Cognitive Cultural Dynamics

Harold P. de Vladar
Centre for Cooperation and Conflict in Evolutionary Systems. Institute for Advanced Studies Kőszeg - iASK.
Chernel Ut. 14.
H-9730, Kőszeg, Hungary
Centre for the Conceptual Foundations of Science. Parmenides Foundation.
Kirchplatz 1
82049, Pullach, Germany
Harold.Vladar@parmenides-foundation.org

ABSTRACT

Based on previous results on language games here I study cultural dynamics extended in spatial environments. The underlying model makes assumptions regarding cognitive aspects of the individuals based on the Neuronal Replicator hypothesis. Although I assume a simple and minimal version of cultures, this model allows exploring the effects of idiosyncratic as well as externally, environmentally, imposed preferences on cultural traits. I also study the case of dispersal of individuals and find that this factor is key for the rapid spread of cultural traits.

CCS CONCEPTS

•Computing methodologies → Cognitive science; Multi-agent systems; Modeling and simulation; •Applied computing → *Anthropology;* Life and medical sciences;

KEYWORDS

Evolution, culture, language, neuronal replicators, cognition, stochastic models, spatial systems

ACM Reference format:
Harold P. de Vladar. 2017. Cognitive Cultural Dynamics. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 7 pages.
DOI: http://dx.doi.org/10.1145/3067695.3082464

1 INTRODUCTION: CULTURES AND MULTI-CULTURES

Evolutionary biologists have long recognised that culture evolves in a way analogous to biological evolution [2]. However, the mathematisation of cultural evolution has remained elusive partly because (a) the mechanisms of cultural transmission have remained largely inaccessible to evolutionary biologists and (b) the units that are evolvable have been hard to identify.

A central question in cultural evolution is: *how to model culture* [2, 10, 11]. But the complication is twofold becaiuse the definition of what *culture* is remains a subject of constant debate [8, 9]. Although culture is a multifaceted human trait that can be hardly fully described mathematically, it is feasible to adopt working definitions that will facilitate understanding some aspects of its dynamics. The sole identification of what aspects are prone to be modelled by using relatively simple formal techniques of mathematics or computer science is already an advance.

In my proposed framework I employ a broad and simple definition of culture: *a culture is the set of concepts that individuals share in a population.* While I agree that this definition is open to discussion, I will adhere to it with the aim of modelling how concepts originate and spread in a population.

Unlike problems in other fields such as statistical physics and evolutionary genetics, it is not obvious what aspects to choose to be modelled and which are the units of evolution. But once having committed to a framework to describe and study cultural evolution in the space of concepts and tags (see below) I ask:

(1) Do some cognitive aspects limit or facilitate multicultural mixability?
(2) What limits the speed of spatial spread of cultures: population diffusion or efficiency of cultural transmission?

I make a distinction between "implied concepts" and their expressions through actions, utterances or tags [11, 14]. Implied concepts, I assume, remain mental and abstract to each individual or agent. Actions or 'tags', are explicit forms that are transmissible across agents (i.e. can be copied) and make reference to the actual, implied meaning. While the tags are transmissible, the implied concepts have to be developed and implied by each individual.

However, culture is changeable, partly because different social groups have different cultures, and across-group communication brings novelty to the populations [8]. In our contemporary highly mobile societies populations are constituted by individuals of different cultures. What are the features of these multi-cultural societies? The approach that I present assumes a population sub-structuring where part of it has one (sub)set of concepts that is disjoint from the (sub)set of concept of the other part. By assuming basic aspects of cognitive processes that are universal to humans based on the Neuronal Replicator hypothesis, thereby I take the first steps in the construction, description and implementation of cultural dynamics. In this first work I use a space of only two concepts. With this model I draw implications and consequences for multi-cultures to provide partial answer to questions 1 and 2 above.

GECCO '17 Companion, Berlin, Germany

DOI: http://dx.doi.org/10.1145/3067695.3082464

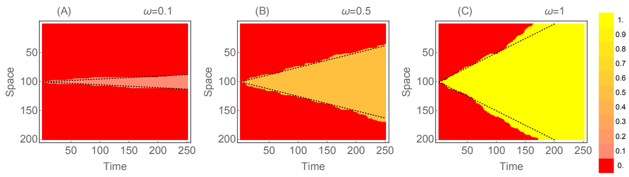

Figure 1: Spread of a new tag in a homogeneous population. (A) Low preference, $\omega = 0.01$. (B) No preference, $\omega = 1/2$. (C) High preference $\omega = 1$. The original tag σ_1 is in red, the coloured region indicates where the new tag σ_2 is known but used with certain preference indicated in the colour bar.

