Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Theoretical Biology 230 (2004) 227–239
Journal of Theoretical Biology
www.elsevier.com/locate/jtbi

Replicator–mutator equation, universality property and population dynamics of learning

Natalia L. Komarova[a,b,*]

[a] Department of Mathematics, Rutgers University, 110 Frelinghuysen Rd., Piscataway, NJ 08854, USA
[b] Institute for Advanced Study, Einstein Dr., Princeton, NJ 08540, USA

Received 10 October 2003; received in revised form 27 March 2004; accepted 5 May 2004

Available online 19 June 2004

Abstract

Replicator–mutator equation is used to describe the dynamics of complex adaptive systems in population genetics, biochemistry and models of language learning. We study "localized", or "coherent", solutions, which are especially relevant in the context of learning and correspond to the existence of a predominant language in the population. There is a coherence threshold for learning fidelity, above which coherent communication can be maintained. We prove the following surprising universality property of coherence threshold: for typical realizations of random coefficients in the fitness matrix, the value of the coherence threshold does not depend on the size of the system.

Keywords: Evolution of language; Quasi-species; Error threshold

1. Introduction

The subject of this study is the so-called *replicator–mutator equation* which describes dynamics of complex adaptive systems (Levin, 2002). This equation appears in many different contexts in biology, such as population genetics (Hadeler, 1981), autocatalytic reaction networks (Stadler and Schuster, 1992) and models of language evolution (Nowak et al., 2001).

Let us suppose that there are n types, $G_1, \ldots, G_n$. Depending on the biological context, these can be different alleles, polynucleotide molecules, or grammars, as will be explained below. The frequency of each of the types is denoted by $x_1, \ldots, x_n$. We will assume that these types undergo selection. In other words, the reproduction rate of each type, G_j, is determined by its fitness, f_j, which in the most general case is a function of all frequencies, $f_i = F_i(x_1, \ldots, x_n)$. If we take a polynomial expansion of fitness in terms of x_i, and only keep the linear terms, we obtain

$$f_j = w_j + \sum_{m=1}^{n} a_{jm} x_m + \cdots, \quad 1 \leqslant j \leqslant n. \tag{1}$$

Here w_j is the inhomogeneous (frequency-independent) part of fitness, and a_{ij} are entries of an $n \times n$ matrix, A. The fidelity of reproduction can be imperfect. Let Q_{ji} be the probability that type G_i is produced by type G_j. Q is a row-stochastic matrix so that

$$\sum_{j=1}^{n} Q_{ij} = 1, \quad 1 \leqslant i \leqslant n.$$

The changes in the frequencies of types $G_1, \ldots, G_n$ in time can then be described by the following model:

$$\dot{x}_i = \sum_{j=1}^{n} f_j x_j Q_{ji} - x_i \phi, \tag{2}$$

where

$$\phi = \sum_{j=1}^{n} f_j x_j$$

is the average fitness of the population. Note that this definition of ϕ leads to the conservation law of the form

$$\sum_{i=1}^{n} x_i = 1.$$

Eq. (2) is the replicator–mutator equation (Bürger, 1998). It is quite general and contains as limiting cases many other important equations in biology. In

*Corresponding author. Department of Mathematics, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854, USA.
E-mail address: komarova@math.rutgers.edu (N.L. Komarova).

0022-5193/$ - see front matter
doi:10.1016/j.jtbi.2004.05.004

particular, setting $Q = I$, the identity matrix, in (2), yields the well known replicator equation (Hofbauer and Sigmund, 1998) used in game theory. On the other hand, letting $a_{ij} = 0$ for all i and j leads to the quasi-species equation of molecular evolution (Eigen and Schuster, 1979). It has been recently observed by Page and Nowak (2002) that replicator–mutator equation is equivalent to a generalized Price equation of evolutionary genetics (Price, 1970).

In the next section, we give a brief description of several biologically relevant contexts where the replicator–mutator equation is used.

1.1. Three biological contexts

1.1.1. Population genetics

Let us suppose that a gene locus has n alleles, $G_1, \ldots, G_n$, and let $x_1, \ldots, x_n$ be relative frequencies of the alleles at the time of mating (Hofbauer and Sigmund, 1998). The relative frequency of a new gene pair, (G_iG_j), is proportional to the product x_ix_j. Due to natural selection, only a fraction, $a_{ij}x_ix_j$ of them will survive to maturity, provided that the types of the two alleles are independent. Here a_{ij} is the fitness parameter. In this context, Q is a mutation matrix, $w_i = 0$ for all i and ϕ is the average fitness of the population. The differential equation describing the production of new genes is (2).

1.1.2. Autocatalytic reaction networks

The n types are n *replicators*—for instance, polynucleotide molecules, RNA or DNA—which are capable of self-replication (Stadler and Schuster, 1992). Their concentrations are denoted as $x_1, \ldots, x_n$. Let f_j be a reaction rate of type G_j. Then expression (1) is the polynomial expansion of reaction rates in terms of concentrations. Let us denote as Q_{ij} the mutation rate which determines the fraction of replications from template G_i which yield G_j as error copy. Then, we can represent the kinetics as the differential equation (2). In this context, the quantity ϕ controls the total concentration of replicating molecules.

1.1.3. Population language learning

Imagine a group of individuals (or linguistic agents), each speaking one of the n possible variants of grammar, $G_1, \ldots, G_n$. Individuals reproduce, and children learn the language of their parents. Denote by x_i the relative abundance of individuals who use grammar G_i. The process of learning is modelled as follows.

Children receive language input (sample sentences) from their parents, as well as their siblings, peers and other adults in the community, and develop their own grammar. The elements Q_{ij} of the stochastic matrix, Q, denote the probability that a child born to an individual using G_i will develop G_j .The main source of change comes from the interactions within the linguistic community; presumably, if exposed to a homogeneous source, children would acquire the source grammar precisely. This leads us to conjecture that in the most general case, the stochastic matrix, Q, is a function of $x_1, \ldots, x_n$

$$Q(x) = Q + \sum_{i=1}^{n} Q_i^{(1)} x_i + \cdots.$$

In the current setting we only consider the first term in this general expression; the effects of the x-dependent terms will be considered elsewhere. The matrix Q sets the rate of language change. The quantities Q_{ii} measure the accuracy (or fidelity) of grammar acquisition.

Let us specify the similarity between grammars by introducing the numbers $0 \leqslant s_{ij} \leqslant 1$ which denote the probability that a speaker who uses G_i will say a sentence that is compatible with G_j. We assume there is a reward for mutual understanding. The payoff for someone who uses G_i and communicates with someone who uses G_j is given by

$$a_{ij} = (s_{ij} + s_{ji})/2. \tag{3}$$

This is simply the average taken over the two situations when G_i talks to G_j and when G_j talks to G_i. Assume that everybody in the population talks to everybody else with equal probability. Therefore, the average payoff for all those individuals who use grammar G_j is given by $\sum_{m=1}^{n} a_{jm}x_m$. We assume that the payoff derived from communication contributes to biological fitness, and w_j is the language-unrelated component of fitness. For simplicity, let us set

$$w_j = f_0, \quad 1 \leqslant j \leqslant n.$$

Individuals leave offspring proportional to their payoff, and the offspring learn (possibly with mistakes) their grammar. The population dynamics of grammar evolution are then given by Eq. (2), where $\phi = \sum_{i=1}^{n} f_ix_i$ denotes the average fitness; its language-related part, $\phi - f_0$, has the meaning of *grammatical coherence* of the population. The grammatical coherence defines the probability that a randomly chosen sentence of one person is understood by another person. It is a measure for successful (coherent) communication in a population. If $\phi - f_0 = 1$, all sentences are understood and communication is perfect. In general, $\phi - f_0$ is a number between 0 and 1.

Note that here, we assume that the language contribution to fitness comes from the individuals' ability to transfer information. This is a simplification. It is well known that language is used for other purposes, e.g. to disguise information, to lie, to identify oneself with a certain group and to separate oneself from other groups. However, despite differences across the population which distinguish class, status etc., the language of a community is *mostly* understandable by all members. This suggests that the predominant mechanism of selection which acts on language is the one that makes it uniform, and therefore is related to the role of language in information transfer. Other mechanisms contribute to the variation which is observed on top of that.

