PERGAMON

Chaos, Solitons and Fractals 14 (2002) 745–758

CHAOS
SOLITONS & FRACTALS

www.elsevier.com/locate/chaos

Entanglement-based self-organization

Michail Zak [a,b,*]

[a] *Ultracomputing group, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, M/S 126-347, Pasadena, CA 91109-8099, USA*

[b] *Center for Engineering Science Advanced Research, Oak Ridge National Laboratory, Tennessee, USA*

Accepted 20 December 2001

Abstract

Based upon quantum entanglement, several paradigms of self-organization (such as inverse diffusion, transmissions of conditional information, decentralized coordination, cooperative computing, competitive games, topological evolution in active systems) are introduced and discussed. © 2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

This paper was motivated by recent discovery and experimental verification of the most fundamental and still mysterious phenomenon in quantum mechanics: quantum entanglement. Formally, quantum entanglement as well as associated with it quantum non-locality follows from the Schrödinger equation; however, its physical meaning is still under extensive discussions.

The most attractive aspect of quantum entanglement, in terms of a new quantum technology, is associated with instantaneous transmission of messages on remote distances. However, practical applications of this effect are restricted by the postulate adopted by many authors [5] that these messages cannot deliver any intentional information. That is why all the entanglement-based communication algorithms must include a classical channel. The main challenge of this study is to evaluate the degree of usefulness of entanglement-based communication technology without any classical channels. The first attempt of this kind was presented in our earlier work [4] where it has been demonstrated how a randomly chosen message can deliver non-intentional, but useful, information under special conditions which include a preliminary agreement between the sender and the receiver. In this paper we are trying to extend this effort by applying the entanglement-based correlations to an active system represented by a collection of intelligent agents. The problem of behavior of intelligent agents correlated by identical random messages in a decentralized way has its own significance: it simulates evolutionary behavior of biological and social systems correlated only via simultaneous sensoring sequences of unexpected events. As shown in this paper that under the condition that the agents have certain preliminary knowledge about each other, the whole system can exhibit emergent phenomena such as topological self-organization, inverse diffusion; it also can perform transmission of conditional information, decentralized coordination, cooperative computing, competitive games.

It is always a temptation to simulate any new quantum phenomenon by classical tools. In the case of quantum entanglement such a possibility was excluded from the very beginning since this is a non-local phenomenon which does not have any classical equivalents. However, one can argue that actually the active system under consideration becomes classical as soon as the quantum message is received and interpreted by the agent; therefore, instead of quantum implementation of the correlations between the agents, one can generate a pool of samples of stochastic processes in advance, make copies and distribute them over the agents, so that any two agents to be correlated would have identical

* Tel.: +1-818-393-5351; fax: +1-818-393-5471.
E-mail address: michail.zak@jpl.nasa.gov (M. Zak).

0960-0779/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0960-0779(02)00021-8

records of "random" messages. However, there is a fundamental flaw in such an implementation: the whole scenario of the active system evolution is fully predetermined, and someone (for instance, those who generated, copied and distributed the messages) can know this scenario in advance. In principle, each agent also can find out his future messages since the knowledge about this future has already existed. The difference between the quantum and classical cases is similar to that between real-time and pre-recorded TV programs: in the first case, future is unpredictable, while in the second case "future" has already happened, although the viewer may not know about that. In a more practical sense, the difference between the quantum and classical implementations becomes important when the communications between the agents are supposed to be confidential: in the classical case, the confidential information, in principle, is available long before it is needed, and that makes such communications less secure.

2. Active systems

Cooperation behavior of active systems is a fast growing field [1,2] comprising methods of statistical mechanics, nonlinear dynamics and information processing. By an active system we will understand here a set of interacting agents capable to process information. In the simplest case, each agent is represented by an inertionless classical point particle moving in d dimensions under the combined influence of Newtonian and non-Newtonian forces:

$$\dot{x}_i^{(j)} = f_i^{(j)}(\{x^{(j)}\}) + g_{ii}^{(j)}\Gamma^{(j)}(t); \quad \{x^j\} = x_1^{(j)}, x_2^{(j)}, \ldots, x_d^j, \qquad i = 1, 2, \ldots, d; \quad j = 1, 2, \ldots, n, \tag{1}$$

where $x_i^{(j)}$ is the position vector, $f(\{x^{(i)}\})$ is the Newtonian force, $\Gamma^{(j)}(t)$ is the non-Newtonian force generated by the agent in response to received information, n is the total number of agents in the system, and $g_{ii}^{(j)} = g_{ii}^{(j)}(\{x^{(j)}\}, t)$.

The variety and the complexity of behavior of active systems depend upon the type of interactions between the agents. So far we have not introduced any interactions on the level of Newtonian forces: in Eq. (1), the force $f^{(j)}$ acting upon the jth agent does not depend upon the coordinates of other agents. Let us turn now to the non-Newtonian forces $\Gamma(t)$. Assuming that these forces result from uncorrelated information flows between the agents, the total effect can be presented in the form of the Langevin force, i.e., white noise:

$$\langle \Gamma^{(j)}(t) \rangle = 0, \quad \langle \Gamma^{(j)}(t)\Gamma^{(k)}(t') \rangle = 2\underset{jk}{\delta}\delta(t - t'). \tag{2}$$

In addition we will also assume that

$$|g_{ii}| \ll |f_i|, \tag{3}$$

i.e., the contribution of the non-Newtonian forces is much smaller than those of the Newtonian ones. Then each agent can predict its own future motion only probabilistically via the corresponding Fokker–Planck equation [3],

$$\frac{\partial \rho^{(j)}}{\partial t} = D^{(j)}\nabla^2\rho^{(j)} - \nabla \cdot (\rho^{(j)} f^{(j)}), \quad j = 1, 2, \ldots, n, \tag{4}$$

where $\rho^{(j)}\mathrm{d}^d x$ is the probability of finding the jth agent obeying Eq. (1) within a cubical volume $\mathrm{d}^d x$ centered at x at time t, and $D^{(j)}$ is the diffusion tensor with the diagonal elements $(g_{ii}^{(j)})^2$ and zero non-diagonal elements.