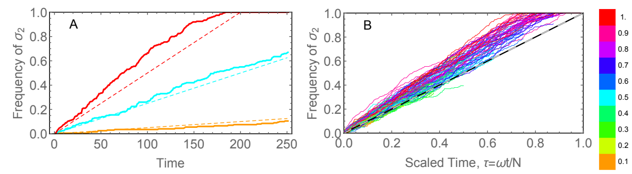

Figure 2: Change in frequency of the alternative tag σ_2 in the population. (A) Different preference values. Solid lines: simulation average; dashed lines: theoretical prediction. (B) Dynamics with time scaled by preference and population size. Dashed black-white: theoertical expectation (identity line). The bar indicates the preference values for both panels.

2 LANGUAGE GAMES AS A FRAMEWORK FOR CULTURAL EVOLUTION

Based on previous frameworks of artificial intelligence originally devised to study language [11] and advances in Darwinian neurodynamics [4, 12] I propose a way to study cultural dynamics in terms of evolutionary concepts and of neurobiological mechanisms.

The approach consists of, first, identifying a sensible way to describe what culture is. Second, it requires an implementation of cognitive processes that allows copying certain patterns across agents or individuals and making an inference of what these patterns mean. Third, the agents must be capable of choosing, idiosyncratically, what to transmit to others. This choice must be cognitively based.

The details of the architecture will be published somewhere else [5]. Here, we use a mathematical model and simulations based on Markov chains which is derived from general properties of cognition based on the Neuronal Replicator hypothesis (NRH) described in the next section.

This paper presents advances regarding cultural dynamics by using simplified models of culture. These models consist of a population of agents that interact pair-wise under a variety of conditions. The differences between the general framework and the simple model are that, in the former, we aim at dealing with a relatively complex models of culture, with a variety of features and where different populations can interact. In contrast, in the simpler version, culture is represented only by a set of two possible contrasting concepts, and the interactions are only pairwise across individuals.

With the description of the general framework we are able to advance some important analogies with evolution (e.g. the Wright-Fisher model which is central in evolutionary genetics [3, see Appendix A]). With this general model we derive the simpler one which we implement to study some interesting scenarios.

3 THE NEURONAL REPLICATOR HYPOTHESIS

Building up the original Neuronal Replicator Hypothesis (NRH) [6] we have developed a cognitive framework for problem solving. In our framework, hypotheses or candidate solutions to a problem play the role of evolutionary units: they are selected based on their fitness just like in evolution and also multiply with heredity (though learning) and introducing variation, virtually becoming an evolutionary search. In last year's GECCO '16 we presented an instance of said cognitive architecture [4]. In that work we showe how the synergy between learning and selection can find optimal patterns in a roughed fitness landscape [12]. Now, I make use of these ideas an apply them as a basic milieux to study cultural transmission across agents. The central goal is to implement the NRH in language games as a cognitive framework in the agents to construct, store and retrieve concepts and forms.

However, a full implementation of Language Games in terms of the NRH is in progress and will be published elsewhere [5]. The mathematical model presented in this paper derives directly from 2-player naming games that implement this NRH. This model properly describes the dynamics for competing tags, constituting an ideal tool to explore cognitive cultural dynamics and state precise hypotheses that are more cumbersome to implement with other AI platforms and embodied agents. However, ultimately the idea is to achieve such computational implementation.

The connection to the NRH here is as follows. In the full implementation, each agent is initialised with one of two different competing tags, and also with a larger concept space that is not explicitly employed for anything else. (This is important since the agents need to discover that the alternative tag can refer to the same concept as their cognate one.) The preference to chose a given tag σ_1 over another σ_2 is given by a scoring mechanism that weights the success of having used each tag. These preferences are defined as the probability of choosing a particular tag given the agent's knowledge for other tags i.e. $\omega_i = \Pr[\sigma_i|\{\sigma_1, \ldots, \sigma_k\}]$. In other words, it is the probability that an agent decides to transmit tag σ_i, amongst its alternative possibilities. Appendix B gives more details about the preference scoring system and argue that, in many cases, these can be constant, result we apply in this work.