1.1.4. Applicability of the replicator–mutator equation

As we have seen, the replicator–mutator equation has relevance to many areas of science including genetics, theoretical biochemistry, language evolution and population biology. Like any universal equation (e.g. the famous nonlinear Schroedinger equation in physics), replicator–mutator equation is an approximation to reality, but it does grasp many important features of the dynamics, common to a wide variety of systems. Every time we use it, we need to be aware of its shortcomings, no matter what area of science we apply it to. However, this does not undermine the fact that this equation can be a first step towards building more sophisticated models of complex systems.

In this paper, we study some general properties of the replicator–mutator equation. For convenience, we choose to use the terminology which comes from language evolution models, so that $G_1, \ldots, G_n$ are the possible grammars, and matrix Q_{ij} defines the learning accuracy of children.

1.2. Evidence for universality property in simple systems

This work is a part of a larger effort to understand the properties of selection–mutation dynamics. Many interesting results obtained in this area are reviewed in Bürger (1998). In this paper, we concentrate on the behavior of fixed points of Eq. (2).

In order to get a feeling of what the behavior of this system may be like, let us consider the following simplified model (Komarova et al., 2001). We impose a symmetry condition on the matrices A and Q such that

$$a_{ij} = a, \quad i \neq j, \quad a_{ii} = 1, \tag{4}$$

$$Q_{ij} = \frac{q}{n-1}, \quad i \neq j, \quad Q_{ii} = 1 - q, \quad 1 \leqslant i, j \leqslant n, \tag{5}$$

where $0 < a < 1$ is some constant and q is a control parameter. For simplicity, we will also assume that $w_i = 0$ for all i. We will refer to the corresponding system as a *fully symmetrical* system.

The quantity q ranges from $q = 1 - 1/n$ to 0. For $q = 1 - 1/n$, the outcome of "learning" is random: the chance for the learner to pick grammar G_k is the same for all k, no matter what the teacher's grammar is. For $q = 0$, learning is error-free; learners always end up speaking the grammar of their teachers. In this simple case, the fixed points of the system can be found exactly and the phase portrait of the system can be obtained analytically, see Fig. 1.

When q is close to 1 (low-learning accuracy), there is only one stable fixed point in the system which we call the *uniform equilibrium*: all the grammars are represented in the population and have an equal abundance, $x_i = 1/n$ for all i. The coherence in this case is given by

$$\phi_{uniform} = a,$$

see Fig. 1. As q decreases, a bifurcation occurs where n equivalent *one-grammar equilibria* appear, each of them corresponds to a particular *dominant grammar*. These solutions can be viewed as *localized* solutions of the system (as opposed to the *de-localized*, uniform solution), because they have the form

$$x_k = X, \quad x_i = \frac{1-X}{n-1}, \quad i \neq k,$$

where G_k is the dominant grammar. The value of X can be expressed in terms of the parameters of the system and tends to $X = 1$ as $q \to 0$.

It is possible to calculate the coherence threshold, i.e. the value of q, q_c, such that for $q < q_c$ one-grammar solutions exist and are stable. It is given by

$$q_c = \frac{1-\sqrt{a}}{1+\sqrt{a}} + O(1/n). \tag{6}$$

The function q_c decreases with a. This is to be expected because as the grammars get harder to distinguish, which corresponds to $a \to 1$, we need higher and higher learning accuracy to reach coherence, i.e. $q_c \to 0$. The coherence of the population corresponding to one-grammar solutions at the threshold value of q is

$$\phi_{one\text{-}grammar} = \frac{2a}{1+\sqrt{a}} > \phi_{uniform}.$$

We can see that as soon as such solutions appear, the population fitness, ϕ, increases with a jump. As q

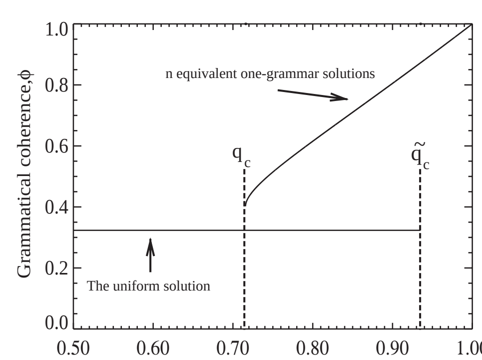

Fig. 1. The bifurcation diagram for fully symmetrical systems. Here, $a = 0.3$, $n = 30$ and $f_0 = 0$. At $q = q_c$, n equivalent one-grammar solutions appear. Each one-grammar equilibrium is characterized by a dominant grammar used by a majority of the population; the rest of the grammars are used by the remaining population and have equal (and small) frequencies. At $q = \tilde{q}_c$, a uniform (low coherence) solution becomes unstable.

continues to decrease to 0, $\phi_{one\text{-}grammar}$ approaches 1, which means perfect coherence. Because of symmetries in the system, the values of ϕ corresponding to one-grammar solutions with different dominant grammar are equal to each other.

We have the following remarkable **Universality property**

The threshold value, q_c, for the learning accuracy tends to a constant for large values of n, i.e. is independent of the size of the system in the limit $n \to \infty$.

Note that a similar universality result can be proved for another particular case of system (Hadeler, 1981). For a general A matrix, let us assume that the matrix Q has the form

$$Q = I + q,$$

where q is a matrix and $q_{ij} = \varepsilon_j$ for $i \neq j$, so the error rates only depend on the *target* grammar. It turns out that system (2) in this case is a gradient flow (see Hofbauer and Sigmund, 1998, Theorems 20.3.3 and 20.3.4), and the localized, one-grammar solutions persist for all $|q_{ii}|$ smaller than a constant which does not depend on n.

It is this universality property that we are going to study for a general class of matrices Q.

1.3. Model assumptions and the main result

In this paper, we will consider the general system (2) under the following mild assumptions:

$$w_j = f_0, \quad 1 \leqslant i, \ j \leqslant n, \tag{7}$$

$$a_{ij} = a_{ji}, \quad 0 \leqslant a_{ij} \leqslant 1. \tag{8}$$

Eq. (7) requires the inhomogeneous part of the fitness function to be the same for all types, and Eq. (8) is a symmetry assumption on the A matrix (which is obviously much weaker than the assumption made for the fully symmetrical system). In addition, we are going to bound the entries of the A matrix by setting

$$0 \leqslant a_{ij} \leqslant 1.$$

Let us introduce a family of matrices, $Q = Q(N)$, such that

$$\lim_{N \to \infty} Q(N) = I. \tag{9}$$

Increasing N is similar to decreasing q in the fully symmetrical case because it increases the learning accuracy of the system. An example of a bifurcation diagram is given in Fig. 2. This is a numerical plot, where the values $a_{ii} = 1$ for all i, a_{ij} with $i \neq j$ are taken from a uniform distribution between zero and one and $Q(N)$ is parameterized according to the so-called memoryless learner algorithm (Niyogi, 1998). This is a typical bifurcation diagram for the system under consideration. As we can see, it is in some sense similar to the diagram in the fully symmetrical case, except localized solutions corresponding to different dominant grammars appear at different points and correspond to different values of grammatical coherence, ϕ.

In this paper, we will study the bifurcation diagram of such systems and find out whether the appearance of localized solutions can be controlled independently of the size of the system. In order to approach this problem, we will

- study the system in the limit $Q = I$ (perfect fidelity),
- employ a perturbation analysis for the Q matrices not too far from identity (imperfect fidelity), and investigate the persistence of stable fixed points,
- get an estimate on how *far* the matrix Q can deviate from I so that the fixed points still exist and are stable; for this we will assume that the coefficients of the A matrix are random numbers and use their statistics in order to get a result for a typical realization.

It is convenient to introduce the parameter,

$$q = \max_i \, (1 - Q_{ii}).$$

The main result of this paper is the universality property: the typical threshold value of q such that localized solutions exist and are stable is independent on the size of the system (Theorem 3.3) for a very wide class of distributions of the entries a_{ij}. This work is a generalization of the analysis of Komarova et al. (2001) and Nowak et al. (2001); in those publications, a severely restricted class of fully symmetrical systems was studied. Here, we identify the universality property in a much more general class of systems and prove that it is indeed universal.