The degree of unpredictability is measured by the entropy:

$$I_j = \sum_{i=1}^{n} \log_2 \sqrt{\pi \mathrm{e} \left(g_{ii}^{(j)}\right)^2 t} = I(\{x^{(j)}\}) \tag{5}$$

if one starts in Eq. (1) with the sharp value of $\{x^{(j)}\}$ at $t = 0$. Hence, no matter how small the diffusion components g_{ii}^2, eventually the entropy or the information capacity becomes large.

Obviously, the information capacity of the system is:

$$I(\{x\}^{(1)}, \ldots, \{x\}^n) = \sum_{j=1}^{n} I_j. \tag{6}$$

Since all the agents are independent, the conditional entropy

$$I_{k|\ell} = I_\ell, \tag{7}$$

i.e., the knowledge of his own motion does not help an agent to predict motion of another agent.

In this paper, based upon the model of active systems represented by Eqs. (1) and (4) we are going to introduce a special correlation between the agents in the following way: each agent simultaneously receives a sequence of random

but identical binary signals of the type (+) or (−) and it converts each signal into a space step Δx to the right or the left, respectively, and these steps are performed during the corresponding time interval Δt.

Several properties of the proposed correlation should be emphasized.

First, there is no centralized source or a sender of the signal: each agent receives it by performing certain measurements synchronized with the measurements of the others. Thereby the signal gets uniformly and simultaneously distributed over the agents in a decentralized way.

Second, the signals transmit no intentional information that would favor one agent over another.

Third, all signals received by different agents are not only statistically equivalent, but also point-by-point identical.

Fourth, it is important to assume that each agent knows that all the other agents simultaneously receive the identical signals.

Finally, the sequences of the signals must be true random so that no agent could predict the next step with the probability different from 1/2.

It turns out that under these quite general assumptions, the active system can perform non-trivial tasks which include transmission of conditional information from one agent to another, simple paradigm of cooperation and self organization such as inverse diffusion, etc.

In order to justify the usefulness of the proposed correlation paradigm, consider an earthquake which is represented by some sequence of totally unpredictable jolts. All the "agents" (humans, animals) receive these unexpected signals simultaneously, and from that moment their activity became correlated and organized: they run to shelter, turn of pipelines, etc.

The second example is more sophisticated (but less idealized). Consider different social communities (or clubs) which receive information from the corresponding (but different) newspapers or television programs. The behaviors of the members of different clubs will be more independent. Now the following question can be asked: under what conditions would members of one club start migrating to other clubs? When does the balance between mixing (i.e., diffusion) and unmixing (i.e., inverse diffusion) occur? The answers to these questions can be started with the highly idealized paradigm of the proposed correlations.

The physical implementation of this paradigm was inspired by the phenomenon of entanglement [5] which is the most fundamental property of non-local quantum world. As will be shown below, all the five requirements listed above can be implemented only via quantum entanglement.

3. Quantum entanglement in active systems

Let us assume that agents A_1 and A_2 possess a set of N particles (say, electrons), which are in a one-to-one correspondence such that each pair is entangled; and suppose that the agent A_1 performs a sequence of measurements: one particle per unit time-step. Each measurement performed by the agent A_1 has two equally probable outcomes. In case of electrons, these outcomes can be spin-up (+) or spin-down (−). If (+) and (−) are converted by the agent into the movements along an axis to the right or to the left, respectively, the sequence of the agent's measurement can be interpreted as a symmetric unrestricted random walk. Hence, by performing these measurements, the agent A_1 selected (randomly) one trajectory out of 2^N equally probable trajectories of the corresponding random walk. Due to entanglement, the agent A_2 instantaneously receives this trajectory (after performing simultaneously the same type of measurements). This paradigm is easily generalizable to n entangled agents if each of them has a set of N particles entangled pairwise with the similar particles of all the other agents. The usefulness of such entanglement-based communications has been discussed in [4]. Here we will pay attention to other aspects of this type of correlations.

Firstly, one can verify that quantum entanglement satisfies all the five conditions formulated in Section 1 concerning decentralized correlations, namely: there is no single sender of the signals; the signals uniformly and simultaneously distributed over the agents; there is no intentional information transmitted; the sequences of the signals are true random. It should be emphasized that any attempt to simulate the same correlations classically would fail. Indeed, let us assume that a stochastic force $\Gamma(t)$ in Eq. (1) is generated classically. Then the probability that two different samples of this force are point-by-point identical is vanishingly small, and therefore, such a classical simulation of the correlations via quantum entanglement is unrealistic. But if one tries to generate one sample of $\Gamma(t)$ copy it and distribute these copies over the agents, then the randomness of the signals is destroyed: each signal becomes fully deterministic. In addition, the idea of decentralized correlations is lost: those who made the copies of the sample of $\Gamma(t)$ are in full control of the evolution of the active system (1).

Secondly, we will discuss the degree of cooperation between the agents due to quantum entanglement. For that purpose we will turn to the system (1). In order to trivialize the situation, we will consider only two one-dimensional

agents subjected to linear deterministic forces. In addition to that, for mathematical convenience, we will assume that the space steps Δx and the time-steps Δt of the random walk are small, i.e.,

$$\frac{\Delta x}{x_0} \ll 1, \quad \frac{\Delta t}{T_0} \ll 1, \tag{8}$$

where x_0 and T_0 are the characteristic lengths and time, respectively, and

$$\frac{(\Delta x)^2}{\Delta t} = 2. \tag{9}$$

This will allow one to preserve the continuous form of the Langevin equations in the system (1).