4 RESULTS: A SIMPLE MODEL FOR CULTURAL EVOLUTION

In this section I assume a population of N individuals that is structured in space; in this case, along a line. Each individual can interact as a hearer or as a speaker only with any of its two neighbours. The update is done in an asynchronous way, that is, the order in which individuals interact is randomly chosed, but making sure that each pair does so only once. Another assumption is that there is only one cultural attribute for which there are two instantiations of referents represented by the tags σ_1, σ_2 with preferences $1 - \omega$ and ω respectively. These preferences are individually based and constitute an important link with the NRH and/or other cognitive strategies. Therefore, this preference can change across individuals

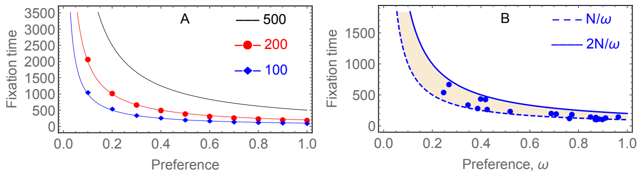

Figure 3: (A) Mean time to fix a new concept as a function of its preference ω for different population sizes. (B) If initial conditions are arbitrary, the fixation time falls, as predicted between the bounds. In this example $N = 100$. Lines: Theory; symbols: numerical averages over 100 randomised simulations with the same initial conditions.

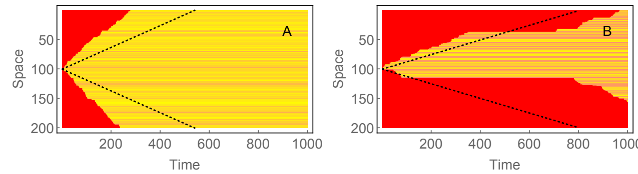

Figure 4: (A) Preferences are distributed as $\beta(4, 1.5)$, which places most of the mass towards large values; small values are very unlikely. (B) Preferences are distributed as $U(0, 1)$, which assigns equal probability to any preference. The black dashed lines are the expected trajectory according to the mean preference value. Otherwise as in Fig. 1.

and/or environments. In each iteration, the agents exchange concepts by uttering σ_1 or σ_2 according to their preference. In some simulations I allow for random spatial movement. This is implemented by randomly choosing pairs (not necessarily the same pairs that interacted) and swapping their position. This swapping occurs only after all interactions in one round have taken place.

In the Appendices A and C I explain in detail the algorithm with which the population is iterated; in this part I concentrate on the results.

Spread of preferred concepts. One of the simplest scenarios is that where almost all individuals express a pre-established concept, say σ_1, and one individual expresses an alternative concept σ_2. When will σ_2 become fixed in the population? Note that I do not allow for forgetting, so individuals will still be able to express σ_1 even if they would prefer σ_2. Here I assume that the preferences are fixed. This assumption is clearly an idealisation since preferences are themselves evolvable. Nevertheless, keeping preferences fixed will allow us to understand in a simple way the kind of dynamics that result, which will facilitate a better understanding of the more complex and realistic case of evolving preferences.

Figure 1 shows the spread of σ_2 in the population under different preference values. In this plot, every horizontal cut is the state of the population at a given time, with each colour representing the likelihood of expressing of each individual. There are three states of an individual according to their knowledge: they can know only one concept or both. However, even if they know both, the preference ω dictates the likelihood of expressing σ_2 over σ_1, as indicated in the colouring in Fig. 1.

We can appreciate that the higher the preference is, the larger the angle between the boundaries. Also, we see that the spread of σ_2 is linear with time. It can be shown (see Appendix A) that spread of the σ_2 is given by

$$x_{\pm}(t) = (N - x_0) \pm \frac{\omega}{2} t \,, \tag{1}$$

where $x_{\pm}$ describe the $(+)$ upper and $(-)$ lower boundaries of the region where σ_2 has been adopted. Note that the slope is $\omega/2$; thus, the angle between the two boundaries increases with preference.

The frequency of σ_2, denoted by p (Fig. 2A) follows directly from this relationship and it gives

$$p = \frac{\omega}{N} t \,. \tag{2}$$

Clearly, p is bounded at time τ_f which is the time for σ_2 to fix in the population. This formula implies that if time is scaled with ω/N all the dynamics will fall over a straight line of slope 1. Figure 2 shows that this is approximately true. Deviations are due to the stochastic nature of the process.