The rest of this paper is organized as follows. Section 2 describes the system with perfect learning fidelity. In this limit, we obtain the well known selection equation of population genetics. We study the stability of various

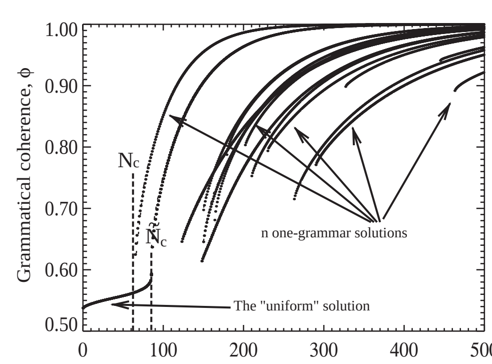

Fig. 2. The bifurcation diagram for a system with a random similarity matrix. Here, $n = 20$. At $N = N_c$, the first one-grammar solution appear. The learning accuracy matrix, Q, is calculated according to the memoryless learner algorithm, and depends on N, the number of learning events. At $N = \tilde{N}_c$, the low coherence solution equivalent to the uniform solution of the fully symmetrical system, disappears.

fixed points and prove in particular that for diagonally dominant matrices, coherent solutions are stable. In Section 3, we consider imperfect learning accuracy and show that solutions which are stable in the absence of learning mistakes, persist in the presence of learning mistakes. We give an estimate for the coherence threshold which guarantees coherence for a typical realization of the matrix, A. We also prove that this quantity tends to a finite number as the size of the system increases to infinity. In Section 4, we discuss some applications.

2. Perfect learning fidelity

We will start by studying the limiting case of system (2) for $N \to \infty$. This corresponds to the case of perfect learning fidelity (no mistakes of learning). Using Eq. (9), we obtain from Eq. (2),

$$\dot{x}_i = (f_i - \phi)x_i. \tag{10}$$

This is the *selection equation of population genetics*, also known as replicator equation for partnership games (Hofbauer and Sigmund, 1998). In Section 2.1, we give an explicit form of all the fixed points of system (10) together with their existence conditions. In Section 2.2, we present their stability conditions and prove some simple facts about coexistence of stable fixed points. In particular, we will show that if the matrix, A, is diagonally dominant, then the only stable solutions are the n one-grammar (coherent) solutions, corresponding to the n-dominant languages.

2.1. m-grammar solutions

We will refer to the domain

$$S_n = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i = 1 \right\} \tag{11}$$

as *simplex* S_n. We define an m-grammar solution of system (10), $\mathbf{x}^{(m)} = \bar{\mathbf{x}}^{(m)} \in S_n$, as the fixed point of the form:

$$\bar{\mathbf{x}}^{(m)} = (\bar{x}_1, \ldots, \bar{x}_m, 0, \ldots, 0)^{\mathrm{T}}, \quad \bar{x}_i = 0, \quad 1 \leqslant i \leqslant m. \tag{12}$$

Here m, $0 \leqslant m \leqslant n$, defines how many grammars have a non-zero share in the solution. Note that any fixed point can be written down in this way by proper renumbering of variables. It is easy to show that for these solutions,

$$f_1 = f_2 = \cdots = f_m = \phi, \tag{13}$$

which together with condition $\sum_{i=1}^{m} x_i = 1$ uniquely defines the m-grammar solution. In order to write down the solutions, let us introduce the following notations. We will call $A(m)$ the $m \times m$ matrix which is the upper left corner of the A matrix. Also, the matrix which is obtained from the $A(m)$ matrix by replacing the ith column by a column of ones will be called $B^{(i)}(m)$. Next, we define

$$T_m^i = det\, B^{(i)}(m), \quad 1 \leqslant i \leqslant m. \tag{14}$$

The m-grammar solution can be written as

$$\bar{x}_i = \frac{T_m^i}{\sum_{k=1}^{m} T_m^k}, \quad 1 \leqslant i \leqslant m, \quad \bar{x}_{m+1} = \cdots = \bar{x}_n = 0. \tag{15}$$

Note that this gives us the existence condition for the m-grammar solution, in order for $\mathbf{x}^{(m)}$ to belong to S_m, we need

$$T_m^i T_m^j \geqslant 0 \quad \forall i, j. \tag{16}$$

The fitness of the dominant grammars and the average fitness of the population corresponding to m-grammar solution is given by

$$\bar{\phi} = \bar{f}_i = f_0 + \frac{det\, A(m)}{\sum_{k=1}^{m} T_m^k}, \quad 1 \leqslant i \leqslant m \tag{17}$$

(the bar denotes that the function is evaluated at the fixed point). For instance, for $m = 1$ (one-grammar solutions) we have

$$\bar{x}_1 = 1, \quad \bar{\phi} = f_0 + a_{11}. \tag{18}$$

For $m = 2$ we have

$$\bar{x}_1 = \frac{a_{22} - a_{12}}{a_{11} - 2a_{12} + a_{22}}, \quad \bar{x}_2 = \frac{a_{11} - a_{12}}{a_{11} - 2a_{12} + a_{22}},$$
$$\bar{\phi} = f_0 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11} - 2a_{12} + a_{22}}. \tag{19}$$

2.2. Stability of m-grammar solutions

Let us perturb the m-grammar solution (12) with the vector $e^{\Gamma t}\mathbf{y}$ so that $\mathbf{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$ and $\sum_{i=1}^{n} y_i = 0$. The latter condition comes from the fact that solutions must stay within a simplex. Linearizing around the m-grammar solution we obtain an equation of the form $\mathscr{L}\mathbf{y} = \Gamma\mathbf{y}$ for the vector $\mathbf{y}$. Next, we write $y_m = -\sum_{i \neq m} y_i$. One of the n equations in the linear system for $\dot{\mathbf{y}}$ follows from the others, so we can get rid of the variable y_m and obtain a linear system for the $(n-1)$-dimensional perturbation vector with the matrix $\mathscr{L}'$ which has the following block structure:

$$\mathscr{L}' = \begin{pmatrix} L & L' \\ L_0 & L'' \end{pmatrix}, \tag{20}$$

where L is an $(m-1) \times (m-1)$ matrix defined as $L_{ij} = \bar{x}_i(a_{ij} - a_{im})$; L' is an $(m-1) \times (n-m)$ matrix with entries $L'_{ij} = \bar{x}_i[a_{ij} - a_{im} - 2\,(\bar{f}_j - \bar{\phi})]$; L_0 is an $(n-m) \times (m-1)$ matrix of zeros and L'' is an $(n-m) \times (n-m)$ diagonal matrix with elements $L''_{ii} = \bar{f}_i - \bar{\phi}$. The stability condition for an m-grammar solution is the *negative definiteness* of the matrix $\mathscr{L}'$. By Sylvester criterion, the positive numbers multiplying the rows of the matrix do not change the definiteness of the matrix, so we have the following

Theorem 2.1. *The necessary and sufficient conditions for the stability of an m-grammar solution* (12) *are given by*

$$A'(m) \quad \text{is negative definite}, \tag{21}$$

$$\bar{f}_j - \bar{\phi} < 0 \quad \forall j\colon\ m+1 \leqslant j \leqslant n, \tag{22}$$

where the matrix $A'(m)$ *is an* $(m-1)\times(m-1)$ *with* $A'(m)_{ij} = a_{ij} - a_{im}$.

Note that the matrix $A'(m)$ is obtained from the matrix $A(m-1)$ by subtracting the mth column of the matrix $A(m)$ from all its columns, and taking the upper left $(m-1)\times(m-1)$ corner.

Applying this theorem for the case of $m=1$, we obtain the following existence and stability condition for one-grammar solutions, need

$$a_{1j} \leqslant a_{11}, \quad 2 \leqslant j \leqslant n. \tag{23}$$

For stability of two-grammar solutions ($m=2$) we need

$$a_{11} < a_{12}, \tag{24}$$

$$a_{22} < a_{12}, \tag{25}$$

$$(a_{1j} - a_{11})\bar{x}_1 + (a_{2j} - a_{21})\bar{x}_2 < 0, \quad 3 \leqslant j \leqslant n. \tag{26}$$

All of the above can be explained if we notice that Eq. (10) possess a strict Lyapunov function (Hofbauer and Sigmund 1998, Theorem 7.8.1):

$$V(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} A \mathbf{x} = \sum_{i,j=1}^{n} a_{ij} x_i x_j. \tag{27}$$

We have $\dot{V} \geqslant 0$, and $\dot{V} = 0$ only at points defined by (12) and (13). This means that the only stable solutions of system (10) are maxima of the function V.

