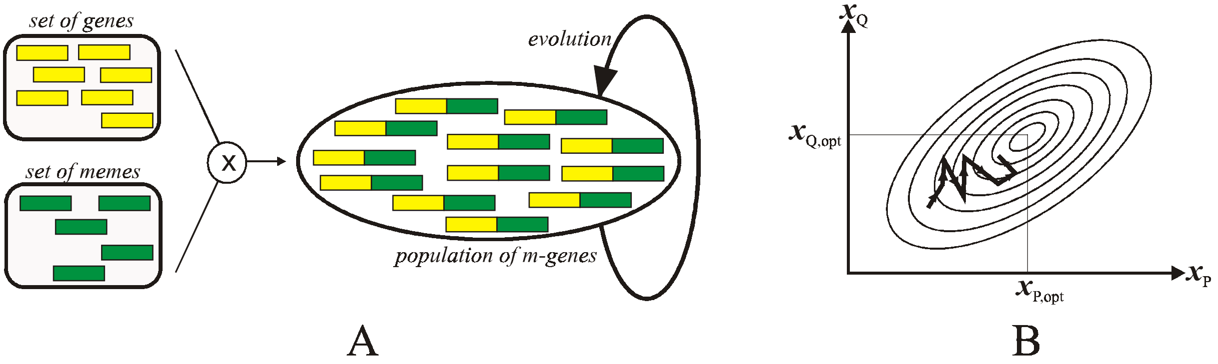


Fig. B2.
(A) Schematic visualization of a cooperative coevolution with one population - chemostat $\mathcal{P}$ composed of ordered pairs $\boldsymbol{x}=(\boldsymbol{x}_{\mathrm{P}},\boldsymbol{x}_{\mathrm{Q}})$. (B) An optimization process in the cooperative coevolution looks like a stochastic line globally directed to the maximum.