From the formula of the linear spread the mean time to fixation of σ_2 can be estimated. That is, the time τ_f when $p = 1$:

$$\tau_f = \frac{N}{\omega} \,. \tag{3}$$

Figure 3A compares this prediction with some simulations. Although in the figures we assume that the new concept is introduced by an individual at the centre of the space, this need not be the case. The formulas are rather similar and although fixation time increases slightly, it can be shown that, independently of where the new concept is introduced, $\tau_f \leq \frac{2N}{\omega}$. Moreover, it is possible to show that, irrespective of where the new concept is introduced, the expected time to fixation lies between these two boundaries:

$$\frac{N}{\omega} \leq \tau_f \leq \frac{2N}{\omega}. \tag{4}$$

Figure 3B shows some results on this bound.

Summarising, in these analyses I have characterised how the preferences affect the mixing of a simple culture. The central finding is that the time that is expected to achieve a mix (i.e. the fixation of a new concept) decreases with the inverse of the preference. If the individuals have a strong preference, then it takes N units of time, which is the number of interactions required for each individual to adopt the new concept. If preferences are intermediate, then it takes longer, but (except if $\omega = 0$), fixation is always reached in linear time.

I have analysed only the case where only one individual introduces a new concept; it could be that several individuals introduce it at the same time, that there are differential preferences, and/or that individuals can move across the space. The results of the next sections reveal that these factors affect the speed of the spread.

Heterogeneous preferences. Now I allow each individual to have a different preference which remains constant in time but introduces spatial heterogeneity. In general lines, the adoption of the new concept spreads according to the average preference (e.g. Fig. 4A). There will be higher variance on the distribution of fixation times, though. However, this is approximation is best when the distribution of preferences does not have extremely small values.

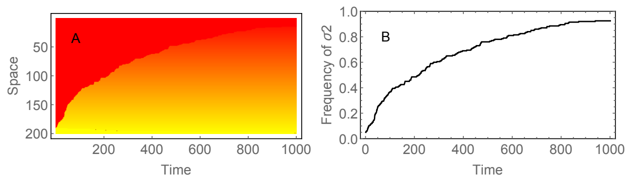

Figure 5: Spread of a new concept in a cline of preferences. The cline ranges linearly from $\omega = 0$ at $x = 0$ to $\omega = 1$ at at $x = 200$. (A) spatial spread of σ_2. Otherwise as in Fig. 1. (B) Population frequency of concept σ_2.

However, if there are individuals with very low preferences, they will act as a barrier and stop the spread (e.g. Fig. 4B). Actually, in the long run, the average still holds (unless a preference that is infinitely small occurs), but as the distribution of preferences is more spread, the variance of fixation time increases (Fig. 6).

Clines of preference. Another perspective that can be addressed is that when the heterogeneity in preference is not dictated individually but is externally fixed in a gradient. This mimics a geographic situation where there are smooth changes in concepts not because idiosyncratic preferences but rather because of externally set cultural or environmental predispositions.

In addition, I will assume that the concept σ_2 exists close to one of the borders of the cline. This is not a necessary assumption, since it does not affect the results, but is a seemingly convenient initial condition to assume.

Figure 5 reveals that under these conditions the spread is sublinear. Whether fixation finally occurs or not depends on the cline values. In the example of Fig. 5 it spans, linearly, from 0 to 1, therefore as the new concept progresses it encounters more and more resistance for further spread. In this case it cannot, strictly speaking, fix, but only asymptotically increases to $\rho = 1$. That is because the preference at one of the boundaries is zero, absolutely rejecting the new concept. However if the cline does not go all the way to zero, the concept σ_2 can of course fully spread.

A subtle aspect is that σ_1 will also spread into the sub-population that only had σ_2, even though this 'deme' will be unlikely to use the concept σ_1.

Dispersing populations. As a final experiment I study what happens when individuals disperse. This is a crucial aspect of this research because human populations are inherently mobile. Dispersion happens at different time scales, all of which are important for cultural and human evolution. For now I focus on dispersal within structured populations which mimics within-lifetime migrations, e.g. Silk Route journeys carrying new artefacts and ideas.

Figure 7 shows that dispersal has a dramatic effect on the fixation of new ideas. The overall effect is that dispersal increments the speed of fixation. Without dispersal, spread of the ideas occurs only at the cultural interface, which in the model above is always driven by at most two individuals –one at each interface.