Now the equations of motion for two agents can be written in the following form:

$$\dot{x}_1 + \alpha_1 x_1 = \Gamma_1(t), \tag{10}$$

$$\dot{x}_2 + \alpha_2 x_2 = \Gamma_2(t). \tag{11}$$

One should recall that without quantum entanglement, the Langevin forces, although statistically equivalent, are not point-to-point identical, i.e.,

$$\Gamma_1(t) \not\equiv \Gamma_2(t). \tag{12}$$

Therefore, one cannot eliminate them from Eqs. (10) and (11). However, in case these agents are entangled

$$\Gamma_1(t) \equiv \Gamma_2(t), \tag{13}$$

one obtains a fully deterministic relationship between the motions of the agents:

$$\dot{x}_1 + \alpha_1 x_1 = \dot{x}_2 + \alpha_2 x_2. \tag{14}$$

In the general case, the entangled version of Eq. (1) yields:

$$\left[\dot{x}_i^{(j)} - f_i^{(j)}(\{x^{(j)}\})\right] g_{kk} = \left[\dot{x}_i^{(k)} - f_i^{(k)}(\{x^{(k)}\})\right] g_{jj}, \quad i = 1, 2, \ldots, d, \;\; k, l = 1, 2, \ldots, n. \tag{15}$$

In the following section we will give an interpretation of Eqs. (14) and (15) from the viewpoint of agents' cooperation.

4. Transmission of conditional information

Let us turn to Eq. (14) and assume that the agent A_1 can detect its own motion (which has been random before its detection), i.e., the agent A_1 detects $x_1^*(t)$. Then, based upon Eq. (14), he can calculate exactly the motion of the agent A_2 (although it is also random):

$$x_2(t) = x_2(0)\,\mathrm{e}^{-\alpha_2 t} + \int_0^t \mathrm{e}^{-\alpha_2(t-t')}\left[\dot{x}_1^*(t) + \alpha_1 x_1^*(t)\right] \mathrm{d}t' \ldots \tag{16}$$

It is implied that the agent A_1 knows the deterministic part of the governing equation of the agent A_2 as well as the initial condition $x_2(0)$. Analogously, if the agent A_2 can detect his own motion, he can calculate exactly the motion of the agent A_1 by the formula similar to Eq. (16). The same is true in the general case (15): if each agent can detect his own motion, he can calculate motions of all other agents. In terms of information it means that

$$I_{k|\ell} = 0, \quad k, \ell, = 1, 2, \ldots, n, \tag{17}$$

i.e., given the motion of the agent A_ℓ, the motions of all other agents A_k are fully predictable, and therefore, the conditional information capacity drops to zero due to entanglement.

It should be recalled that without the entanglement, the same information capacity, or the measure of unpredictability, is expressed by Eq. (5). However, based upon Eq. (5), one can argue that the gain in predictability due to the entanglement is small since the coefficients g_{ii} are small (see Eq. (3)). In order to remove this argument, we will give the following counter-example. Let us modify Eqs. (10) and (11) as follows:

$$\dot{x}_1 + x_1(1 + x_1)(1 - x_1) = g\Gamma_1(t), \quad g > 0, \tag{18}$$

$$\dot{x}_2 + x_2(1 + x_2)(1 - x_2) = g\Gamma_2(t), \quad g > 0. \tag{19}$$

The deterministic version of Eq. (18)

$$\dot{x}_1 + x_1(1+x_1)(1-x_1) = 0 \tag{20}$$

has three fixed points:

$$x_1^{(1)} = 0, \quad x_1^{(2)} = 1, \quad x_1^{(3)} = -1. \tag{21}$$

The first point is a repeller, and the others are attractors. If the motions starts at $x_1 = 0$, with the probability 1/2 it will be trapped at $x_1 = 1$ or $x_1 = -1$. The same result will be obtained for Eq. (18) since the Langevin force $\Gamma_1(t)$ with equal probability 1/2 can have sign (+) or (−) at $t = 0$. Since Eqs. (18) and (19) are identical, similar conclusion can be made about the solution to Eq. (19). In case when agents A_1 and A_2 are not entangled, the outcome of the solution to Eq. (18) does not help to predict the outcome of the solution to Eq. (19). However, if the agents A_1 and A_2 are entangled, and the agent A_1 found himself in the well $x_1 = 1$, then with the probability one he can conclude that the agent A_2 is in the well $x_2 = 1$ since, as a result of the identity (13),

$$\operatorname{Sign}\Gamma_1(0) = \operatorname{Sign}\Gamma_2(0). \tag{22}$$

It should be emphasized that this result does not depend upon the value of g in Eqs. (18) and (19): since the Langevin force triggers the instability, g can be vanishingly small. At the same time, the gain in predictability is significant: from total unpredictability (without the entanglement) to complete predictability (with the entanglement).

5. Topological complexity

So far we have only discussed two extreme cases when the active system is totally uncorrelated (see Eq. (1)) and when it is fully entangled (see Eq. (15)). However, there are a lot of other possibilities when certain groups of agents are entangled with each other, but not with the rest of the agents. Any such associations will form a particular topological structure. The number of structures is equal to the number of different partitions of a sequence

$$n = n_1 + n_2 + \cdots + n_m \quad (m \leqslant n) \tag{23}$$

of objects into m non-overlapping classes if each class n_α consists of indistinguishable n_α objects. This number is given by the multinomial coefficients:

$$p_m = \frac{n!}{n_1! n_2! \cdots n_m!}, \quad m = 1, 2, \ldots, n \tag{24}$$

and it can be associated with the number of topological complexions. Another type of partition can be based upon the master–slave principle: one group includes s independent (master) agents, and the second group includes the rest $n - s$ (slave) agents which are entangled with the masters such that in order to detect their motions, one has to know motions of all s masters,

$$I_{n-s|s} = 0. \tag{25}$$

The number of complexions in this case is

$$p' = \frac{n!}{s!(n-s)!}. \tag{26}$$

In order to include the topological changes into the active system, one can assume that each agent has n sets of N particles such that an rth set corresponds to possible entanglement with the rth agent. Then Eq. (1) can be rewritten as follows:

$$\dot{x}_i^{(j)} = f_i^{(j)}(\{x^{(j)}\}) + \varepsilon \sum_{k=1}^{m} g_{jk}\Gamma^{(k)}, \quad i = 1, 2, \ldots, d, \;\; j = 1, 2, \ldots, n; \;\; m \leqslant n, \tag{27}$$

$$\varepsilon = \text{Const.} > 0, \quad |\varepsilon| \ll |f|. \tag{28}$$

As it follows from Eqs. (27) and (28), the contribution of Langevin forces $\Gamma^{(k)}$ is still small in comparison to the deterministic forces $f^{(j)}$. The Langevin force $\Gamma^{(j)}$ of the agent A_j is now a linear weighted combination of forces $\Gamma^{(k)}$ resulting from possible entanglement with the agent A_k. By choosing a weight g_{jk}, the agent A_k is actually chooses a degree of a contribution from the agent A_j.