$V(\mathbf{x})$ is defined on the simplex S^n, $\sum_{i=1}^{n} x_i = 1$, and it can reach its maximum (maxima) in the interior or on the boundary of the simplex. An interior maximum of the function V corresponds to the n-grammar solution, and a maximum reached at a boundary, or, more precisely, at an interior point of the projection of the simplex to m dimensions, corresponds to an m-grammar solution.

There is a simple geometric interpretation of stability criterion (21) and (22) of m-grammar solutions. In order for a fixed point to be stable, it must correspond to a maximum of the function V. An m-grammar solution $x^{(m)}$ (of the form (28)) is a maximum if

(i) it corresponds to an interior maximum of the function V_m defined on simplex S_m, $\sum_{i=1}^{m} x_i = 1$, and

(ii) the function V_j defined on any simplex which includes S_m as a subset, satisfies $V_j(\mathbf{x}) < V_m(\mathbf{x}^{(m)})$ for any $\mathbf{x}$ in the vicinity of $\mathbf{x}^{(m)}$.

Lemma 2.2. *Condition* (i) *is equivalent to criterion* (21), *and condition* (ii) *is equivalent to criterion* (22).

Proof. Condition (i) implies that the function V is concave in simplex S_m. Let us express x_m as $1 - \sum_{i=1}^{m-1} x_i$ and form the matrix $A''(m) = \partial^2 V / (\partial x_i \partial x_j)$. The function V is concave in S_m iff the matrix $A''(m)$ is negative definite. But matrix $A''(m)$ is defined as $A''(m)_{ij} = a_{ij} - a_{im} - a_{jm} + a_{mm}$ and is formed by taking the matrix $A(m)$, subtracting the last column from all the columns, then subtracting the last row from all the rows, and taking the upper left $(m-1)\times(m-1)$ corner. It is easy to show that the rows of matrix $A''(m)$ are linear combinations of the rows of the matrix $A'(m)$, namely, $row''_i = row'_i + \sum_{j=1}^{m-1} \bar{x}_j \, row'_j / \bar{x}_m$, where row'_i is the ith row of the matrix $A'(m)$ and similarly with matrix $A''(m)$. The coefficients in these linear combinations are *positive*, and therefore the definiteness of the matrices $A''(m)$ and $A'(m)$ is the same. Thus condition (ii) is equivalent to condition (21).

Condition (ii) can be written down as $\Delta V < 0$, where $\Delta V = V(\mathbf{x}^{(m)} + \delta\mathbf{x}) - V(\mathbf{x}^{(m)})$. By taking $\delta\mathbf{x} = (\delta_1, 0, \ldots, 0, \delta_j, 0, \ldots, 0)^{\mathrm{T}}$ such that $m+1 \leqslant j \leqslant n$ and $\delta_1 + \delta_j = 0$, we obtain exactly conditions (22) for each j. $\square$

Remark 2.3. Condition (21) is equivalent to the condition of Kingman (1961) that the matrix $A(m)$ has strictly one positive eigenvalue and $m-1$ negative eigenvalues.

Proof. Let us replace the last row of the matrix $A(m)$ by the following linear combination of its rows: $row_m \to \sum_{j=1}^{m} \bar{x}_j \, row_j = (\phi, \ldots, \phi)$. Next, let us subtract the last column of this new matrix from the rest of its columns. These operations do not change the signs of eigenvalues of the matrix. On the other hand, the eigenvalues of the resulting matrix are ϕ and the eigenvalues of the matrix $A'(m)$. Therefore, the negative definiteness of A is equivalent to having only one positive and $m-1$ negative eigenvalues of the matrix $A(m)$. $\square$

With the geometric representation in mind, we can prove some simple statements about stable fixed points of the replicator equations.

Theorem 2.4. *Suppose that system* (10) *has a stable m-grammar solution*, (12), *with grammars* G_1 *through* G_m, *and k is some integer* $k < m$. *Then system* (10) *cannot have a stable k-grammar solution, where the k grammars involved are a subset of* $G_1, \ldots, G_m$.

Proof. Let us assume that both solutions,

$$\hat{\mathbf{x}}^{(m)} = (\bar{x}_1, \ldots, \bar{x}_m, 0, \ldots, 0), \tag{28}$$

$$\tilde{\mathbf{x}}^{(k)} = (\tilde{x}_1, \ldots, \tilde{x}_k, 0, \ldots, 0) \tag{29}$$

are stable (note that we can always enumerate grammars in such a way that an k-grammar solution involving a subset of $G_1, \ldots, G_m$ can be written in this way). The point $\hat{\mathbf{x}}^{(m)}$ is an interior maximum of function V_m defined on simplex S_m. Solution $\tilde{\mathbf{x}}^{(k)}$ is an exterior maximum reached on the boundary of this simplex. Let us connect the points $\hat{\mathbf{x}}^{(m)}$ and $\tilde{\mathbf{x}}^{(k)}$ with a straight line (it will be contained inside the simplex). Along this line, the function V_m has the form $V_m(\xi) = a\xi^2 + b\xi + c$, where ξ is the coordinate along the line, so that $\xi = 0$, corresponds to point $\hat{\mathbf{x}}^{(m)}$, and $\xi = \xi'$ is the boundary point of $S_m, \tilde{\mathbf{x}}^{(k)}$. The quantities a, b and c are some constants. Since $\xi = 0$ corresponds to an inner maximum, we have $N = 0$ and $a < 0$. It is easy to see that the function $V_m(\xi)$ cannot have a second maximum at the end point $\xi' \neq 0$. We have a contradiction which proves the statement of the theorem. □

Corollary 2.5. *If system* (10) *has n distinct stable one-grammar solutions, it can have no other stable fixed points.*

Proof. Let us assume that system (10) has a stable m-grammar solution with $m > 1$. Therefore, by Theorem 2.4, it cannot have a stable one-grammar solution $x_1 = 1$, which proves the corollary. □

Corollary 2.6. *If $a_{ii} = 1$ for all i, and $a_{ij} < 1$ for all $i \neq j$, then the only stable fixed points of system* (10) *are the n one-grammar solutions.*

Proof. The stability of one-grammar solutions follows from condition (23), and the uniqueness from corollary (2.5). □

2.3. The meaning of m-grammar solutions

m-grammar solutions correspond to the coexistence of several spoken languages. There are several reasons why many grammars are observed in the real world. The foremost reason is the existence of barriers separating groups of people from each other; these could be mountains, oceans, state borders. People mostly communicate within finite (and relatively small) groups. In this paper, we do not take account of such spacial effects; they have been addressed within our framework by several authors (Sasaki and Nowak, 2003; Solan et al., 2002). What we show here is the following subtle and counterintuitive point. There may be situations where an m-grammar solution exists even in communities where all people *do* talk among each other. The necessary and sufficient requirements for this are given in the previous section; they are conditions on pairwise eligibility of the languages.

Note that in population genetics and in autocatalytic reaction networks, m-grammar solutions (that is, interior solutions) are rather common and correspond to the coexistence of several types in a community with mass-action-type interactions.

2.4. Example: stability of one-grammar solutions

Let us perform directly a stability analysis of a one-grammar solution,

$$\bar{x}_k = 1, \quad \bar{x}_i = 0, \quad i \neq k,$$

assuming for simplicity that $a_{kk} = 1$. We introduce a perturbation in the form $x_k = 1 + y_k e^{\Gamma t}$, $x_i = y_i e^{\Gamma t}$, $i \neq k$. Substituting into system (10) and linearizing we obtain the following eigenvalue problem:

$$\Gamma y_i = y_i(a_{ki} - 1), \quad i \neq k, \tag{30}$$

$$\Gamma y_k = -\sum_{m=1}^{n} (f_0 + a_{km}) y_m. \tag{31}$$

Because of the condition $\sum_{i=1}^{n} y_i = 0$, Eq. (31) follows from the previous equations. The resulting system can be rewritten in the form

$$\hat{K} Y = 0, \tag{32}$$

where $Y = (y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n)^{\mathrm{T}}$ and $\hat{K}$ is the $(n-1) \times (n-1)$ diagonal matrix given by

$$K_{ij} = \begin{cases} K_i \equiv a_{ki} - 1 - \Gamma, & i = j, \\ 0, & \text{otherwise}. \end{cases} \tag{33}$$

The system has non-trivial solutions if the matrix $\hat{K}$ has a zero determinant. This results in the condition

$$det\, \hat{K} = \prod_{i \neq k} K_i = 0. \tag{34}$$

The $n - 1$ expressions for the growth rate, Γ, are given by

$$\Gamma_i = a_{ki} - 1, \quad i \neq k. \tag{35}$$

We can see that under the condition of strict diagonal dominance, $a_{ki} < a_{kk}$, all the expressions for Γ are negative, and the solution under consideration is stable.