In this model dispersal is implemented by randomly swapping the positions of individuals. Therefore individuals that are at the interface can be moved away from it. Thus individuals that carry the concept σ_2 are sometimes placed beyond the cultural interface.

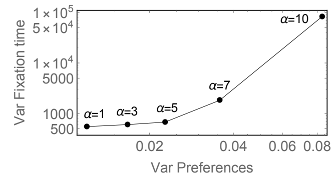

Figure 6: The variance in fixation time increases with the variance of the preferences (in log-log scale). Average preference is set to 1/2. Each point computed from 100 simulations. Preferences drawn from a $\beta(\alpha, \alpha)$ distribution, $\alpha = 1, 3, 5, 7, 10$. Population size $N = 100$. σ_2 is introduced at position $x_o = 50$.

At the same time individuals that were beyond the interface are placed inside the range where σ_2 has been adopted.

As a consequence, there are more than two agents that are effectively spreading the new concept. How many individuals scatter the interface, depends on the amount of dispersal 7A-C.

The effect of dispersal is stronger for populations with lower preferences (Fig. 7D), which suggests that mixability is facilitated by mobility.

5 DISCUSSION

One of the seminal contributions in cultural evolution [2] was the prediction of the rate of spread of an innovation. In the original proposal Fisher waves were predicted and used to study the spread of cultural traits mapped with objects and languages in the earlier periods of the cultural history of our species. However, in that approach the speed of spread (a traveling wave) [1] is a parameter set in the model. The model here presented is consistent with these results and it is possible to derive a traveling wave from it (unfortunately space limitations do not allow developing this line here). Like in those previous works, speed of spread is determined by the preference ω which I have set to a constant. However, the crucial factor is that, unlike in the previous approaches, this factor directly derives from a cognitive model.

This merits some discussion from the Language Game perspective. These AI systems to study language and cognition typically focus on a different kinds of problems as those addressed here, in evolutionary biology, anthropology or cognitive psychology. However, I have shown that there is consistency and convergence of these approaches. The implication is that it is not only possible but desirable to use grounded agents to study some aspects of culture, as it is done with language [11]. Moreover, this work shows that there is a direct analogy between selection on variation and preferences (closing the circle with the previous traveling-wave cultural models). In the Naming Game [14], this cognitive step employs a scoring mechanisms where constructions are rewarded by a fixed amount γ upon communicative success. This score system is $\omega' \propto \omega(1-\gamma)+\gamma$. This necessarily converges to $\omega \to 1$ implying that if a construction is used very often its use is reinforced, becoming more prone to be used. The same scoring mechanism is used for every construction, therefore what determines which construction is adopted by the population is of a stochastic nature. Yet, this ensures that a population of agents converge to a common lexicon (alignment). Using the NR this system is replaced by a cognitive

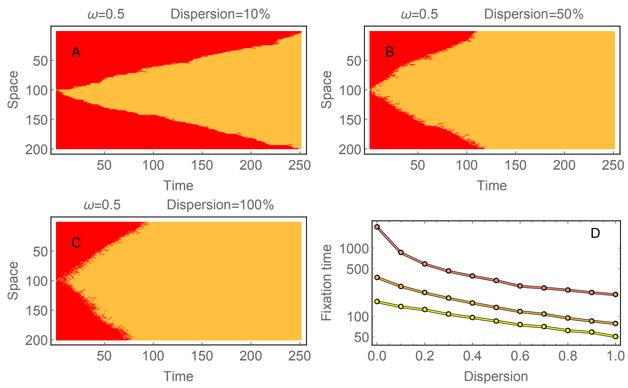

Figure 7: Spread of a novel concept σ_2 under different values of individual dispersion (A) 10%, (B) 50% and (C) 100% In this example neutral preference is assumed ($\omega = 1/2$). Otherwise as in Fig. 1. (D)Effect of dispersal probability in the fixation time if the new concept σ_2. Note the graph is in semi-log scale. Pink: $\omega = 0.1$, orange $\omega = 0.5$, yellow $\omega = 1$, $N = 100$.

version that allows each agent to score different tags (constructions) differentially, based on the properties of the networks (e.g. quality of the output or energy of the attractor) and which allow tuning of this preference (see Appendix B).

6 SUMMARY AND CONCLUSIONS

I have introduced a spatial version of the imitation game with the aim to study cultural dynamics and cultural mixability. Although I have taken a very minimal culture cased in two concepts, this has allowed understanding how (a) cognitive restrictions affect the spread of cultures and (b) how individual dispersal facilitates cultural mixability.