The condition that each topological structure consists of non-overlapping classes of agents can be now expressed as follows:

$$g_{ik} = 0 \quad \text{only if } g_{kj} = 0. \tag{29}$$

Indeed, under this condition, the entanglement of any two agents is reciprocal.

The type of topological structure is defined by the rank of the determinant of the matrix $\|g_{ik}\|$.

If $m = n = r$, then the active system (27) is totally uncorrelated. Indeed, in this case one cannot eliminate the Langevin forces $\Gamma^{(k)}$ in order to correlate the deterministic forces, and Eq. (27) can be reduced to Eq. (1). Another extreme is when the system is fully correlated, and is obtained if $r = 1$. Then Eq. (27) can be reduced to Eq. (15). In the case

$$1 < r < n, \tag{30}$$

one arrives at the variety of different master–slave topological structures, the total number of which follows from Eq. (26). Using the terminology of the statistical mechanics, one can introduce the Boltzmann entropy:

$$E = -\left(\frac{s}{n}\ln\frac{s}{n} + \frac{n-s}{n}\ln\frac{n-s}{n}\right). \tag{31}$$

If the determinant of the matrix $\|g_{ik}\|$ is representable as a product:

$$\det\left\|g_{jk}\right\| = \prod_{q=1}^{m} \det\left\|g_{jk}^{(q)}\right\|, \tag{32}$$

then the active system is partitioned into m non-overlapping classes (see Eq. (23)), and the number of complexions for a fixed m is presented by Eq. (24). The corresponding Boltzmann entropy takes the form:

$$E_m = -\sum_{i=1}^{m} \frac{n_i}{n}\ln\frac{n_i}{n}. \tag{33}$$

In the context of the entangled active systems, the entropy E_m can be associated with the information capacity due to the topological complexity.

6. Topological evolution

In this section we will discuss possible ways in which the topology of the active system (in terms of the degree of its entanglement) can change due to decentralized random choices made by the agents. So we will assume that each agent A_j, after certain time interval ΔT makes random choices for the weights g_{jk}. In order to separate the topological evolution from the dynamics at a fixed topology, one should choose

$$\Delta T \gg \Delta t, \tag{34}$$

where Δt is the timescale of the Langevin forces (see Eqs. (8) and (9)).

Since the topological changes do not require from the agents any "non-physical" properties (such as measurements and their conversion into the corresponding movements), the topological evolution will obey the laws of statistical mechanics, namely: the system will approach such a topological structure which can be arranged in the largest number of different ways, i.e., which is characterized by the maximum number of complexions P. As it follows from Eq. (24), P approaches its maximum at $m = n$, i.e., when

$$p = n! \tag{35}$$

This means that irrespective of the initial topology, the active system eventually will approach the topology of totally uncorrelated active system (1).

Let us assume now that after the same time interval ΔT, each agent A_j chooses equal contributions from all other agents, i.e.,

$$a_{jk} = a_{jq}. \tag{36}$$

Then the matrix $\|a_{jk}\|$ will be idempotent, and its rank

$$r = 1, \tag{37}$$

hence, one arrives at another extreme: as noticed above, the active system becomes fully correlated, and it is described by Eq. (16).

7. Topological self-organization

Let us now assume that the agents have an additional channel of entanglement that coordinates their global topology: each agent performs the corresponding measurements and receives the binary sequence of N pluses or minuses which can encode 2^N different messages. One should recall that due to the entanglement, all the agents receive identical sequences. Suppose that, according to a preliminary agreement between the agents, each of 2^N sequences encodes a certain topology in terms of the corresponding choice of the weights a_{jk}. In order to count all the possible topological structures, we will turn to Eq. (32): according to Eq. (24), there are P_m ways in which the matrix $\|a_{jk}\|$ can be partitioned into the product of m sub-matrixes, when $m = 1, 2, \ldots, n$; in addition to that, each sub-matrix can have the rank $r = 1, 2, \ldots, s_m$ where s_m is the number of master agents in the mth sub-matrix. Hence, the total number of different topological structures is:

$$P = \sum_{m=1}^{n} \frac{n!}{n_1! n_2! n_m!} \frac{n_m!}{s_m!(n_m - s_m)!}, \quad \sum_{m=1}^{n} n_m = n, \tag{38}$$

where n_m is the size of the mth sub-matrix. This number must be equal to the number of different entangled sequences 2^N, i.e., the number of entangled particles N,

$$N = \log_2 P. \tag{39}$$

As soon as an encoded topological structure is selected, its actual implementation requires from each agent to assign certain values to the weights g_{ik} which would satisfy the constraints imposed upon the matrix $\|a_{jk}\|$ by the selected topology. Obviously there are an infinite number of ways in which these weights can be chosen, and each agent can choose them randomly. Hence, although the additional channel of entanglement uniquely defines the topological structure of the active system, it leaves the agent free to choose an actual implementation. In order to specify such an implementation the agents must have an additional preliminary agreement, for instance: all the coefficients a_{jk} can have only binary values 0 or 1. Then there is only

$$P = \frac{(n^2)!}{((n^2/2)!)^2} \tag{40}$$

different matrixes $\|a_{jk}\|$ and each of them can be encoded by the entangled sequences generated by $\tilde{N}$ pairs of particles where

$$\tilde{N} = \log_2 \tilde{P}. \tag{41}$$

In this case, the structure of the active system, in terms of Eq. (27), is defined uniquely.