3. Imperfect learning fidelity

In this section, we will see how existence and stability properties for m-grammar solutions change for finite values of N, that is, when the learning fidelity is not perfect. In Section 3.1, we will find corrections to m-grammar solutions which come from mistakes of learning, and derive the existence conditions. In Section 3.2, we will prove that there is a region where these extended solutions are stable. In Section 3.3, we will prove the universality property, which states that the size of the stability domain is independent of the system size, n, in the limit $n \to \infty$.

3.1. Existence of m-grammar solutions

Let us study how the stable fixed points corresponding to m-grammar solutions change for finite values of N. In this case, the matrix Q is given by

$$Q_{ij} = \delta_{ij} + q_{ij}, \quad \|q\| \ll 1 \tag{36}$$

(here $\|\cdot\|$ is some matrix norm). Note that $q_{ii} \leqslant 0$, $q_{ij} \geqslant 0$ for all $i \neq j$ and $\sum_{j=1}^{n} q_{ij} = 0$ (to keep the matrix row-stochastic). Let us check if m-grammar solutions depend on the quantities q_{ij} continuously in the vicinity of $\|q\| = 0$. We start by assuming that they do, and write down their expansion in terms of q_{ij} around zero, keeping only the first-order terms

$$x_1 = \bar{x}_1 + y_1 + O(\|q\|^2), \ldots, x_m = \bar{x}_m + y_m + O(\|q\|^2), \tag{37}$$

$$x_{m+1} = y_{m+1} + O(\|q\|)^2, \ldots, x_n = y_n + O(\|q\|^2), \tag{38}$$

where $\bar{x}_j, 1 \leqslant j \leqslant m$, are the components of the m-grammar solution for $N = \infty$. We will have to check that condition

$$y_k \sim \|q\| \quad \forall k \tag{39}$$

is satisfied, otherwise our assumption of smoothness breaks down. Let us find y_k explicitly. We have

$$f_j = \bar{\phi} + \sum_{k=1}^{n} a_{jk} y_k, \quad 1 \leqslant j \leqslant m, \tag{40}$$

$$f_i = \sum_{j=1}^{m} a_{ji}\bar{x}_j + \sum_{k=1}^{n} a_{ik} y_k, \quad m+1 \leqslant i \leqslant n, \tag{41}$$

$$\phi = \bar{\phi} + \phi', \quad \phi' \equiv 2\sum_{k=1}^{n}\sum_{j=1}^{m} a_{jk}\bar{x}_j y_k. \tag{42}$$

We first calculate the correction to the average fitness, ϕ', coming from a finite probability of mutation. We can write down the first m-evolutionary equations using expressions (37), (38) and (42). They are

$$0 = \bar{x}_1 \sum_{k=1}^{n} a_{1k} y_k + \bar{\phi}(q_{11}\bar{x}_1 + \cdots + q_{m1}\bar{x}_m) - \bar{x}_1\phi', \tag{43}$$

$$0 = \bar{x}_2 \sum_{k=1}^{n} a_{2k} y_k + \bar{\phi}(q_{12}\bar{x}_1 + \cdots + q_{m2}\bar{x}_m) - \bar{x}_2\phi', \tag{44}$$

$$\ldots$$

$$0 = \bar{x}_m \sum_{k=1}^{n} a_{mk} y_k + \bar{\phi}(q_{1m}\bar{x}_1 + \cdots + q_{mm}\bar{x}_m) - \bar{x}_m\phi'. \tag{45}$$

Adding up these equations and using $\sum_{j=1}^{m} \bar{x}_j = 1$ and definition (42), we obtain the following expression:

$$\phi' = 2\bar{\phi}\sum_{j=1}^{m}\sum_{k=1}^{m} q_{jk}\bar{x}_j. \tag{46}$$

The corrections to x_i look like

$$y_i = \frac{\bar{\phi}\sum_{j=1}^{m} \bar{x}_j q_{ji}}{\bar{\phi} - \sum_{j=1}^{m} a_{ji}\bar{x}_j} = \frac{\bar{\phi}\sum_{j=1}^{m} \bar{x}_j q_{ji}}{\bar{\phi} - \bar{f}_i}, \quad i > m. \tag{47}$$

Finally, in order to calculate corrections y_k for $1 \leqslant k \leqslant m$, we use Eqs. (43)–(45). We have $\sum_{k=1}^{m} a_{ik} y_k + \sum_{j=m+1}^{n} a_{jk} y_j = \frac{\bar{\phi}}{\bar{x}_i}\sum_{j,k=1}^{m} \bar{x}_j(2q_{jk} - q_{ji})$. Let us express $y_m = -\sum_{j \neq m} y_j$, and denote $\mathbf{y}^{(m-1)} = (y_1, \ldots, y_{m-1})^{\mathrm{T}}$. The corrections $\mathbf{y}^{(m-1)}$ are obtained by solving an inhomogeneous linear system of equations with a non-singular matrix, $A'(m)$

$$\mathbf{y}^{(m-1)} = [A'(m)]^{-1}\mathbf{g}^{(m-1)}, \quad \mathbf{g}^{(m-1)} = (g_1, \ldots, g_{m-1})^{\mathrm{T}} \tag{48}$$

with $g_i = \bar{\phi}\sum_{k=1}^{m} \bar{x}_k \left(\frac{1}{\bar{x}_i}\sum_{j=1}^{m}(2q_{kj} - q_{ki}) - \sum_{j=m+1}^{n} (q_{kj}/\bar{\phi} - \bar{f}_j)\right)$.

Note that since the m-grammar solution is stable for $N \to \infty$, condition (22) is satisfied which guarantees the finiteness of the expression (47). Similarly, the matrix $A'(m)$ is negative definite (see Eq. (21)) and therefore the corrections $y_1, \ldots, y_{m-1}$ are also finite. This means that condition of smoothness (39) holds, which in turn proves the existence of the fixed point for finite N.

Now it is possible to estimate the domain of existence of m-grammar solutions. Let us concentrate on the important case $m = 1$. We have from (46) and (47)

$$\phi = (f_0 + a_{11})(1 + 2_{q11}), \quad y_i = \frac{(f_0 + a_{11})q_{1i}}{a_{11} - a_{1i}}, \quad i > 1,$$

$$y_1 = -\sum_{i=2}^{n} y_i. \tag{49}$$

Fig. 3 illustrates formula (49). The horizontal axis is N, the number of learning events, and the vertical axis is ϕ. The diamonds represent the stable fixed points of the system found numerically, and the solid lines are formula (49) with k from 1 to n. The theoretically calculated values predict the limiting behavior of the fixed points for large values of N.

Let us take for simplicity $f_0 = 0$ and $a_{11} = 1$ (the case where the diagonal elements of the matrix A are equal to one is important because it corresponds to "self-consistent" languages where the probability to understand a grammatically correct sentence is one). It is clear from formula (49) that in order for our approximation to hold we need to require

$$q_{1i} \ll (1 - a_{1i}) \tag{50}$$

and also that

$$\left|\sum_{i=2}^{n} \frac{q_{1i}}{1 - a_{1i}}\right| \ll 1. \tag{51}$$

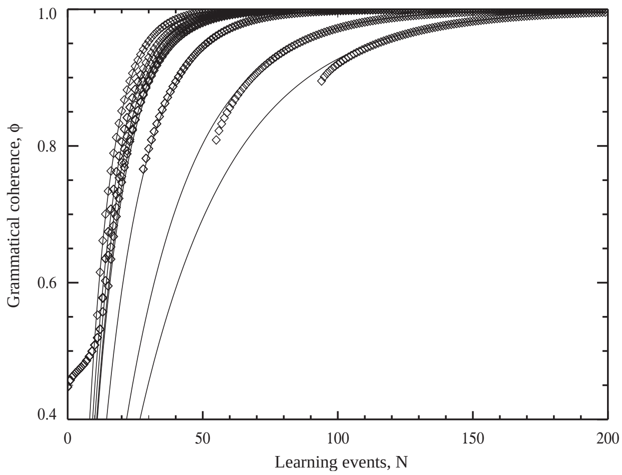

Fig. 3. The bifurcation diagram for a system with a random similarity matrix. The diamonds represent the stable fixed points of the system found numerically, and the solid lines represent formula (49) with k from 1 to n.