With this simple model I have shown a connection between cognitive factors and the speed of evolution of culture. Even though in this model said connection figures only through a parameter, ω, I remind that this simplification results from a more comprehensive AI model that takes into consideration a broader dimension of interesting factors that range from neuronal and learning models to complexity and richness of cultures. Despite the exciting nature of this complexity, a simpler and more conservative approach allows making precise quantifications and harder to make with more complex and realistic scenarios.

Naturally, it desirable to develop models for several concepts and with populations structured in demes, so that each deme has its own cultural dynamics (i.e. sub-populations, rather than only individuals in an array) and, in that sense, truly address multi-cultural dynamics. This constitutes the following steps of this work. As mentioned above, the ultimate goal is to have a richer implementation on cognitive agents with NRH architecture, so that we can address complex cultural traits instead of only two competing tags. Nevertheless, the resultspresented in this paper has demonstrated that there is a precise analogy between cultural dynamics and other models of culture and of genetic evolution, which sets an important agenda to further understand cognition, its applications in AI and the origins and evolution of the *Homo* genus.

A MATHEMATICAL MODEL

Markov model of transmission. The derivation of the mathematical model is based on properties of language games. I assume a population of N individuals spatially arranged in an array. Each individual can be in one of three states according to their knowledge, namely (σ_1), (σ_2) or (σ_1, σ_2). Interactions occur amongst pairs of individuals which can be occur in the six possible combinations of these states, namely:

$$\begin{aligned}
X_1 &= (\sigma_1) \quad , \quad (\sigma_1)\\
X_2 &= (\sigma_2) \quad , \quad (\sigma_2)\\
X_3 &= (\sigma_1) \quad , \quad (\sigma_2)\\
X_4 &= (\sigma_1) \quad , \quad (\sigma_1, \sigma_2)\\
X_5 &= (\sigma_2) \quad , \quad (\sigma_1, \sigma_2)\\
X_6 &= (\sigma_1, \sigma_2) \quad , \quad (\sigma_1, \sigma_2)
\end{aligned}$$

(order is ignored). These are the states that will be described since they represent the cultural pool of a population. The transition between each of these states for a pair of individuals represents the acquisition of a new concept by one of the agents resulting from their interaction. Despite the elaborate computational nature of the agents, it is simple to describe the space of utterances by using a Markov Chain that has a transition matrix:

$$\mathbb{M} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1-\frac{\omega}{2} & 0 & \frac{\omega}{2} \\ 0 & 0 & 0 & 0 & \frac{\omega+1}{2} & \frac{1-\omega}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5}$$

This transition matrix contains the probabilities of going from an actual state X_i to another state X_j. In a real cognitive system the value of ω is expected to change as individuals interact and, idiosyncratically, develop preferences for σ_1 or σ_2 (see Appendix B). However, in this case ω is kept constant. This transition matrix is derived from naming games that implement a Neuronal Replicator dynamics [5]. In that case ω is not constant, but is given by cognitive dynamics. Here we take ω fixed on each individual.

Boundary expansion. The expected trajectory can be predicted on the basis of these elements, assuming that the preferences are homogeneous in space. First note that if interacting neighbours have the same concepts their state remains unchanged because they can only choose to express this single one. For example, in Fig. 1 there are two regions, the yellow where neighbours are in state X_1. In the yellow region the neighbouring states will be X_6. In the matrix $\mathbb{M}$ we can see that these are absorbing states, that is, they cannot change. Therefore, changes can only happen at the interface between the two regions. Here a pair of individuals are always in state X_4. This either they remain in the same state because σ_1 is expressed (with prob. $\mathbb{M}_{4,4} = 1-\omega/2$), or σ_2 is expressed (with prob. $\mathbb{M}_{4,6} = \omega/2$) (Fig. 8A). Therefore, the interface never shrinks and it either stays at the same point, or increases by one step. This means that we only need to take into account the position of the interface which is a directed random walk with probability of advancing of $\omega/2$ (Fig. 8B). After n iterations the average number of successes is, following a binomial distribution, $n\omega/2$. Thus, the average slope after n steps is $(n\omega/2)/n = \omega/2$.