Thus one arrives at topological self-organization triggered by quantum entanglement in a random and decentralized way.

One should recall that transmission of conditional information via encoding of random messages delivered by quantum entanglement and exploited above has been discussed earlier in [4].

8. Entanglement-based cooperative computations

In this section we will return from the topological evolution to the dynamics with a fixed topology. Suppose that the agents have another preliminary agreement: to minimize a certain functional:

$$\varphi = \varphi\left[\{x(t)\}^1, \{x(t)\}^2, \ldots, \{x(t)\}^{(n)}\right]; \quad \{x\}^{(j)} = x_1^{(j)}, \ldots, x_\alpha^{(j)}, \tag{42}$$

where $\{x(t)\}^\alpha$ is the motion of the agent A_j.

For that purpose, each agent has to be equipped with a control parameter which changes its deterministic force:

$$f_j = f_j(x_j, \lambda_j). \tag{43}$$

Then, as it follows from Eq. (1), the motion of an agent A_j $(j = 1, 2, \ldots, n)$ will be a function of λ_j and a functional of the Langevin force $\Gamma^{(j)}$ while each agent will have its own version of it,

$$\varphi = \tilde{\varphi}_j\left[\lambda_1, \ldots, \lambda_n \Gamma^{(1)}(t), \ldots, \Gamma^{(n)}(t)\right], \quad j = 1, 2, \ldots, n. \tag{44}$$

Its minimum as a function of λ_j can be found by each agent A_j from the system of the following equations:

$$\frac{\partial \tilde{\varphi}^j}{\partial \lambda_k} = 0, \quad j, k = 1, 2, \ldots, n. \tag{45}$$

However, in the case (1) when the agents are not entangled, the explicit form of Eq. (45),

$$\lambda_j = \tilde{\lambda}_j\left[\Gamma^{(1)}(t); \ldots; \Gamma^{(n)}(t)\right] \tag{46}$$

defines λ_j as a stochastic process driven by the Langevin forces. This means that the control parameters are not known exactly, but they can be only predicted with some probability. As demonstrated in the example presented by Eqs. (18) and (19), such a prediction can be totally useless even if the variance of the Langevin force is vanishingly small. Thus, the cooperation of non-entangled agents, in principle, is ineffective.

Let us turn now to the fully entangled case described by Eq. (15). As it follows from these equations, due to the entanglement, the Langevin forces can be eliminated from the functional (44). Indeed, if the agent A_j $(j = 1, 2, \ldots, n)$ knows his own motion $\{x(t)\}^j$, and assuming again that he knows the deterministic parts of the models of all other agents as well as the corresponding initial conditions he can calculate exactly (using Eq. (15)) the motions of the rest of the agents. As a result of that, each agent will have the same functional in the form of the function of the control parameters λ_1,

$$\varphi_j = \tilde{\varphi}_0(\lambda_1, \ldots, \lambda_n) \tag{47}$$

and the same equations

$$\frac{\partial \tilde{\varphi}_0}{\partial \lambda_k} = 0, \quad \lambda_k = \lambda_k^0, \;\; k = 1, 2, \ldots, n. \tag{48}$$

Hence, the entangled agents minimize the functional (42) in a fully cooperative and the most effective way. In many cases, the function (47) has several local minima, and therefore, Eq. (48) has the corresponding number of roots. Then a dynamical approach to find the global minimum is more effective. For that purpose, Eq. (48) is replaced by the dynamical equation:

$$\dot{\lambda}_k = -\frac{\partial \tilde{\varphi}_0}{\partial \lambda_k}, \tag{49}$$

whose solution converges to the roots of Eq. (48). (This convergence is guaranteed by the fact that Eq. (49) forms a gradient system with $\tilde{\varphi}_0$ playing the role of a Lyapanov function). Depending upon the initial conditions, this system will converge to one of the local minima. In order to continue the search for the global minimum, the following strategy is applied: Eq. (49) is solved many times with different initial conditions,

$$\lambda_k^0 = \lambda_k(t = 0), \tag{50}$$

which are selected at random; then for each solution the value of the function $\tilde{\varphi}_0$ in (47) is computed, and the lowest value is accepted as the solution to the problem. However, if the agents are acting independently, the randomly chosen initial conditions (50) are different for each agent, and the cooperation is lost: eventually each agent will arrive at its own "version" of the global minimum subject to its probabilistic evaluations of the motions and the decisions of other agents.

However, if the agents are entangled, they can have an additional channel that encodes the initial condition (50) such that each agent receives the same randomly chosen values λ_k^0. Indeed, suppose that there is a set of possible s values for λ_k^0:

$$\lambda_{k_1}^0, \lambda_{k_2}^0, \ldots, \lambda_{k_s}^0. \tag{51}$$

Then this sequence can be encoded via a special channel of entanglement with N_S entangled particles.

$$N_S = \log_2 s \tag{52}$$

in the same way as described in Section 6 (see Eq. (41)), so that each agent will receive the same value λ_{k_i} after decoding the sequence of random entangled signals, and therefore all the agents will arrive at the same global minimum in a fully cooperative way.

The entanglement-based cooperation is even more important in case when the objective function (44) results from a combinatorial problem, for instance, a traveling salesman TSP problem. The main characteristic of such problem is that there is no analytical formulation for a gradient, and therefore, Eqs. (45) and (49) cannot be written down explicitly. One of the most reliable approaches to the solution of these problems is Monte Carlo method combined with the simulated annealing optimization technique. The algorithm starts with some initial configuration (for instance, initial tour in case of TSP) of the system and then changes are made to the system (for instance, the swap pairs of cities). These changes generate a numerical version of the gradient,

$$\frac{\partial\tilde{\varphi}^0}{\partial\lambda_k} \cong \frac{\Delta\tilde{\varphi}^0}{\Delta\lambda_k}, \quad k = 1, 2, \ldots, n. \tag{53}$$

Changes that decrease $\tilde{\varphi}^0$ are accepted unconditionally, while those that increase it are also accepted, but with a conditional probability,

$$w(\Delta g) = \mathrm{e}^{-\Delta g|t}, \tag{54}$$

where Δg is the amount of increase in the objective function, and T ("temperature") is a parameter that is subject to control. Condition (54) is supposed to prevent the process from being trapped in a local minimum since by occasionally accepting increases in the function $\tilde{\varphi}^\alpha$, the process can climb out of local minimum. The problem with this approach is in developing adequate annealing schedule, i.e., the prescription for the number of iterations required at each temperature T as this temperature is gradually lowered.