This guarantees the smallness of the correction y_1 in formula (49).

3.2. Stability domains

Next we will derive some stability results for the m-grammar solutions. We have the following:

Theorem 3.1. *Let us suppose that for $N = \infty$ there exists a stable m-grammar solution. Then there exists a finite value $N = N_0$ such that for all $N > N_0$ the m-grammar solution is stable.*

Proof. Let us solve the stability problem of m-grammar solutions for finite values of N (in the region of their existence). We take the exact steady-state solution (we do not have an analytical formula for it), and follow the usual steps of a linear analysis: perturb it with z_i, linearize with respect to small z_i, and obtain a system of homogeneous linear equations for z_i. The m-grammar solution is stable if all the eigenvalues of the corresponding linear matrix are negative. Next, we recall that the value $\|q\| \ll 1$ which means that the linear matrix is equal to matrix $\tilde{\mathscr{L}}$ of the stability problem, plus a correction. In other words, the new matrix $\mathscr{L}$ satisfies

$$\mathscr{L} = \tilde{\mathscr{L}} + \mathscr{N}, \quad \|\mathscr{N}\| \sim \|q\|. \tag{52}$$

The matrix $\mathscr{N}$ is expressed in terms of corrections to the m-grammar solution found in (47) and (48). We need to solve the equation

$$det(\tilde{\mathscr{L}} + \mathscr{N} - \Gamma I) = 0, \tag{53}$$

where Γ is the growth rate of the perturbation and is equal to the ith eigenvalue of $\mathscr{L}$. Let us write

$$\Gamma_i = \tilde{\Gamma}_i + \gamma_i, \quad |\gamma_i| \sim \|q\|, \quad 1 \leqslant i \leqslant m-1, \tag{54}$$

where the first approximation to the growth rate, $\tilde{\Gamma}_i$, is found to be the ith eigenvalue of the linear stability problem for $\|q\| = 0$. We assume that all $\tilde{\Gamma}_i < 0$, i.e. the m-grammar solution is stable at $N = \infty$. Linearizing Eq. (53) for different values of $\tilde{\Gamma}_i$, we obtain $m-1$ linear equations for γ_i, which yield exactly $m-1$ corrections to the growth rate.

Now in order to guarantee the stability of the m-grammar solution, we need to have

$$\tilde{\Gamma}_i + \gamma_i < 0 \quad \forall i. \tag{55}$$

Note that γ_i is a linear functional of q_{ij} and is equal to zero at $q_{ij} = 0$ thus making expressions (55) negative. It follows that in the $\mathbf{R}^{n\times n}$ space of all q_{ij}, there exists a neighborhood of the origin, where all the functions (55) are negative. In this neighborhood the m-grammar solution is stable if it exists. Since $q_{ij} \to 0$ and $N \to \infty$, we can find such $N = N_0$ that the matrix $[q_{ij}]$ belongs to the neighborhood where the m-grammar solution is stable. □

3.3. Stability of one-grammar solutions

Let us consider the case $m = 1$ more closely. We can calculate the explicit values of corrections γ_i. We have for $2 \leqslant i \leqslant n$

$$\gamma_i = a_{1i}(q_{1i} + q_{ii}) + \frac{a_{11}q_{1i}(a_{ii} - 2a_{1i})}{a_{11} - a_{1i}} + a_{11}\sum_{k=2}^{n} \frac{q_{1k}(a_{ik} - 2a_{1k} - a_{1i} + 2a_{11})}{a_{1i} - a_{1k}}. \tag{56}$$

For stability we need conditions (55) with $\tilde{\Gamma}_i = -(1 - a_{1i})$, see Eq. (35). Let us define regions χ_i in the space $(\mathbf{R}^+)^{(n-1)}$ of all $q_{1i}, 2 \leqslant i \leqslant n$

$$\chi_i = \{q_{1j} > 0 : \gamma_i - a_{11} + a_{1i} < 0\} \tag{57}$$

(this set depends on the quantities $q_{ii}, 2 \leqslant i \leqslant n$, see formula (56)). Outside the set $\chi = \bigcap_{i=2}^{n} \chi_i$ the one-grammar solution is unstable. The boundaries of the set χ indicate the boundaries of stability (up to higher order terms).

We are going to use inequalities (56) to find a sufficient condition of stability of 1-grammar solutions. Let us rewrite conditions (57) with the simplification $a_{ii} = 1$

$$q_{1i}\left(1 - \frac{a_{1i}^2}{1 - a_{li}}\right) + \sum_{k=2}^{n} \frac{q_{1k}}{1 - a_{1k}} \times (a_{ik} - 2a_{1k} - a_{1i} + 2) < 1 - a_{1i} + a_{1i}|q_{ii}|. \tag{58}$$

It is convenient to use the notation

$$b_i = 1 - a_{1i}, \quad \xi_{ik} = 1 - a_{ik}. \tag{59}$$

We have

$$q_{1i}\left(1-\frac{(1-b_i)^2}{b_i}\right)+2\sum_{k=2}^{n} q_{1k} + \sum_{k=2}^{n} q_{1k}\frac{b_i-\xi_{ik}}{b_k} < b_i + (1-b_i)|q_{ii}|. \tag{60}$$

These inequalities contain linear functions of q_{1i} and thus define $n-1$-dimensional hyperplanes in $(\mathbf{R}^+)^{n-1}$. The set χ_i is the half-space defined by the ith inequality that contains the origin, restricted to the positive "quadrant". The domain of stability of the one-grammar solution with G_1 the dominant grammar is the intersection of all χ_i.

3.3.1. Fully symmetrical systems

Let us consider the case where $a_{ij}=a$, $i\neq j$, and $a_{ii}=1$. At this point we do not make any assumptions on the matrix Q, i.e. it is not necessarily of the form (5).

First, we note that if $b_i \equiv b = 1-a$ for $2\leqslant i\leqslant n$, inequalities (60) follow from the following single inequality:

$$\frac{4}{b}\sum_{k=2}^{n} q_{1k} < 1. \tag{61}$$

This proves the following

Theorem 3.2. *For fully symmetrical systems with $a_{ij}=a$ for $i\neq j$, and $a_{ii}=1$, there exists a ball, $\mathcal{B}$, of radius $\mathcal{R}=(1-a)/4$ in the space of all $q_{1i}, 2\leqslant i\leqslant n$, with the l_1 metric, centered at the origin, such that for any $q_{12},\ldots,q_{1n}$ inside $\mathcal{B}$, the 1-grammar solution is stable.*

Note that the radius $\mathcal{R}$ does *not* depend on the size of the system, n. Since $|q_{11}|=\sum_{k=2}^{n} q_{1k}$, we have a sufficient condition for stability

$$|q_{11}|<(1-a)/4. \tag{62}$$

Let us compare this result with the exact calculation for a fully symmetrical learner defined by (5). We can see that condition (62) is stronger than condition (6), and they coincide in the limit $a\to 1$.

3.3.2. Random coefficients

Next, we assume that the coefficients b_i are random variables. In this case, a statement like Theorem 3.2 is hardly possible, and the question we pose must change slightly. Namely, a "universal" ball inside which stability is guaranteed for *any* realization of the coefficients, b_i, might not exist (or, more precisely, its radius might be $\mathcal{R}=0$). However, for a typical realization we can hope to find a ball of stability which has a finite size.

Let us assume that $a_{ii}=1$ for all i and a_{ij} are distributed between zero and one. Further, we assume that the distribution has a mean and a variance, and also the mean of the quantity $(1-a_{ij})^{-1}$ exists for $i\neq j$.