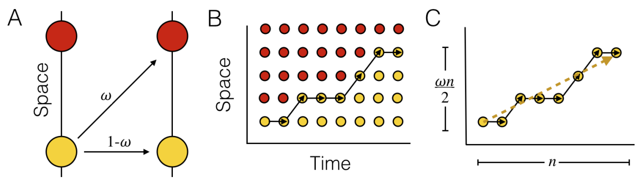

Figure 8: Tag transmission in time and space at the cultural boundary. (A) Only two transitions are allowed: First, with probability ω the tag σ_2 is transmitted to an individual (red) that only knows σ_1 increasing the boundary (yellow). Second, with probability $1-\omega$ the tag σ_1 is transmitted, but since this agent already has knowledge of σ_1 there is no change in state and the boundary does not expand. (B) Successive iterations result in a stochastic increase of the boundary. The black arrows delineate the boundary between the two cultures. (C) After n iterations, the average increase is given by $\omega n/2$ (yellow dotted arrow).

Fixation probability. if the new concept is introduced at a location x_o that is far from the boundaries, the spread is given by the line

$$x_+(t) = x_o + \frac{\omega}{2}t\ . \tag{6}$$

There is another boundary that decreases on a similar way:

$$x_-(t) = x_o - \frac{\omega}{2}t\ . \tag{7}$$

This allows calculating the frequency of concepts at each time, given by the proportion of the two regions. Calling ρ the proportion of individuals that know σ_2, $\rho(t) = (x_+(t) - x_+(t))/N$, giving $\rho(t) = \frac{\omega}{N}t$, as reported in Eq. 2.

From this last equation we compute the fixation time τ_f as the moment when the frequency of σ_2 is 100%, which is directly $\rho(\tau_f) \equiv 1 = \frac{\omega}{N}\tau_f, \Rightarrow \tau_f = N/\omega$, as indicated in Eq. 3.

However, note that this equation is assuming an underlying symmetry, which is that both interfaces cover equal areas. This is not true because the population is finite and once one of the interfaces hits a boundary (either N or 0) then said boundary stays at that value. In other words, these are absorbing boundaries for the random walk that describes the interface. The time at which each interface is expected to hit the boundaries are

$$\tau_+ = \frac{2}{\omega}(N - x_o) \qquad \text{and} \qquad \tau_- = \frac{2}{\omega}x_o$$

and because fixation requires that both interfaces have been absorbed it implies that the fixation time is the largest of both, or

$$\tau_f = \max\{\tau_-, \tau_+\} = \frac{2}{\omega}\max\{x_o, N - x_o\}\ , \tag{8}$$

which implies Eq. 3 if the new concept is introduced at the centre of the space. Moreover, for any $0 < N < \infty$ we can bound below

$$\tau_f \geq \frac{2}{\omega}\min_{x_o \in [0,N]}\max\{x_o, N - x_o\} = \frac{N}{\omega}$$

and, similarly, from above:

$$\tau_f \leq \frac{2}{\omega}\max_{x_o \in [0,N]}\max\{x_o, N - x_o\} = \frac{2N}{\omega}$$

Both inequalities together imply Eq. 4.

B PREFERENCES AND THE NEURONAL REPLICATOR HYPOTHESIS.

Language Games in AI [11] use a specific scoring system for agents to determine which utterances or constructions to employ when communicating. These are ad hoc updates that reward constructions that are used more often, but otherwise make no use of any cognitive property. I have taken a different stance for this, coupling the reward system to cognitive aspects. In short, the update of the weight for a given construction is implemented through a Bayesian learning scheme, namely,

$$\omega_i' \equiv \Pr[\sigma_i|\mathcal{V}] = \omega_i \frac{\mathcal{L}(\sigma_i)}{\sum_j \omega_j \mathcal{L}(\sigma_j)} \tag{9}$$

where $\omega_i \equiv \Pr[\omega_i]$, is the weight before the update and $\mathcal{L}(\sigma_i)$ is the likelihood of choosing σ_i amongst other possibilities. This Bayesian learning is itself a closer step to cognition [7, 13]. However, the real twist comes by making $\mathcal{L}(\sigma_i)$ a function of cognitive parameters.