Thus, if the agents perform the cooperative computing, they must coordinate the changes $\Delta\lambda_k$, the number of iterations S required for each temperature T, the sequence of temperatures T_i, and the samples from the conditional probability (54).

The coordination of the first three sequences is similar to those considered in Eqs. (52) and (53) if one starts with the sets of acceptable values of $\Delta\lambda_k$, S and T_i, encodes them via the random entangled sequences and then decodes them after performing the appropriate measurements of the entangled particles: after this procedure, each agent will receive exactly the same (but randomly chosen) values of $\Delta\lambda_k$, S and T_i.

The coordination of the samples out of the conditional probability (54), i.e., the coordination of the decisions about acceptance or rejection of the increases in the objective function $\tilde{\varphi}^0$ requires several additional procedures. Suppose that each agent generates a Wiener process by establishing the following Langevin equation:

$$\dot{y}_j = \Gamma_j(t), \tag{55}$$

where y_i is an additional variable possessed by the agent A_j, and $\Gamma_j(t)$ is the Langevin force for the same type as those considered in Eq. (1).

The solution of Eq. (55) describes a Wiener process whose probability obeys the simplest version of the Fokker–Planck equation:

$$\frac{\partial R_j}{\partial t} = \frac{\partial^2 R_j}{\partial y^2} \tag{56}$$

subject to the boundary conditions

$$\frac{\partial R_j}{\partial y} = 0 \quad \text{at } y = 0 \text{ and } y = 1. \tag{57}$$

The solution to Eq. (57) eventually approaches the equal probability distribution:

$$R_j^0 = \text{Const.}, \tag{58}$$

i.e., the values of y_i will be uniformly distributed over the interval $(0, 1)$.

Then the variable

$$Z_j = -\ln(1 - y_j) \tag{59}$$

will be exponentially distributed:

$$W(Z_j) = \mathrm{e}^{-z_j}. \tag{60}$$

Since the agent A_j receives the value of $T_j = T$ and he can calculate the increase of the objective function $\Delta g_j = \Delta g$, by making the substitution in Eq. (60),

$$Z_j = \frac{\Delta g}{T}, \tag{61}$$

he can find the probability of acceptance of the increase Δg

$$W - \mathrm{e}^{-\Delta g|T} \tag{62}$$

and simulate the corresponding stochastic process by using Eqs. (59) and (61).

As soon as the Langevin forces $\Gamma_j(t)$ are entangled, they become point-to-point identical:

$$\Gamma_j(t) \equiv \Gamma_k(t). \tag{63}$$

The same property is automatically transmitted to the stochastic process y_i and Z_j, i.e.,

$$y_j(t) \equiv y_k(t), \quad Z_j(t) \equiv Z_k(t). \tag{64}$$

Hence all the agents will make identical decisions about the acceptance or the rejection of the increase Δg, and therefore their full cooperation in minimization of the function $\tilde{\varphi}$ is preserved.

One should recall that most of the computational problems can be reduced to the minimization of a function or a functional of the type (44), and all of these problems, in principle, can be solved by entangled agents in a fully cooperative way.

9. Team games

In this section we will turn from the entanglement-based cooperation to the entanglement-based competition.

Suppose that the agents are divided into two entangled groups:

$$A_1, \ldots, A_n \quad \text{and} \quad B_1, \ldots, B_{n_2}, \qquad n_1 + n_2 = n \tag{65}$$

under the condition that the agents from group A are not entangled with the agents from group B, and assume that the task of the agents A is to minimize the functional (42) using the control parameters λ'_j (see Eq. (44)), while the task of the agents B is to maximize the minimum of this functional using similar control parameters λ''_j.

One should recall that each agent A_j can exactly calculate the motion $x'_k(t)$ of the agent A_k given that he can detect his own motion $x'_j(t)$ (see Eq. (15)), but he can calculate the motion of the agent B_g only probabilistically using Eq. (4). For instance, he can find the expected motion of the agent B_g:

$$\tilde{x}''_q(t) = \int_{-\infty}^{\infty} \xi \rho(\xi, t)\, \mathrm{d}\xi, \tag{66}$$

where $\rho(x''_q, t)$ is defined by the corresponding Fokker–Planck equation (4) which, according to the preliminary agreement between the agents, is known in advance to both entangled and non-entangled agents.

Thus, each agent can reduce the original functional to the following functions to be optimized:

$$\varphi^A_j = \varphi^A\left(\lambda'_1, \ldots, \lambda'_{n_1},\ \tilde{\lambda}''_1, \ldots, \tilde{\lambda}''_{n_2}\right), \quad j = 1, 2, \ldots, n_1, \tag{67}$$

$$\varphi^B_j = \varphi^B\left(\lambda''_1, \ldots, \lambda''_{n_2},\ \tilde{\lambda}'_1, \ldots, \tilde{\lambda}'_{n_1}\right), \quad j = 1, 2, \ldots, n_2. \tag{68}$$