Theorem 3.3 (Universality property). *There exists a constant, $q_c>0$, such that if $|q_{ii}|<q_c$ for all i, then for a typical realization of a_{ij} at least one of the n possible one-grammar solutions is stable. Moreover,*

$$\lim_{n\to\infty} q_c > 0.$$

Proof. Let us assume without loss of generality that

$$|q_{11}|<|q_{ii}|, \quad 1<i\leqslant n. \tag{63}$$

This means that in some sense, the grammar whose stability we are considering is easier to learn than the others. In what follows, we will show that the stability threshold for such "easiest" grammar does *not* depend on the number n. Next, let us set

$$q_{1i}=q\varkappa_i, \quad 1<i\leqslant n, \quad \sum_{j=2}^{n}\varkappa_j=1, \quad |q_{11}|=q. \tag{64}$$

Using assumption (63), we obtain from (60):

$$q\left[\varkappa_i\left(1-\frac{(1-b_i)^2}{b_i}\right)+1+b_i+\sum_{k=2}^{n}\varkappa_k\frac{b_i-\xi_{ik}}{b_k}\right]<b_i, \quad 1<i\leqslant n. \tag{65}$$

This inequality must be satisfied for all i. Note that the coefficient multiplying $\varkappa_i$ is negative if

$$b_i<b_* \equiv \frac{3-\sqrt{5}}{2}\approx 0.3820. \tag{66}$$

Let us define

$$q_c(b_i)=\begin{cases}\frac{b_i}{1+b_i+\sum_{k=2}^{n}\varkappa_k(b_i-\xi_{ik})/b_k}, & 0\leqslant b_i\leqslant b_*,\\ \frac{b_*}{2+b_i+\sum_{k=2}^{n}\varkappa_k(b_i-\xi_{ik})/b_k}, & b_*\leqslant b_i\leqslant 1,\end{cases} \tag{67}$$

Then the conditions on q can be written as

$$0<q<q_c(b_i) \quad \text{if } q_c(b_i)>0, \tag{68}$$

$$q>0 \quad \text{otherwise} \tag{69}$$

(if $q_c(b_i)$ is negative, we do not need to impose any condition on q, because for such b_i condition (65) is satisfied for any $q>0$).

Let us fix i and calculate the expected value of $q_c(b_i)$. We will assume that the variable

$$y_k \equiv \frac{b_i-\xi_{ik}}{b_k} \tag{70}$$

is distributed with density f_y and has the mean $\mu(y)$ and the variance $\sigma^2(y)$. By an analogue of the central limit theorem, the distribution of $S_n \equiv \sum_k y_k\varkappa_k$ tends to

$$f_S(x)=\frac{\sqrt{n}}{\sqrt{2\pi}\sigma'(y)}\exp\left[-\frac{(x-\mu(y))^2 n}{2(\sigma'(y))^2}\right], \tag{71}$$

where

$$\sigma'(y) = \sigma(y)\sqrt{n\sum_j \varkappa_j^2}, \tag{72}$$

here and below we ignore the difference between $n-1$ and n because we assume that n is large. The expected value of $q_c(b_i)$ can now be calculated

$$\mathbf{E}(q_c(b_i)) = \frac{b_i}{1+b_i+\mu(y)}\left(1+O\left(\frac{1}{n}\right)\right), \quad b_i < b_*. \tag{73}$$

Similarly, we get

$$\mathbf{E}(q_c(b_i)) = \frac{b_*}{2+b_i+\mu(y)}\left(1+O\left(\frac{1}{n}\right)\right), \quad b_i > b_*. \tag{74}$$

The threshold value of q can be found as

$$q_c = \min_{b_i}[\mathbf{E}(q_c(b_i)) \mid \mathbf{E}(q_c(b_i)) \geqslant 0]. \tag{75}$$

Let us consider the quantity $\mu(y)$. We have

$$\mu(y) = b_i\mathbf{E}\left(\frac{1}{b_k}\right) - \mathbf{E}\left(\frac{\xi_{ik}}{b_k}\right), \tag{76}$$

i.e. it is a linear function of b_i. Let us first look at expression (73). If the denominator was always positive, then its minimum would be achieved at $b_i = 0$ and equal 0, which means that the coherence threshold would depend on n. However, if the denominator was negative at zero, then we would have a constant threshold value of q_c. We have at $b_i = 0$

$$\mathbf{E}(q_c(0)) = \frac{b_i}{1-\mathbf{E}(\xi_{ik}/b_k)}. \tag{77}$$

The random variable b_k and ξ_{ik} come from the same distribution, f_b. It is easy to show that

$$\mathbf{E}\left(\frac{\xi_{ik}}{b_k}\right) = \int x f_b(x)\,\mathrm{d}x \int \frac{f_b(y)}{y}\,\mathrm{d}y = \mathbf{E}(b_k)\mathbf{E}(1/b_k). \tag{78}$$

The latter quantity is always bigger than 1 which means that the denominator in (77) is negative. As b_i increases, the denominator may become positive at some finite value of b_i. The function $\mathbf{E}(q_c(b_i))$ is always a decreasing function of b_i, so its minimum is reached at $b_i = b_*$ and equals $b_*/(1+b_*(1+\mathbf{E}(1/b_k)) - \mathbf{E}(\xi_{ik}/b_k))$. The minimum of expression (74) is achieved at $b_i = 1$ and equals $b_*/(3+\mathbf{E}(1/b_k) - \mathbf{E}(\xi_{ik}/b_k))$. Therefore, by taking the smaller of these two expressions, we obtain the following estimate:

$$q_c = \frac{3-\sqrt{5}}{2[3+\mathbf{E}(1/b_k)(1-\mathbf{E}(b_k))]}. \tag{79}$$

Assuming that $\mathbf{E}(1/b_k)$ and $\mathbf{E}(b_k)$ exist, we conclude that q_c is a finite constant which does not depend on n. □

4. Applications, examples, conclusions

We have studied the behavior of the "coherence threshold", the critical value, q_c, of the learning accuracy which corresponds to the emergence of "localized", or "coherent" solutions. For learning accuracy higher than critical, the system has stable fixed points corresponding to the majority of the population speaking the same (dominant) language. In this paper, we concentrated on the limiting behavior of the threshold as the size of the system (n, the possible number of languages) becomes high.

The phase portrait of the system is defined by the elements of the $n \times n$ matrix, A, whose entries indicate how *different* the languages are from each other. It is not possible to find a universal threshold value of the accuracy which would work for any matrix A. Therefore, we formulated the problem where the entries of the A matrix are random numbers generated by a certain distribution. Then the definition of the coherence threshold, q_c, can be adjusted such that it guarantees the existence and stability of a localized solution for a *typical* realization of this system. We have been able to show that q_c tends to a non-zero constant as n tends to infinity.

If the learning accuracy matrix, Q, can be parameterized in such a way that it satisfies (9), then by controlling its diagonal entries $(1-q_{ii})$ by q_c we can guarantee coherence for a typical realization of the parameters. In the next sections, we will work out an example where the learning accuracy matrix, Q, can be varied parametrically in a natural way.

4.1. The batch learner as an example of parameterization

Recall that the elements Q_{ij} of the learning accuracy matrix, Q, give probability that a child learning the language from a teacher with language G_i will end up speaking language G_j. There are many ways how the process of individual learning can be modelled. The typical learning theory setting is as follows. We have an ideal speaker–hearer pair. The speaker (the teacher) formulates sentences using a specific grammar, the hearer (the learner) receives these sentences and has to infer the underlying grammar. There are various algorithms that the learner could use, which differ in efficiency and memory requirements. Here, we will consider the so-called *batch learner* (Komarova et al., 2001; Nowak et al., 2002). It works in the following way. The learner starts by (randomly) choosing one of the n grammars as an initial state. He then receives N sentences from the teacher and evaluates them for consistence. The learner sticks with his current hypothesis until the input unambiguously suggests that *all* the sentences come from one grammar, which is *not* the current guess. At this point the learner switches to the correct grammar.

Note that the batch learner algorithm is not an approximation to the actual, yet unknown mechanism of language acquisition used by children. In this study, we use the batch algorithm as an example to demonstrate, how exactly the method developed here can be applied. In fact, the universality property will hold for *any* learning mechanism which satisfies the mild conditions specified in Section 3.

Suppose that the teacher's grammar is G_k. We assume that the probability for a sentence from G_k to belong to G_i is given by s_{ki}. We further assume that the probability for a sentence of G_k to belong to G_i and G_j *simultaneously* is given by $s_{ki}s_{kj}$, and similarly with intersections of a higher number of grammars. This is the independence assumption introduced by Rivin (2001).