For instance for two tags or utterances $\omega_i = \omega^k(1-\omega)^{1-k}$ where $k = 0$ if σ_1 is chosen and $k = 1$ if σ_2 is chosen. We are free to use any $\mathcal{L}(\sigma_i)$ we want, as for example a beta distribution of the form

$$\mathcal{L} \propto \omega^{\alpha+S\Delta V_1-1}(1-\omega)^{\beta+S\Delta V_2-1},$$

where $\Delta V_1 = -\delta V_2 = V_1 - V_2$ is the difference in the energy of the attractors and S is a parameter tuning the initial bias. This leads to an update rule of the form

$$\omega' = \frac{\alpha + S\Delta V_1 + k}{\alpha + \beta + 1}. \tag{10}$$

As the attractors learn, $\Delta V_1 \to 0$ (in average), so the NR converge to a simple Bayesian scheme. However at early stages of learning $S\Delta V_1$ introduce a bias toward the cognate concepts. Under this scheme, the preferences can converge to any value $0 \leq \omega \leq 1$, with the outcomes probabilistically depend on S. Thus, except for an initial period of learning, the games proceed as if ω were fixed in each agent.

C SIMULATIONS.

The simulations are numerical realisations based on the model above and proceed executing the following steps:

(1) Assign a preference and an initial state (σ_1 or σ_2) to each individual.
(2) Draw a random permutation Ω of $1 \ldots N$ which denotes the order in which interactions occur.
(3) Compute the outcome of each interaction. For this, note that a focal pair i, i.e. individuals at positions $x = \Omega_i, \Omega_i + 1$ unambiguously determine the state X; from this state compute using $\mathbb{M}$ the new state X' and update the individual states.
(4) If dispersion is implemented, after all interactions occurred draw another random permutation Ω' of $1 \ldots N-1$ and swap sequentially individuals at positions $x = \Omega_i', \Omega_i' + 1$. Depending on the model preferences are also swapped or not.
(5) Go back to Step 2 for T times or until there is no further population change.

REFERENCES

[1] Kenichi Aoki. 1987. Gene-culture waves of advance. *J. Math. Biol.* 25, 5 (1987), 453–464. DOI: http://dx.doi.org/10.1007/BF00276192

[2] Luigi L. Cavalli-Sforza and Marcus W. Feldman. 1981. *Cultural transmission and evolution: a quantitative approach.* Princeton University Press, Princeton.

[3] James F. Crow and Motoo Kimura. 1970. *An Introduction to Population Genetics Theory.* Harper & Row, New York and London.

[4] Harold P. de Vladar, Anna Fedor, András Szilágyi, István Zachar, and Eörs Szathmáry. 2016. An Attractor Network-Based Model with Darwinian Dynamics. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion.* ACM, 1049–1052.

[5] Harold P. de Vladar, Emilia García-Casademont, and Luc Steels. In preparation. Imitation games on a population of cognitive agents with a Darwinian neurodynamics architecture. (In preparation).

[6] Chrisantha Fernando, Eörs Szathmáry, and Phil Husbands. 2012. Selectionist and evolutionary approaches to brain function: a critical appraisal. *Front. Comput. Neurosci.* 6 (2012), 24.

[7] Karl Friston. 2010. The free-energy principle:a unified brain theory? *Nature Rev.* 11, 2 (Jan. 2010), 127–138.

[8] Alfred L. Kroeber and Clyde Kluckhohn. 1952. Culture: A critical review of concepts and definitions.

[9] Renato I. Rosaldo. 2006. Foreword: defining culture. In *Redefining Culture: Perspectives Across the Disciplines*, John R. Baldwin, Sandra L. Faulkner, Michael L. Hecht, and Sheryl L. Lindsley (Eds.). Lawrence Erlbaum Associates, New Jersey.

[10] Luc Steels. 2011. Modeling the cultural evolution of language. *Phys. Life Rev.* 8, 4 (Dec. 2011), 339–356.

[11] Luc Steels. 2015. *The Talking Heads experiment: Origins of words and meanings.* Language Science Press, Berlin.

[12] András Szilágyi, István Zachar, Anna Fedor, Harold P. de Vladar, and Eörs Szathmáry. 2016. Breeding novel solutions in the brain: a model of Darwinian neurodynamics. *F1000Research* 5 (2016), 2416.

[13] Tomer D. Ullman, Noah D. Goodman, and Joshua B. Tenenbaum. 2012. Theory learning as stochastic search in the language of thought. *Cogn. Dev.* 27, 4 (2012), 455–480.

[14] Remi Van Trijp and Luc Steels. 2012. Multilevel alignment maintains language systematicity. *Adv. Complex Syst.* 15, 03n04 (2012), 1250039.