The agent A_j can find exact dependence of the function φ^A upon the control parameters λ'_k $(k = 1, 2, \ldots, n_1)$ which belong to the agents of the same group, as he could do in the case (47). But the dependence of the function φ^A upon the control parameters λ''_k $(k = 1, 2, \ldots, n_2)$ which belong to agents of the different group is approximate since these parameters enter Eq. (67) via probabilistic relationships such as expectations (66). That is why, in general

$$\lambda''_q \neq \tilde{\lambda}''_q, \tag{69}$$

where $\tilde{\lambda}''_q$ results from the substitution in Eq. (42) the expectations $\tilde{x}''_q(t)$ (see Eq. (66)) instead of the exact function $\tilde{x}''_q(t)$. The same is true for the function φ^B_j where, in general,

$$\lambda'_q \neq \tilde{\lambda}'_q. \tag{70}$$

Now each agent A_j has to minimize the function (68) subject to the control parameters λ'_k $(k = 1, 2, \ldots, n_1)$ by solving the system of equations of the type (50),

$$\dot{\lambda}'_k = -\frac{\partial \varphi^A}{\partial \lambda'_k}, \quad k = 1, 2, \ldots, n_1. \tag{71}$$

However, in contradistinction to Eq. (50), the system (71) is not closed since the function (67) contains the additional control parameters λ''_k (or, to be more precise, their approximations $\tilde{\lambda}''_k$) which are operated by the adversary agent B.

The simplest choice for the agent A_j is not to vary $\tilde{\lambda}''$ at all:

$$\dot{\tilde{\lambda}}''_k = 0, \quad k = 1, 2, \ldots, n_2. \tag{72}$$

However, the more effective choice would be to assume that

$$\dot{\tilde{\lambda}}''_k = \frac{\partial \varphi^A}{\partial \tilde{\lambda}''_k}, \quad k = 1, 2, \ldots, n_2, \tag{73}$$

since that simulates the attempt of the agents B to maximize the minima of the function φ^A by moving its value against the corresponding gradient-descent.

The same is true for the agents B:

$$\dot{\lambda}''_k = -\frac{\partial \varphi^B}{\partial \lambda''_k}, \quad k = 1, 2, \ldots, n_2, \tag{74}$$

$$\dot{\tilde{\lambda}}'_k = \frac{\partial \varphi^B}{\partial \dot{\tilde{\lambda}}'_k}, \quad k = 1, 2, \ldots, n_1. \tag{75}$$

Strictly speaking, Eqs. (71)–(75) should be solved simultaneously with the equations of motions (1) or (15). However, practically the timescale of Eqs. (71)–(75) is supposed to be much smaller than those of Eqs. (1) and (15), i.e., the latter can be considered as "frozen" in time. Under those conditions, Eqs. (71), (73) and Eqs. (74), (75) are decoupled. Each of them does not represent a gradient system anymore because it simulated conflicting objectives of the agents A and B. Therefore, the solutions to these systems (which, in general, will be different) may approach stationary, periodic, or chaotic attractors, and that, eventually will effect the motions of the agents as well as the values of the functionals to be optimized.

10. Maxwell's demon and inverse diffusion

An appearance of Maxwell's demon effects in the system under consideration could be predicted from the very beginning since the agents are represented by physical particles with intelligence. Although this intelligence is restricted to receiving and interpretation only totally random messages, it is sufficient to demonstrate phenomena that are in apparent "violation" of the second law of thermodynamics. For that purpose we will turn to the governing equations of the active system in the form (27), and assume that the weights g_{ij} depend upon the agents' velocities as follows:

$$g_{jj} = 1, \quad g_{jk} = \begin{cases} 1 & \text{if } \upsilon_j^2 < \upsilon_0^2 \\ 0 & \text{if } \upsilon_j^2 > \upsilon_0^2 \end{cases}, \quad \upsilon_0 = \text{Const.}, \quad \upsilon_j^2 = \sum_{i=1}^{d} \left(\dot{x}_i^{(j)}\right)^2, \tag{76}$$

where υ_j is the velocity of the jth agent, and υ_0 is the velocity chosen to separate slow and fast agents.

The rule (76) for g_{jk} can be implemented in a dynamical way. Indeed, suppose that each g_{jk} is governed by the following dynamical equation:

$$\dot{g}_{jk} = g_{jk}(g_{jk} - 1)\left(\upsilon_j^2 - \upsilon_0^2\right), \quad k = 1, 2, \ldots, n-1; \; k \neq \upsilon_0. \tag{77}$$

This equation has two equilibrium points:

$$g_{jk} = 0 \quad \text{and} \quad g_{jk} = 1. \tag{78}$$

When $\upsilon_j^2 < \upsilon_0^2$, the first point in Eq. (78) is a repeller, and the second one is an attractor, i.e., $g_{jk} = 1$. Conversely, when $\upsilon_j^2 < \upsilon_0^2$, the first point is an attractor, and the second one is a repeller, i.e., $g_{jk} = 0$. Therefore, the rule (76) is implemented regardless of the initial values $g_{jk}(t = 0)$. It is implied that the dynamics in Eq. (77) must have much smaller scale than those in Eq. (27) so that the agent' velocities υ_j can be treated as frozen during the transient dynamics in (77).

One should notice that in order to implement the rule (76), the jth agent does not need to know the weights of the other agents, i.e., he acts autonomously.

After the rule (76) is implemented, the matrix of the weights attains the following form:

$$\|g_{jk}\| = \begin{pmatrix} 1 & 1 & \cdots & \cdots & 1 \\ \vdots & & & & \\ 1 & 1 & \cdots & \cdots & 1 \\ 0 & \cdots & 010 & \cdots & 0 \\ 0 & \cdots & \cdots & 010 & \cdots \\ 0 & \cdots & \cdots & \cdots & 01 \end{pmatrix}. \tag{79}$$

Its determinant can be factored as

$$\det \|g_{jk}\| = \det \begin{pmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{pmatrix} \det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix}. \tag{80}$$

The first sub-matix is idempotent, and its rank

$$r_1 = 1. \tag{81}$$

Hence, according to the results presented in Section 4, the corresponding sub-group of slow agent A_1 becomes fully entangled regardless of the initial values of the corresponding weights g_{jk} $(j = 1, 2, \ldots, n_1;\ k = 1, 2, \ldots, n)$.