The probability that after N sentences the input will unambiguously point toward grammar G_k (and no other grammar) is given by

$$\pi_N^{(k)} = \prod_{i \neq k} (1 - s_{ki}^N).$$

The transition matrix after N sentences looks like

$$T_{ij}^{(k)}(N) = \begin{cases} 1, & i = j = k, \\ 0, & i = k,\ j \neq k, \\ \pi_N^{(k)}, & i \neq k,\ j = k, \\ 1 - \pi_N^{(k)}, & i \neq k,\ j = i, \\ 0, & \text{otherwise}. \end{cases}$$

Therefore, we have

$$Q_{kj} = [(\mathbf{p}^{(0)})^{\mathrm{T}}(T^{(k)})(N)]_j.$$

This gives us the following parameterization of the learning accuracy matrix with the parameter N, the number of sampling sentences:

$$Q_{kk} = 1 - \frac{n-1}{n}(1 - \pi_N^{(k)}), \quad Q_{kj} = \frac{1 - \pi_N^{(k)}}{n}.$$

Let us denote $s_k = (1 - \pi_N^{(k)})(n-1)/n$. Then we have

$$q_{kj} = \frac{s_k}{n-1}, \quad q_{kk} = s_k.$$

This gives us the natural parameterization of the learning accuracy matrix, Q, with the number of sampling events, in the case of the batch learner algorithm. Obviously, $\lim_{N \to \infty} \pi_N^{(k)} = 1$, so that

$$\lim_{N \to \infty} Q_{kk} = 1,$$

consistent with (9).

4.2. *Two-way communication—example of a distribution*

Let us suppose that the quantities $a_{ij} = a_{ji}$ come from symmetrizing the one-way communicability functions, so that $a_{ij} = (s_{ij} + s_{ji})/2$. Further, we will assume that s_{ij} and s_{ij}, $1 \leqslant i, j \leqslant n$, $i \neq j$ are independent random variables drown from a uniform distribution between zero and one.

Denoting $b_i = 1 - a_{1i}$, we obtain that the distribution function of b_i is given by

$$f_b(x) = \begin{cases} 4x, & 0 \leqslant x \leqslant 1/2, \\ 4(1-x), & 1/2 \leqslant x \leqslant 1. \end{cases} \tag{80}$$

First, we note that

$$\mathbf{E}(b_i) = \tfrac{1}{2}, \quad \mathbf{F}(1/b_i) = 4 \log 2 \approx 2.77. \tag{81}$$

Let us find the region of existence and stability of a coherent solution for a typical realization of the matrix A described above, in the scenario of the batch learning algorithm. Applying formula (81) together with Eq. (79) we obtain $q_c \approx 0.087$. In order to guarantee that our method works, we also need the applicability conditions, (50) and (51). We need to require

$$\frac{s_k}{n-1} < 1 - a_{kj} \quad \forall j.$$

This condition is easy to satisfy for large enough n because $\mathbf{E}\,min(1 - a_{kj}) \sim \frac{1}{\sqrt{n}}$, see appendix.

Appendix A. Statistics of a_{ik}

Let us calculate the value of $\mathbf{E}\,min(1 - a_{kj})$, where the entries of the matrix A, a_{kj}, are random numbers drawn from distribution (80), and G_k is the teacher's grammar. We will use the notation $b_i \equiv 1 - a_{ki}$. We have

$$P(\min_{i=1}^{n} b_i > \alpha) = \begin{cases} (1 - 2\alpha^2)^n, & 0 \leqslant \alpha \leqslant 1/2, \\ 2^n(1-\alpha)^{2n}, & 1/2 \leqslant \alpha \leqslant 1. \end{cases} \tag{A.1}$$

We have

$$\mathbf{E}(\min_i b_i) = -\int_0^{1/2} \alpha \frac{\mathrm{d}}{\mathrm{d}\alpha}(1 - 2\alpha^2)^n \, \mathrm{d}\alpha - \int_{1/2}^{1} \alpha \frac{\mathrm{d}}{\mathrm{d}\alpha} 2^n (1-\alpha)^{2n} \, \mathrm{d}\alpha. \tag{A.2}$$

The second integral can be evaluated to give $1 + n/(2^n(1+2n))$, i.e. its contribution is exponentially small. The first integral can be taken by parts, yielding

$$1/2^n + \int_0^{1/2} (1 - 2\alpha^2)^n \, \mathrm{d}\alpha. \tag{A.3}$$

The boundary contribution (the first term in the expression above) is exponentially small above. We will now estimate the remaining part, $\int_0^{1/2} (1 - 2\alpha^2)^n \, \mathrm{d}\alpha$.

First, we note that for $\alpha \ll 1/\sqrt{2}$, the function under the integral can be approximated as

$$(1 - 2\alpha^2)^n \approx \mathrm{e}^{-2n\alpha^2}. \tag{A.4}$$

Let us introduce $\alpha_1 = \sqrt{\log n / 2n}$. The function under the integral is equal to $1/n$ at $\alpha = \alpha_1$. Since it is a monotonically decaying function of α, we can bound the integral

$$\int_{\alpha_1}^{1/2} (1 - 2\alpha^2)^n \, d\alpha < \frac{1}{2n}. \qquad (A.5)$$

Next, we note that for $\alpha < \alpha_1$, the condition $\alpha \ll 1/\sqrt{2}$ holds and we can use approximation (A.4) to estimate

$$\int_0^{\alpha_1} (1 - 2\alpha^2)^n \, d\alpha \approx \int_0^{\alpha_1} e^{-2n\alpha^2} d\alpha = \frac{1}{\sqrt{2n}} \int_0^{\sqrt{\log n}} e^{-y^2} \, dy, \qquad (A.6)$$

where we introduced the variable $y = \sqrt{2n}\alpha$. The last integral is easy to estimate. We have

$$\mathbf{E}\left(\min_i b_i\right) = \frac{1}{\sqrt{2n}}\left[\sqrt{\frac{\pi}{2}} + O\left(\frac{1}{n\sqrt{\log n}}\right)\right]. \qquad (A.7)$$

References

Bürger, R., 1998. Mathematical principles of mutation–selection models. Genetica 102–103, 279–298.

Eigen, M., Schuster, P., 1979. The Hypercycle. A Principle of Natural Self-Organisation. Springer, Berlin.

Hadeler, K.P., 1981. Stable polymorphisms in a selection model with mutation. SIAM J. Appl. Math. 41, 1–7.

Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Replicator Dynamics. Cambridge University Press, Cambridge.

Kingman, J.F.C., 1961. A mathematical problem in population genetics. Math. Proc. Cambridge Philos. Soc. 57, 574–582.

Komarova, N.L., Niyogi, P., Nowak, M.A., 2001. Evolutionary dynamics of grammar acquisition. J. Theor. Biol. 209 (1), 43–59.

Levin, S.A., 2002. Complex adaptive systems: exploring the known, the unknown and the unknownable. Bull. Amer. Math. Soc. 40 (1), 3–19.

Niyogi, P., 1998. The Informational Complexity of Learning. Kluwer, Boston.

Nowak, M.A., Komarova, N.L., Niyogi, P., 2001. Evolution of universal grammar. Science 291, 114–118.

Nowak, M.A., Komarova, N.L., Niyogi, P., 2002. Computational and evolutionary aspects of language. Nature 417, 611–617.

Page, K.M., Nowak, M.A., 2002. Unifying evolutionary dynamics. J. Theor. Biol. 219, 93–98.

Price, G.R., 1970. Selection and covariance. Nature 227, 520–521.

Rivin, I., 2001. Yet another zeta function and learning. arXiv.org preprint cs.LG/0107033.

Sasaki, A., Nowak, M.A., 2003. Spatial language diversity. Preprint.

Solan, Z., Ruppin, E., Horn, D., Edelman, S., 2002. Evolution of language diversity: the survival of the fitness. Extended Abstract of a Talk Presented at the Fourth International Conference on the Evolution of Language, Cambridge, MA, March 2002.

Stadler, P.F., Schuster, P., 1992. Mutation in autocatalytic reaction networks—an analysis based on perturbation theory. J. Math. Biol. 30, 597–632.