The second sub-matrix in Eq. (80) is the identity matrix, and its rank is equal to its dimensionality:

$$r_2 = n_2. \tag{82}$$

This means that the subgroup A_2 of fast agents becomes totally uncorrelated.

Let us now turn to the statistical interpretation of the situation. Suppose, for simplicity, that the initial state of the agents was characterized by the weights:

$$g_{jk} = \delta_{jk} = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases} \tag{83}$$

Then as it follows from Eq. (27), the rank of the weight matrix

$$\text{rank } \|g_{ik}\| = n \quad \text{at } t = 0. \tag{84}$$

This means that initially all the agents were independent, and the corresponding active system performed Brownian motion. Its temperature as a macroscopic characteristic could be found from the average kinetic energy (*for* $d = 3$):

$$T = \frac{\varepsilon}{\kappa} \sum_{i=1}^{3} \sum_{j=1}^{n} \left\langle \left(\dot{x}_i^{(j)} \right)^2 \right\rangle. \tag{85}$$

After the topological self-organization (77), the slow agents will have the temperature lower than (85), while the fast agents will raise their temperature so that

$$T_1 < T < T_2. \tag{86}$$

But in addition to that, the groups of agents A_1 and A_2 will depart from each other in space. In order to show that, start with the group of fast agent A_2 who will still perform a Brownian motion (see Eq. (82)). If n_2 is sufficiently large $(n_2 \gg 1)$, it is fair to say that the center of mass of this group O_2 will not change its position in space.

Let us turn now to the group A_1. In order to trivialize the situation, we will assume that in Eq. (27):

$$f_i^{(j)}\left(\{x^{(j)}\}\right) \equiv 0. \tag{87}$$

Then one arrives at the following (simplified) version of Eq. (27):

$$\dot{x}_j^{(j)} = \varepsilon \sum_{k=1}^{n_1} \Gamma^{(k)}, \quad i = 1, 2, 3. \tag{88}$$

As it follows from Eq. (86),

$$\dot{x}_i^{(j)} = \dot{x}_i^{(q)}, \tag{89}$$

i.e., all the agents of group A_1 will move along equidistant trajectories, and therefore, the whole group will perform a rigid translationary motion. At the same time, each particular trajectory can be found from the Langevin equation (88) in the form of a realization of the Wiener stochastic process (see Eq. (4)):

$$\frac{\partial \rho_i^{(j)}}{\partial t} = \varepsilon n_1 \frac{\partial^2 \rho_n^{(j)}}{\partial x_i^2}, \quad j = 1, 2, \ldots, n_1; \;\; i = 1, 2, 3, \tag{90}$$

whence

$$\rho_i^{(j)} = \frac{1}{\sqrt{4\pi\varepsilon n_1 t}} \exp\left[-\frac{x_i^2}{4\varepsilon n_1 t}\right], \quad i = 1, 2, 3. \tag{91}$$

Thus, the rigid translationary motion of agent A_1 can be characterized by only one trajectory, for instance, the trajectory of the center of mass O_1 of these agents.

Now let us try to evaluate the distance O_1O_2 between the centers of mass of the groups of agents A_1 and A_2. Since at $t = 0$, the total system performed a Brownian motion, both the agents A_1 and A_2 were indistinguishable forming a homogeneous mixture. Consequently, their centers of mass coincided:

$$O_1O_2 = 0 \quad \text{at } t = 0. \tag{92}$$

After the topological self-organization (77), the center O_2 has not changed its position (as was noticed earlier), but the center O_1 started its random motion according to Eq. (91). (Here we have chosen the case when the motions along each of three space coordinates x_1, x_2, and x_3 are identical.)

The expected distance between the centers of mass:

$$O_1O_2 = \sqrt{\sigma} = \sqrt[4]{2\pi\varepsilon n_1 t} \to \infty \quad \text{as } t \to \infty, \tag{93}$$

where σ is the variance of the stochastic process (91). Thus, eventually the agents A_1 and A_2 will be separated in space, and one arrives at apparent violations of the second law of thermodynamics. Indeed, the entropy of the total system decreases (since the agent A_1 becomes organized), heat flows from a cold place T to a hot place T_2, the difference of the temperatures $T_2 - T_1$ can generate work, the initially homogeneous mixture of agents A_1 and A_2 turns into two separated groups, and that can be characterized as an inverse diffusion, etc. The resolution is that the active system under consideration is not isolated: it receives information (via the quantum entanglement) from the external world, and this information is shared and processed by the agents. However, according to Landauer's principle, any erasure of information is necessarily a dissipative process. But any information processing performed by the agents must include these erasures. Similar energy dissipation can be associated with other components of the information processing such as measurement, recording, etc.

11. Conclusions

The main purpose of this study was to better understand the degree of usefulness of entanglement-based communications without classical channels. Obviously that imposed certain limitations upon the correlations between the agents; in particular, it forced us to create some special abstract assumptions which, at the first sight, are far from practical applications. Nevertheless it has been demonstrated that even under a severe constraint imposed by the absence of a classical channel, and on the contrary to a well-established belief that entanglement-based messages are useless, the active system can perform a number of emerging phenomena which lead to cooperation, competition, self-organization, etc.

Acknowledgements

The research described in this paper was performed by the Jet Propulsion Laboratory, California Institute of Technology. This research was performed in association with the Center for Engineering Science Advanced Research (CESAR), Oak Ridge National Laboratory. Funding was provided by the Engineering Research Program of the DOE

Office of Science, under contract DE-AC05-000R22725 with UT-Battelle LLC. The author likes to thank the reviewer for some helpful comments.

References

[1] Mikhailov AS. Foundations of synergetics. New York: Springer; 1990.
[2] Zak M. Dynamics of intelligent systems. Int J Theor Phys 2000;39(8).
[3] Risken H. The Fokker–Planck equation. New York: Springer; 1989.
[4] Zak M. Entanglement-based communications. Chaos, Solitons & Fractals 2002;13:39–41.
[5] Mittelstaedt P. Ann Phys (Leipzig) 1998;7(7/8):710–5.